

Mesoscopic Physics of Quantum Systems and Neural Networks

Von der Fakultät für Physik und Geowissenschaften

der Universität Leipzig

genehmigte

DISSERTATION

zur Erlangung des akademischen Grades

Doctor rerum naturalium

Dr. rer. nat.

vorgelegt

von M. Sc. Matthias Thamm

geboren am 17.08.1994 in Zwenkau

Gutachter:	Prof. Dr. Bernd Rosenow	(Universität Leipzig)
	Prof. Dr. Evert van Nieuwenburg	(Leiden University)

Tag der Verleihung: 25.09.2023

Contents

1	Introduction	7
1.1	Random matrix theory	9
1.1.1	Applications of random matrix theory	9
1.1.2	Gaussian orthogonal ensemble and Wishart matrices	10
1.1.3	Wigner surmise and level number variance	11
1.1.4	Marcenko-Pastur and Porter-Thomas distributions	14
1.2	Neural networks and the CMA-ES machine learning algorithm	18
1.2.1	Neural networks	19
1.2.2	Applications of random matrix theory to neural networks	22
1.2.3	The CMA-ES machine learning algorithm	24
1.3	One-dimensional Majorana wires	26
1.3.1	Majorana zero modes in topological superconductor nanowires	26
1.3.2	Majorana zero modes for quantum computation	30
1.3.3	Challenges for realization and detection of Majorana zero modes	31
1.3.4	Amplitude of coherent transmission	33
1.4	Entanglement entropy in one dimensional fermion chains	35
1.4.1	Particle entanglement entropy	36
1.4.2	The J-V model	37
1.4.3	Luttinger liquid theory and bosonization	39
1.5	Structure of this thesis and publications	46
2	Majorana zero modes in topological superconductors	47
2.1	Topological superconductors: BdG Hamiltonian, topological invariants, and zero modes	47
2.2	Realization of Majorana zero modes	56
2.2.1	Spin-polarized p-wave superconductor and the Kitaev chain	56
2.2.2	Majorana zero modes in Rashba wires	62
2.2.3	Finite energy Majorana modes	78
2.2.4	Robust Andreev bound states	81
2.2.5	Two-dimensional Majorana wires	84
2.3	Scattering matrix formalism for coherent transport in Coulomb-blockade	96
2.3.1	Aharonov-Bohm interferometer setup	97
2.3.2	Coulomb-blockade	100
2.3.3	Scattering matrix approach at finite temperatures	104
2.3.4	Microscopic model for dot-lead couplings	113
2.3.5	Extracting Weidenmüller formula parameters from KWANT	115
2.4	Coherent transmission amplitude through a Coulomb-blockaded nanowire	117
2.4.1	Magnetic field and wire length dependence of the transmission amplitude	118
2.4.2	Disorder in the wire	121
2.4.3	Microscopic model and Andreev zero modes	122
2.4.4	Comparison with direct conductance measurement	125
2.4.5	Comparison to approximate analytical solution and connection to experiment	125
2.4.6	Conclusion	127

2.5	Using machine learning to cancel disorder effects	128
2.5.1	Setup and CMA-ES algorithm	128
2.5.2	Optimization results for one dimensional wires	134
2.5.3	Optimization with noisy signal	137
2.5.4	Optimization of average voltage	137
2.5.5	Choice of the metric	139
2.5.6	Number of levels included in the scattering calculation	140
2.5.7	Optimization for two-dimensional wires	142
2.5.8	Conclusion	143
3	Random matrix theory of deep neural network weight matrices	145
3.1	Neural networks	145
3.1.1	Network architectures and performance for image classification	146
3.1.2	Details on preprocessing and training	149
3.1.3	Singular value decomposition of weights	154
3.2	Random matrix theory as a zero-information hypothesis	156
3.2.1	Interpretation of Kolmogorov-Smirnov tests	156
3.2.2	Universal level spacing statistics and level number variance	158
3.2.3	Singular values and vectors: Marcenko-Pastur and Porter-Thomas distribution	161
3.2.4	Random matrix analysis in different learning regimes	165
3.2.5	Tails of weight spectra: Hill estimators and power-law fits	169
3.3	Boundary between noise and information applied to filtering neural network weight matrices	173
3.3.1	Boundary between noise and information in weight spectra	173
3.3.2	Filtering of neural network weight matrices	177
3.4	Conclusions	182
4	Entanglement entropy in one dimensional quantum chains of interacting fermions	185
4.1	Schmidt decomposition	186
4.2	The J-V model	187
4.2.1	Phases from mapping onto a spinful XXZ-chain	188
4.2.2	A simple example: two fermions on four sites	191
4.3	Mapping onto an effective low-energy Luttinger liquid model	200
4.3.1	One-particle entanglement entropy in the equilibrium case	201
4.3.2	Time evolution of the one-body density matrix after a quantum quench	206
4.3.3	Steady state after quench	209
4.4	Numerical computations with exact diagonalization	210
4.4.1	Symmetries	211
4.4.2	Integer fermion basis	212
4.4.3	Structure matrix	215
4.5	Numerical computations with DMRG and ITensors.jl	216
4.5.1	Advantages of tensor networks and stabilizing DMRG	217
4.5.2	Time evolution using GPUs	219
4.5.3	Higher order particle entanglement entropies	220
4.5.4	Benchmarking DMRG with exact diagonalization	220
4.6	Numerical results and comparison to analytical bosonization results	221
4.6.1	Equilibrium case	221

4.6.2	Interaction quantum quench	223
4.7	Conclusions	228
5	Summary	229
	Bibliography	232
	Original publications and author contributions	265
	Acknowledgement	267
	Appendices	268
A.1	Mean field approximation for the superconductor Hamiltonian	269
A.2	Zero energy solutions at a domain wall in a spin-polarized one dimensional p-wave superconductor	270
A.3	Eigenstates of the normal-conducting Rashba wire	273
A.4	Finite Majorana energy	275
A.5	CMA-ES algorithm	277
B.1	Spectra and singular vectors for more networks	280
C.1	Consistency check of the refermionized field operator	281
C.2	Finite size corrections of the entropy in the clustered solid limit	281
	Bibliographische Beschreibung	283
	Erklärung	284

1. Introduction

Many systems in everyday life are macroscopic, such that we can describe them in terms of coarse grained, averaged quantities. By contrast, quantum effects determine the properties of objects with microscopic dimensions, such as electrons or photons. Mesoscopic physics takes place in the intermediate region between macroscopic and microscopic scales, in systems with a large number of quantum objects, as for example in nano-scale electronic devices. Here, quantum properties and fluctuations play a crucial role, but the individual microscopic degrees of freedom often cannot be individually traced due to the huge number of interacting constituents. In the 1950s, E. Wigner also faced this problem [1–5] when trying to theoretically describe the energy-level spacings of compound nuclei. In the absence of a coherent theoretical description – despite the existence of numerous experimental data [6, 7] – he proposed the use of a completely random Hamiltonian as a zero-information hypothesis for comparison with experiments [1–3]. The striking success of this description [7–12] paved the way for the application of random matrix theory (RMT) – some of whose statistical predictions had been developed earlier [13–16] – in numerous fields [7, 11, 12, 17]. Besides the aforementioned application to many-body systems in nuclear physics [11, 18–21], RMT proved fruitful for describing quantum chaos [22, 23], disordered mesoscopic systems [11], transport through nanowires [24, 25], in field theoretical descriptions of interacting fermions [11, 26–29], and even for estimating correlations between assets in the stock market [30–34].

Recently, it has also been proposed to apply RMT to neural networks [35–37], where the situation is similarly complicated: Deep neural networks contain huge sets of parameters whose fitting to correctly predict a training dataset is not easily traceable [37], and a cohesive theory and deep understanding of the training process is elusive [37, 38]. Nevertheless, these networks have an excellent generalization ability beyond the training dataset while also being able to memorize large amounts of random data [39]. A further problem is obtaining clean training datasets, since noise in the labels is often inevitable [40, 41].

Here, we consider RMT as a zero-information hypothesis to study the weight matrices of deep neural networks, which store the parameters of the networks: Since these weights are initialized as random matrices before training, they initially precisely follow the predictions as random matrices. After training, we can then use deviations from the predictions to locate the information stored in the weights during training. We show that the bulk of the weights remains random after training, that the relevant information is encoded in a small portion of the spectrum of the weight matrices, and that this information, for suitable training, is also stored separately from the information about incorrect, noisy labels. This further allows us to propose a filtering algorithm that can partially remove the influences of the noisy labels from the already trained weights and thus increase the generalization accuracy of the networks.

Machine learning techniques, such as neural networks, have numerous applications [42–47] from image and pattern recognition [44] and protein folding [48] to physics applications such as quantum state tomography [49, 50] and localizing phase transitions [51–55]. Moreover, in recent years, increasingly complex mesoscopic quantum systems have been proposed and also implemented [56–60], which requires very sophisticated tuning. Automating this tuning is profitable and in some cases even necessary [61–63], and machine learning approaches have proven to be particularly fruitful and flexible for this purpose [60, 62–67]. Especially for the implementation of large scale quantum computation [68–71], tuning of parameters and gates is required, and numerous automations in quantum dot based qubits have been proposed [61–63, 65, 66, 72–78]. A

popular foundation for scalable qubit architectures is based on Majorana zero modes (MZMs) in topological superconductors [56, 59, 79–84], whose advantages lie in the topologically protected encoding of quantum information in the strongly non-local states as well as their manipulation by anyonic braiding [79–82, 85]. These MZMs have been proposed in semiconductor-superconductor hybrid systems [79, 86–89], and many of their predicted signatures have also been found experimentally: For example, zero-bias conductance peaks [90–93], the fractional Josephson effect [94], and the suppression of the even-odd spacing difference between conductance resonances in hybrid wires with Coulomb blockade [95]. For clean wires, it has been theoretically shown that specifically chosen magnetic [96–99] and potential profiles [97] can make MZMs more robust, and also the geometry of Josephson junctions has been optimized, with the aim of obtaining a large topological gap [100]. Recently, however, it has turned out that the implementation is not as simple as initially thought [101]: The main problem in these systems is the almost inevitable presence of strong disorder [102–106], which can not only destroy the topological phase and thus the MZMs [107] but can even mimic most of the previously mentioned signatures in the topologically trivial regime entirely without requiring any MZMs [102, 104, 108–111]. As a result, a large portion of samples currently have to be discarded in the fabrication of Majorana devices: In a recent study conducted by the Microsoft Quantum group, it was found that no more than 20% of their elaborately manufactured and tested Majorana devices have a chance for “a high probability [...] of a topological phase hosting Majorana zero modes” [106].

We here propose to bring an array of gates in proximity to a Majorana wire and optimize the gate voltages using the CMA-ES machine learning algorithm to compensate for disorder while making MZMs more robust. For this purpose, we need a suitable metric that achieves the desired effects when it is maximized by the algorithm, which itself does not require any other system-specific information. One major problem is that many signatures can be mimicked by trivial Andreev bound states, which can be pinned at zero energy, so called pseudo-MZMs [111–128], such that the optimization of metrics derived from these signatures does not favor MZMs. A suitable signature should therefore be based on the non-locality of the MZMs, for which transport through the Majorana wire in Coulomb blockade, referred to as electron teleportation [129], is a suitable choice.

In a recent experiment, Whiticar et al. [130] considered the coherent transmission amplitude through a Majorana wire in Coulomb blockade, which can be determined by embedding the wire in one arm of an electron interferometer and measuring the oscillations of the current with the magnetic flux through the interferometer loop [130]. Using an adapted scattering matrix formalism and combining numerical and analytical calculations, we explain the experimentally observed occurrence of a maximum of the coherent transmission amplitude as a function of an external Zeeman field soon after entering the topological phase as a direct consequence of the localization properties of the MZMs [131] – a signature that also allows distinguishing MZMs from pseudo-MZMs [131, 132]. This makes the coherent transmission amplitude ideal for optimization with the CMA-ES algorithm: We show that the machine learning algorithm is capable of very efficiently improving topological signatures, making MZMs more stable, and even learning disorder profiles and fully restoring MZMs that have been completely destroyed by strong disorder by using the optimized gate voltages. Such optimization can therefore potentially increase the yield in the production of Majorana devices significantly.

To implement large scale quantum computing, in addition to a scalable qubit architecture, a deep understanding and detailed control of quantum entanglement are necessary [133, 134], which we further study here in the context of one-dimensional interacting fermions. If one prepares a quantum system in a pure state and performs an interaction quantum quench [135],

that is, a sudden change in the interaction strength between the particles, the entanglement entropy between two spatial regions of the system [136, 137] describes how information about the quench propagates through the system: Immediately after the quench, this spatial entanglement entropy increases linearly and saturates once the information has propagated through the whole system, thereby playing the role of a thermal entropy with the help of which local observables can be determined [138–141]. Instead of the well-understood spatial entanglement entropy, we consider here entanglement in an alternative bipartition of the system into subgroups of n particles and the remaining $N - n$ particles [142–145]. This particle entanglement entropy differs from the spatial entanglement entropy in that it is sensitive to particle statistics and interactions to first order [146–157] and directly linked to the reduced n -particle density matrix and the n -point correlation matrix [145]. The latter also opens ways to measure the particle entanglement entropy in experiments [158–160]. Moreover, the equivalence of particle and spatial entanglement entropy in the thermodynamic limit has been demonstrated for long times after an interaction quench [161].

For accurate finite-size scaling to the thermodynamic limit, we perform large-scale, state-of-the-art numerical computations of the one-particle entanglement entropy using both exact diagonalization and approximative methods. This also allows us to make a connection to a field theoretic bosonization calculation in the Luttinger liquid phase [162–164], where the one-dimensional problem is made tractable by linearizing the dispersion around the Fermi level and considering low-energy excitations due to density fluctuations against an average density background [163]. By comparing with the numerical results, we are able to determine an interaction cutoff of the field theory such that the analytical results can make accurate predictions for weak interactions, going far beyond a mere determination of universal scale coefficients.

In this chapter, we provide an introduction and outline of the topics covered in this dissertation, as well as a brief description of the publications that have resulted from the research presented in the following chapters.

1.1 Random matrix theory

1.1.1 Applications of random matrix theory

The application of RMT is promising for systems with dynamics that are too complicated to be traced, for instance because of disorder, chaotic behavior, or due to a complex, random initial state. In many cases, it is then possible to derive a probability distribution that describes a quantity of interest, such as a Hamiltonian, a scattering matrix, or a parameter matrix based on symmetries or the shape of the initial state. Application of RMT for these distributions allows us to make predictions for observables based on the universal properties of the underlying symmetries. Systematic deviations from the RMT predictions, on the other hand, reveal system-specific properties that are not random.

The first successes of RMT in physics were achieved in the description of the energy spectra of complex nuclei and molecules [7, 11, 12]: In the 1950s, neutron scattering experiments enabled the measurement of energy spectra of composite nuclei with good resolution [6, 7], but a theoretical description of the observed results was elusive due to the large number of interacting constituents. Here, RMT brought the first breakthrough [1, 11] by describing the Hamiltonian with a random matrix from the Gaussian orthogonal ensemble (GOE) - random matrices with normally distributed entries that have the form of time-reversal invariant Hamiltonians. The distributions of level spacing and level broadening determined from the statistics of the random

Hamiltonian showed excellent agreement with the observed measurements [1, 11]. Later in the 1980s, improved experimental techniques enabled the creation of large data sets, so that the description of the systems with the GOE could be statistically justified in retrospect [8, 11, 165].

Subsequently, RMT has been successfully applied in a variety of fields [11, 12, 166, 167]. It gained great popularity in the description of quantum transport [167]: Here, in addition to random Hamiltonians, random scattering matrices were also considered, from which universal predictions about transport properties can be obtained [167]. For example, for scattering from quantum dots, it was shown that the Hamiltonian for a disordered metal grain can be described by the Wigner-Dyson statistic [167, 168], which allowed determining a random ensemble for the scattering matrix [166]. These considerations have further been extended to wires with strong disorder [166, 169, 170]. Here the applications range from descriptions of normal-superconductor contacts and chaotic Josephson junctions [166], to the calculation of universal properties of disordered superconductors [11, 167]. In this context, RMT has even been applied to topological superconductors with Majorana zero modes [167], for which generalized random matrix ensembles for Hamiltonians and scattering matrices have been developed that also take into account particle-hole symmetry [167]. Here it is possible to distinguish between symmetry classes and even topological phases by comparison with predictions of RMT [167].

In addition, RMT has also been used to study the scaling of entanglement entropy at critical points in quantum chains [171, 172], and there are numerous applications outside of physics, e.g., for the description of financial markets [30, 32–34, 173], where RMT has been used in the analysis of return statistics and in the optimization of portfolios and investment strategies. In fact, there are so many other applications of RMT – both in physics [11, 12, 166, 167] and outside [174–178] – that it is impossible within the scope of this thesis to provide a fully comprehensive list. We explore another use case in detail in Ch. 3, where we use RMT for the analysis of neural networks. We postpone the discussion of earlier applications of RMT in this field to Sec. 1.2.2, as we first want to give a brief introduction to the main predictions of RMT and to neural networks.

1.1.2 Gaussian orthogonal ensemble and Wishart matrices

Because the first applications of random matrix theory in physics were developed for the description of time-reversal invariant systems whose Hamiltonians are real symmetric matrices, the so-called Gaussian Orthogonal Ensemble (GOE) is of particular importance [11]. A matrix A , from the GOE universality class satisfies the following conditions [12]:

- (i) The elements A_{ij} are real.
- (ii) The matrix is symmetric, $A_{ij} = A_{ji}$.
- (iii) The diagonal elements A_{ii} are normally distributed with mean 0 and variance $2\sigma^2$.
- (iv) The off-diagonal entries $A_{ij}, i > j$ are normally distributed with mean 0 and variance σ^2 .

Therefore, the probability distribution of the GOE matrices is given by [11]

$$P_{\text{GOE}}(A) \propto \exp \left[-\frac{1}{4\sigma^2} \text{tr}(A^2) \right], \quad (1.1)$$

and to generate a matrix from the GOE, one can first draw a random matrix B with i.i.d. normal entries with mean zero and variance σ^2 and construct the GOE matrix as $A = (B + B^T)/2$.

While numerous properties have been computed for matrices of the GOE [7, 11, 12], the matrices that are used as weights in neural networks, which we study in Ch. 3, are in general neither symmetric nor quadratic. However, for a rectangular $n \times m$ matrix (w.l.o.g. let $n \geq m$) with i.i.d. Gaussian entries of zero mean and standard deviation σ , we can consider the so-called Wishart matrix $X = W^T W$ whose entries are symmetric,

$$X_{ij} = \sum_{k=1}^n W_{ki} W_{kj} = X_{ji} , \quad (1.2)$$

have $\text{mean}(X_{ij}) = n\sigma^2\delta_{ij}$ and variance $\text{var}(X_{ij}) = \delta_{ij}2\sigma^4n + (1 - \delta_{ij})\sigma^4n$. In the limit $n \rightarrow \infty$, the central limit theorem thus ensures that the matrix $X - \mathbb{1}n\sigma^2$ belongs to the GOE with variance $n\sigma^4$. It is therefore not surprising that many predictions of the GOE apply to Wishart matrices as well. In particular, the eigenvalues λ_i of X coincide with the squares of the singular values $\lambda_i = \nu_i^2$ of W , which is why later we primarily consider the singular value decomposition $W = USV^T$. Here U and V are orthogonal matrices and S is the diagonal matrix of the singular values. In fact, the limit $n \rightarrow \infty$ is not strictly necessary to find agreement between empirical spectra of random Wishart matrices and GOE predictions [179]. Below, we consider examples for comparing Wishart matrices of different sizes with the GOE predictions and find that $n \sim 300$ already yields excellent correspondence.

In the following two sections, we sketch the derivation of RMT predictions, first those of universal nature that depend only on the ensemble, and those with non-universal properties that depend on the individual realizations, e.g., the chosen standard deviation of the matrix entries.

1.1.3 Wigner surmise and level number variance

As a first universal signature, we consider the distribution of level spacings in the spectra of random GOE matrices, i.e. the distribution of the differences of neighboring eigenvalues, $|\lambda_i - \lambda_{i+1}|$. For many different types of matrices, complicated analytic derivations and results exist in the literature [11, 180–182]. However, it turns out that the simple model proposed by E. Wigner to describe the energy level spacings in composite nuclei [1] provides a surprisingly accurate and robust approximation for GOE matrices of arbitrary size [182]. Here, Wigner’s derivation is based on a simple two-level system [183]: consider the general 2×2 GOE matrix¹

$$X = \begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix} , \quad (1.3)$$

where the entries are drawn from a normal distribution with zero mean, variances of the diagonal entries given by $\text{var}(x_1) = \text{var}(x_2) = 2$, and variance of the off-diagonal elements given by $\text{var}(x_3) = 1$. The corresponding probability density is therefore given by

$$P(x_1, x_2, x_3) = \frac{1}{2(2\pi)^{3/2}} e^{-\frac{1}{4}(x_1^2 + x_2^2 + 2x_3^2)} . \quad (1.4)$$

¹For the derivation of the probability distribution of the level spacings, we follow an exercise problem from the Leiden University available at <https://www.lorentz.leidenuniv.nl/RMT/RMTproblems.pdf>.

We can diagonalize the real, symmetric matrix using the Bogoliubov transformation

$$\begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \quad (1.5)$$

$$= \begin{pmatrix} \lambda_1 \cos^2 \varphi + \lambda_2 \sin^2 \varphi & \cos \varphi \sin \varphi (\lambda_2 - \lambda_1) \\ \cos \varphi \sin \varphi (\lambda_2 - \lambda_1) & \lambda_2 \cos^2 \varphi + \lambda_1 \sin^2 \varphi \end{pmatrix}, \quad (1.6)$$

where λ_1 and λ_2 (w.l.o.g. $\lambda_1 > \lambda_2$) are the eigenvalues. The Jacobian for this transformation is given by

$$J(\lambda_1, \lambda_2, \varphi) = \begin{pmatrix} \partial_{\lambda_1} x_1 & \partial_{\lambda_2} x_1 & \partial_{\varphi} x_1 \\ \partial_{\lambda_1} x_2 & \partial_{\lambda_2} x_2 & \partial_{\varphi} x_2 \\ \partial_{\lambda_1} x_3 & \partial_{\lambda_2} x_3 & \partial_{\varphi} x_3 \end{pmatrix} = \begin{pmatrix} \cos^2 \varphi & \sin^2 \varphi & 2 \cos \varphi \sin \varphi (\lambda_2 - \lambda_1) \\ \sin^2 \varphi & \cos^2 \varphi & -2 \cos \varphi \sin \varphi (\lambda_2 - \lambda_1) \\ -\cos \varphi \sin \varphi & \cos \varphi \sin \varphi & \cos(2\varphi)(\lambda_2 - \lambda_1) \end{pmatrix} \quad (1.7)$$

with determinant $|\det J(\lambda_1, \lambda_2, \varphi)| = \lambda_1 - \lambda_2$. Fortunately, the determinant is already given by the level spacings, and transforming the coordinates of the probability density from (x_1, x_2, x_3) to $(\lambda_1, \lambda_2, \varphi)$ yields

$$P(\lambda_1, \lambda_2, \varphi) = P(x_1, x_2, x_3) |\det J| = \frac{\lambda_1 - \lambda_2}{2(2\pi)^{3/2}} e^{-\frac{1}{4}(x_1^2 + x_2^2 + 2x_3^2)} = \frac{\lambda_1 - \lambda_2}{2(2\pi)^{3/2}} e^{-\frac{1}{4}(\lambda_1^2 + \lambda_2^2)}. \quad (1.8)$$

To further determine the distribution of the level spacings, we first transform to the relative and center of mass coordinates $d = \lambda_1 - \lambda_2$, $r = (\lambda_1 + \lambda_2)/2$ and find $\lambda_1^2 + \lambda_2^2 = (4r^2 + d^2)/2$, such that the probability density is given by

$$P(r, d, \varphi) = \frac{d}{2(2\pi)^{3/2}} \exp \left[-\frac{1}{8}(4r^2 + d^2) \right]. \quad (1.9)$$

To obtain the probability density of the level spacing d alone, we integrate out the coordinates r and φ :

$$P(d) = \int_0^\infty dr \int_0^{2\pi} d\varphi P(r, d, \varphi) = \frac{d}{4} e^{-d^2/8}. \quad (1.10)$$

Now, to get a truly universal property, we consider the distribution of the standardized level spacings s , computed from the so-called unfolded spectrum, for which the mean level density is normalized to unity. For the simple two-level system, this unfolding is trivial, since we only need to divide by the mean level spacing

$$\langle d \rangle = \int_0^\infty dP(d) dd = \sqrt{2\pi}. \quad (1.11)$$

The resulting distribution, $P_{\text{WS}}(s) = \langle d \rangle P(s \langle d \rangle)$ [183],

$$P_{\text{WS}}(s) = \frac{s\pi}{2} \exp \left[-\frac{\pi s^2}{4} \right], \quad (1.12)$$

is known as the Wigner surmise, whose validity – at least as an excellent approximation [182] – can be shown numerically even for large matrices that are far from the simple two-level system.

As mentioned earlier, in order to apply RMT to the weight matrices of neural networks, we need predictions for the singular values of non-symmetric matrices. Thus, the question about

the distribution of the singular value distances of a general 2×2 matrix with i.i.d. normally distributed random numbers arises. For such a matrix W , we consider the singular value decomposition

$$W = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}^T \begin{pmatrix} \nu_1 & 0 \\ 0 & \nu_2 \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad (1.13)$$

where the singular values $\nu_1 > \nu_2$ can always be chosen to be positive by parameterizing the symmetric matrices through $\phi \in [0, 2\pi), \theta \in [0, \pi)$. For this case, the corresponding probability distribution of the matrix entries is given by

$$P(x_1, x_2, x_3, x_4) = \frac{1}{4\pi^2} e^{-(x_1^2 + x_2^2 + x_3^2 + x_4^2)/2}, \quad (1.14)$$

where we obtain the Jacobi determinant as $|\det J| = (\nu_1 - \nu_2)(\nu_1 + \nu_2)$ after a few algebraic steps. Thus, we find the probability density of the singular values as

$$P(\nu_1, \nu_2, \phi, \theta) = \frac{(\nu_1 - \nu_2)(\nu_1 + \nu_2)}{4\pi^2} e^{-(\nu_1^2 + \nu_2^2)/2} \quad (1.15)$$

$$P(d, r, \phi, \theta) = \frac{2dr}{4\pi^2} e^{-(d^2 + 4r^2)/4}, \quad (1.16)$$

where $d = \nu_1 - \nu_2$, $r = (\nu_1 + \nu_2)/2$ are again relative coordinates. Integrating out the parameters d, ϕ , and θ yields a similar distribution as obtained before for the eigenvalue spacings of GOE matrices:

$$P(d) = \frac{d}{2} e^{-d^2/4}, \quad (1.17)$$

such that the distribution of the unfolded singular values $s/\langle d \rangle = s/\sqrt{\pi}$ is again given by the Wigner surmise:

$$P_{\text{WS}}(s) = \frac{s\pi}{2} \exp \left[-\frac{\pi s^2}{4} \right]. \quad (1.18)$$

In addition to the observation that the Wigner surmise is also good at describing large matrices, the numerical considerations in Ch. 3 show that the level spacing statistic is also robust to small changes in parts of the singular value spectrum – as long as the bulk of the spectrum is equivalent to that of a random GOE matrix.

Another universal property based on the unfolded spectrum is the level number variance, i.e. the variance Σ^2 in the number of unfolded eigenvalues in intervals of length ℓ . Here, unfolding refers to normalizing the mean level density to unity such that for a given spectrum of n singular values ν_i described by the probability density $P(\nu)$, the unfolded spectrum is given by $\xi_i = n \int_{-\infty}^{\nu_i} P(x) dx$ [7, 11, 12, 19, 33, 182]. It has been shown [11, 21, 182] that for the singular values of an i.i.d. Gaussian matrix with entries with zero mean and also for the eigenvalues of GOE matrices, the level number variance grows logarithmically, i.e. $\Sigma^2(\ell) \propto \ln(2\pi\ell) + \text{const.}$.

These universal properties are independent of the realization of the matrices and are fully determined by the underlying universality class, i.e. by the symmetry of the random matrix ensemble². While a GOE matrix has a logarithmically growing level number variance, for a

²In fact from Eq. (1.1) one can see that the weight function of the GOE is invariant under orthogonal transformations $A \rightarrow \tilde{A} = UAU^\dagger$ with $UU^\dagger = U^\dagger U = \mathbb{1}$, where the cyclicity of the trace ensures that $\text{tr}(A^n) = \text{tr}(\tilde{A}^n)$.

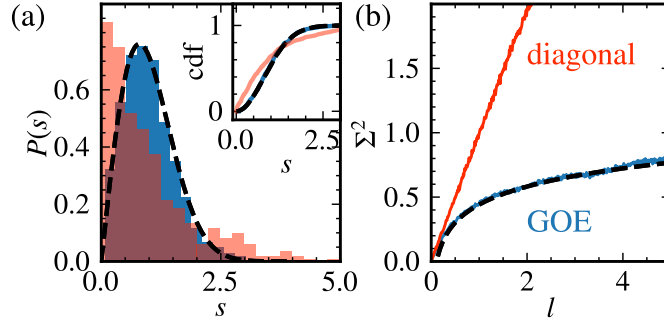


Figure 1.1: Universal properties of random matrices. We show theory curves for the GOE universality class (dashed, black) and empirical results for 512×512 matrices for unfolded singular values of a matrix with i.i.d. normal entries (approximately belonging to the GOE ensemble) in blue and for an i.i.d. normal diagonal matrix in red. (a) The level spacing distribution $P(s)$ of unfolded singular values for the GOE matrix follows the Wigner surmise $\propto se^{-\pi s^2/4}$, while the diagonal matrix has a qualitatively different level spacing distribution. (b) The level number variance $\Sigma^2(\ell)$ of the GOE spectrum only grows logarithmically, while the diagonal matrix has a linearly increasing level number variance of the unfolded singular values.

diagonal matrix with i.i.d. normal entries the growth of the level number variance is linear in ℓ [33]. Similarly, while the unfolded eigenvalue spacings of a GOE matrix follows the Wigner surmise, $\propto se^{-\pi s^2/4}$, the diagonal random matrix has a qualitatively different level spacing distribution. This is shown for realizations of random 512×512 matrices in Fig. 1.1. In Chapter 3, we therefore use the universal properties to decide whether the bulk of a random matrix is consistent with a considered random matrix ensemble, where we compare the weight matrices of neural networks before and after training the network.

To allow distinguishing different realizations within the same ensemble, e.g. matrices with different variances of the entries, we will study some non-universal RMT predictions in the following as well.

1.1.4 Marcenko-Pastur and Porter-Thomas distributions

The singular value decomposition of a real $n \times m$ matrix W is defined as $W = USV^T$, where U and V are orthogonal matrices, S is a diagonal matrix containing the singular values ν_i , and we assume w.l.o.g. that $n \geq m$. This decomposition exists for any real matrix and is also unique if one chooses the ν_i to be non-negative. Then, the number of non-zero singular values defines the rank of the matrix, and one says that W has full rank if this number is equal to m . For the study of weight matrices of neural networks, the singular value decomposition is advantageous as it decomposes the matrices into directions – the vectors in U and V – and scale factors – the singular values – allowing us to (i) identify relevant directions, (ii) compress matrices, and (iii) manipulate the spectrum, reconstruct the weight matrix, and use it in the network again.

Before training, the weight matrices are initialized with i.i.d. random numbers of zero mean and variance σ^2 such that their singular value spectra follow the Marcenko-Pastur distribution [184]. In the following, we sketch the derivation of this distribution closely following Sengupta et al. [185]. We will skip a few details during the derivation as the purpose of this section is to give an overview of the requirements and the limits that need to be taken for the Marcenko-Pastur law to be valid. Ultimately, we would like to obtain an expression for the probability

density $P_{\text{MP}}(\nu)$ of the singular values ν of W . It is convenient to first obtain the distribution $p(\omega) = \sum_{j=1}^m \delta(\omega - \nu_j^2)$ of the squared singular values $\omega = \nu^2$, which coincide with the eigenvalues of $W^T W$, and then simply use that $P_{\text{MP}}(\nu) = 2\nu p(\nu^2)$. If we think of $W^T W$ as the Hamiltonian of a time-reversal invariant system, we can use some insights and methods we are used to from physical systems. The first trick in the derivation is to consider the Green function (trace of the resolvent operator)

$$\mathcal{G}(\omega + i\eta) = \text{tr}([\omega + i\eta - W^T W]^{-1}) \quad (1.19)$$

which can be rewritten using the singular value decomposition of W , i.e. $W^T W = V S^2 V^T$ with $S^2 = \text{diag}(\nu_i^2)$, and the cyclicity of the trace as

$$\mathcal{G}(\omega + i\eta) = \text{tr}(V[\omega + i\eta - S^2]^{-1}V^T) = \text{tr}([\omega + i\eta - S^2]^{-1}) = \sum_{j=1}^m \frac{1}{\omega + i\eta - \nu_j^2} . \quad (1.20)$$

Here, η is just a regularization we will send to zero in a moment. The form of \mathcal{G} on the right-hand side allows to use the Dirac identity

$$\lim_{\eta \rightarrow 0^+} \frac{1}{\omega + i\eta - \nu_i^2} = \mp i\pi \delta(\omega - \nu_i^2) + \mathcal{P} \frac{1}{\omega - \nu_i^2} , \quad (1.21)$$

which reveals the familiar connection of \mathcal{G} with the density

$$p(\omega) = \lim_{\eta \rightarrow 0^+} \frac{-1}{\pi} \text{Im}[\mathcal{G}(\omega + i\eta)] . \quad (1.22)$$

One next considers the limit of large matrices $n, m \rightarrow \infty$ while the ratio $m/n \leq 1$ is kept finite, and for the variance, we assume $\sigma^2 \sim 1/n$ such that $n\sigma^2 \equiv \tilde{\sigma}^2$ also stays finite. Under these conditions, Sengupta et al. argue that \mathcal{G} is self-averaging when taking this limit, $\mathcal{G}(z) \rightarrow G(z) \equiv \langle \mathcal{G}(z) \rangle$, such that $p(\omega) \rightarrow -\lim_{\eta \rightarrow 0^+} \text{Im}[G(\omega + i\eta)]/\pi$ can be computed from an ensemble average. The next trick for obtaining an expression that can be evaluated in this average, is to consider

$$\partial_z \ln \det[z - W^T W] = \partial_z \ln \prod_{j=1}^m (z - \nu_j^2) = \sum_{k=1}^m (z - \nu_k^2)^{-1} = \mathcal{G}(z) . \quad (1.23)$$

For circumventing the logarithm, one further uses the replica trick

$$\begin{aligned} G(z) &= \partial_z \langle \ln \det[z - W^T W] \rangle = -\lim_{k \rightarrow 0} \frac{1}{k} \partial_z \langle \det[z - W^T W]^{-k} - 1 \rangle \\ &= -\lim_{k \rightarrow 0} \frac{1}{k} \partial_z \langle \det[z - W^T W]^{-k} \rangle . \end{aligned} \quad (1.24)$$

At this point Sengupta et al. consider k replicas of a vector with m entries $\mathbf{Y}^{(\alpha)} = (y_{\alpha,1}, \dots, y_{\alpha,m})$ and the auxiliary Gaussian integral

$$Z_k(z) = \int \left[\prod_{\alpha=1}^k \prod_{a=1}^m dy_{\alpha a} \right] \left\langle \exp \left(-\frac{m}{2} \sum_{\alpha=1}^k Y^{(\alpha)T} (z - W^T W) Y^{(\alpha)} \right) \right\rangle \quad (1.25)$$

$$= \left(\frac{2\pi}{m} \right)^{mk/2} \langle [\det(z - W^T W)]^{-k/2} \rangle \quad (1.26)$$

such that

$$G(z) = \lim_{k \rightarrow 0} \frac{-1}{k} \left(\frac{2\pi}{m} \right)^{mk/2} \partial_z Z_{2k}(z) . \quad (1.27)$$

The idea behind this trick is that the auxiliary quantity $Z_k(z)$ can be analytically computed in the limit $m, n \rightarrow \infty$, $\tilde{\sigma} < \infty$, $m/n < \infty$ by explicitly evaluating the ensemble average and then using a saddle point approximation for the remaining integral that becomes exact in this limit. This evaluation requires extensive algebraic steps and one needs to exchange the limits $k \rightarrow 0$ and $m \rightarrow \infty$ which is explained in Ref. [185] (and also in [186]). We will therefore skip these steps for the purpose of this introductory section and quote the result of the auxiliary integral in the saddle point approximation [185]

$$Z_k(z) = \frac{1}{2^{k^2}} \left(\frac{m}{2\pi} \right)^{-mk/2+k^2} \exp \left[-mk \left(\ln(z - \sigma^2 \tilde{Q}(z)) + \frac{m}{n} \ln(1 - R(z)) + Q(z)R(z) \right) \right] \\ R(z) = \frac{\tilde{\sigma}^2}{z - \tilde{\sigma}^2 Q(z)} \quad (1.28)$$

$$Q(z) = \frac{n/m}{1 - R(z)} . \quad (1.29)$$

Taking the partial derivative in Eq. (1.27), where we need to be careful that the derivative does not act on $G(z)$ and $Q(z)$ like a total derivative would (see Eq. (1.24)), yields

$$G(z) = \lim_{k \rightarrow 0} \frac{2mk}{2k2^{k^2}} \left(\frac{m}{2\pi} \right)^{(2k)^2} e^{-2mk[\ln(z - \sigma^2 \tilde{Q}(z)) + \frac{m}{n} \ln(1 - R(z)) + Q(z)R(z)]} \frac{1}{z - \tilde{\sigma}^2 Q(z)} \\ = \frac{m}{z - \tilde{\sigma}^2 Q(z)} = \frac{m}{\tilde{\sigma}^2} R(z) . \quad (1.30)$$

This implies that $Q(z) = (n/m)/(1 - \tilde{\sigma}^2 G(z)/m)$ such that one obtains the implicit expression

$$G(z) = \frac{m}{z - \tilde{\sigma}^2 \frac{m/n}{1 - \tilde{\sigma}^2 G(z)/m}} , \quad (1.31)$$

which can be directly solved, resulting in

$$G(z) = \frac{m}{2\tilde{\sigma}^2 z} \left[z + \tilde{\sigma}^2(1 - n/m) \pm \sqrt{[z - \tilde{\sigma}^2(1 + n/m)]^2 - 4\tilde{\sigma}^4 n/m} \right] . \quad (1.32)$$

The last step³ for computing $p(\omega)$ is to use $z = \omega + i\eta$ and compute the imaginary part of G :

$$p(\omega) = -\frac{1}{\pi} \lim_{\eta \rightarrow 0^+} \text{Im}[G(\omega + i\eta)] \quad (1.33)$$

$$= \frac{m}{\pi \tilde{\sigma}^2 2\omega} \sqrt{\omega^2 - 2(1 + n/m)\omega \tilde{\sigma}^2 - 4\tilde{\sigma}^4 n/m + (1 + n/m)\tilde{\sigma}^4} . \quad (1.34)$$

By introducing the boundary values at which the expression under the square root is non-negative, $\nu_{\pm} = \tilde{\sigma}^2(1 \pm \sqrt{n/m})$, one can rewrite the expression in the usual Marcenko-Pastur form [184]

$$p(\omega) = \frac{m}{2\pi \tilde{\sigma}^2 \omega} \sqrt{(\omega - \nu_-^2)(\nu_+^2 - \omega)} . \quad (1.35)$$

³Here the \pm sign in the solution of G just results in an overall sign, $\pm p(\omega)$, such that only the sign which gives a positive probability density is relevant.

As discussed at the beginning of the derivation, this allows us to directly write down the distribution of the singular values ν as

$$P_{\text{MP}}(\nu) = 2\nu p(\nu^2) = \frac{m}{\pi\tilde{\sigma}^2\nu} \sqrt{(\nu^2 - \nu_-^2)(\nu_+^2 - \nu^2)} \quad (1.36)$$

which is the final result of this derivation used in Ch. 3 to compare the singular value spectra of weight matrices of neural networks before and after training.

For the $n \times m$ random matrix W of i.i.d. normal entries with $n \geq m$, we showed that the eigenvalues of the Wishart matrix $X = W^T W = V S^2 V^T$ are distributed according to the Marcenko-Pastur distribution. Here, the eigenvalues of X are the squared singular values of W , and the eigenvectors correspond to the right singular vectors V . The entries V_{ik} of these eigenvectors are distributed according to the Porter-Thomas distribution [187], i.e. a normal distribution with mean zero and variance $1/m$. To show this, we consider the probability density of the matrix W , which is given by

$$P(W) \propto \prod_{ij} e^{-\frac{1}{2\sigma^2} W_{ij}^2} = e^{-\frac{1}{2\sigma^2} \text{tr}(W^T W)} , \quad (1.37)$$

such that it is invariant under rotations. If R is an orthogonal $m \times m$ matrix, then the cyclicity of the trace $\text{tr}(RW^T R^T R W R^T) = \text{tr}(W^T W)$ ensures that W and WR are equally distributed. Therefore, the singular value decompositions $W = U S V^T$, $WR = U S V^T R$ imply that also V^T and $V^T R$ follow the same distribution. As the singular vectors $\mathbf{V}^{(k)}$ in $V = (\mathbf{V}^{(1)}, \dots, \mathbf{V}^{(m)})$ are normalized, i.e. $\sum_{j=1}^m |V_j^{(k)}|^2 = 1$, the singular vector entries are therefore uniformly distributed on the m -sphere of unit radius. As it was shown in Ref. [11], the probability density of a coordinate x_i of a vector uniformly distributed on the m -sphere is given by

$$P(x_i) = \mathcal{N} \left(1 - x_i^2 \right)^{(m-3)/2} . \quad (1.38)$$

This can be found by first noting that uniform distribution on the m -sphere means that the angles in hyperspherical coordinates are uniformly distributed, which allows us to explicitly write down the expression for the probability density of the angles, $P(\varphi_1, \dots, \varphi_{m-1}) d\varphi_1 \cdots d\varphi_{m-1} \propto d\varphi_1 \cdots d\varphi_{m-1}$. Using the Jacobi determinant for the transformation between hyperspherical and Euclidean coordinates and the inverse transformation to Euclidean coordinates x_1, \dots, x_m allows us to get an expression for $P(x_2, \dots, x_m)$ (where one coordinate x_1 is fixed by the constraint $r = 1$) and integrating out all coordinates except for x_i , one arrives at Eq. (1.38).

From Eq. (1.38), we obtain the mean and variance of x_i as

$$\text{mean}(x_i) = \int_{-1}^1 dx_i P(x_i) x_i = 0 \quad (1.39)$$

$$\text{var}(x_i) = \int_{-1}^1 dx_i P(x_i) x_i^2 = 1/m . \quad (1.40)$$

To keep the variance finite in the limit $m \rightarrow \infty$, we rescale $\xi_i = x_i \sqrt{m}$ and find [11]

$$\lim_{m \rightarrow \infty} P(x_i) = \mathcal{N} \lim_{m \rightarrow \infty} \left(1 - \xi_i^2/m \right)^{(m-3)/2} \propto \lim_{m \rightarrow \infty} \left(1 - \xi_i^2/m \right)^{m/2} \propto e^{-\xi_i^2/2} = e^{-m x_i^2/2} , \quad (1.41)$$

which is the Porter-Thomas distribution – a normal distribution with mean zero and standard deviation $1/\sqrt{m}$.

As the derivations of the Marcenko-Pastur and Porter-Thomas law required rotational invariance of the distribution of the matrices and to take the limit $n, m \rightarrow \infty$, with fixed $m/n < 1$, we expect them to be applicable to both large matrices with i.i.d. normal entries and large matrices with i.i.d. uniformly distributed entries. The latter is of particular relevance here, since the default initialization for weight matrices in the popular neural network python package `tensorflow` [188], is the so-called Glorot uniform distribution [189]. For the Glorot uniform initialization, the entries of the random $n \times m$ matrix are drawn uniformly from the interval $I = [-\sqrt{6/(n+m)}, \sqrt{6/(n+m)}]$, i.e. the probability density is given by

$$P_{\text{Gu}}(x) = \begin{cases} 2^{-1}(6/(m+n))^{-1/2} & \text{for } x \in I \\ 0 & \text{else} \end{cases} \quad (1.42)$$

with the first and second moment

$$\text{mean}(x) = \int_I dx x P_{\text{Gu}}(x) = 0 \quad (1.43)$$

$$\text{var}(x) = \int_I dx x^2 P_{\text{Gu}}(x) = \frac{2}{m+n} . \quad (1.44)$$

Therefore, the Glorot distribution is also translationally invariant and has a variance $\sim 1/n$ in the limit for $n, m \rightarrow \infty$.

To get an intuition for how large n and m need to be to describe random Glorot uniform matrices using the Wigner surmise, Porter-Thomas distribution, Marcenko-Pastur distribution, and logarithmic level number variance, we consider square matrices of sizes $n = 80, 300, 512$ in Fig. 1.2. It becomes apparent that even $n = 80$ is sufficient to see good visual agreement with the Marcenko-Pastur and Porter-Thomas distribution (panels a and b). Here deviations from the Wigner surmise and logarithmic grow of the level number variance is due to the fact that unfolding the spectrum is very inaccurate for small matrices. For matrix dimensions of $n, m \gtrsim 300$, we find excellent agreement with all the predictions of RMT described before. We therefore conclude that we can apply these predictions for weight matrices of neural networks, which typically have dimensions of several hundred to several thousand entries per row.

1.2 Neural networks and the CMA-ES machine learning algorithm

In this thesis, we focus on two different types of machine learning algorithms – neural networks and the Covariance Matrix Adaptation Evolution Strategy algorithm (CMA-ES). Both have in common that they train parameters based on data shown to them over several iterations to subsequently solve a problem. Here, neural networks consist of layers of neurons connected by weights. When the network is presented with a data example, it activates the neurons in the first layer, and the weights determine how this information is propagated through to the output layer, whose activations are then the answer of the network to the problem. To train the network, one needs a large dataset of examples and the desired output for each example – the labels of the training dataset. Initially, the weights are chosen randomly, and during training, one presents the network the training examples and adjusts the weights such that the predictions match the labels. The hope is that the trained network then generalizes, i.e. that it makes correct predictions for completely new examples.

The CMA-ES algorithm, on the other hand, aims to find the minimum of any function of many variables. To do this, it does not need to know the function itself, but the algorithm

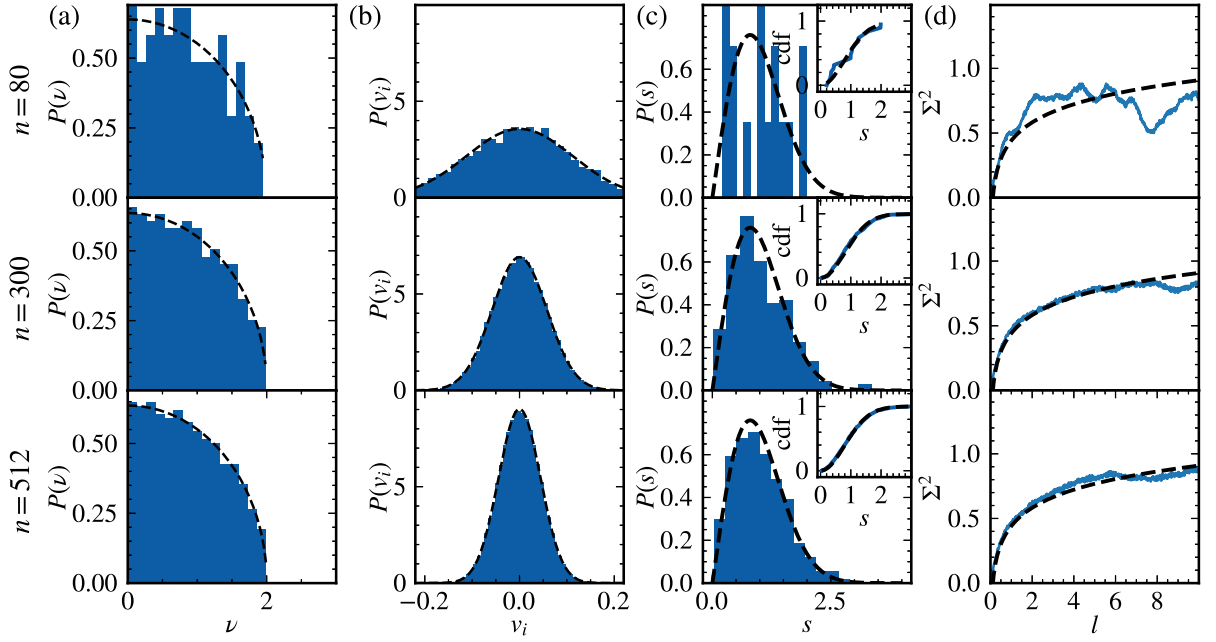


Figure 1.2: Comparison of some results of random matrix theory for i.i.d. Glorot uniform matrices of different sizes (blue) and theory curves (dashed, black): (a) Histogram of the singular values and Marcenko-Pastur distribution; (b) Histogram of the entries of all singular vectors and Porter-Thomas distribution; (c) Histogram of spacings of unfolded singular values and Wigner surmise with the corresponding cummulative distributions shown in the insets; and (d) level number variance of unfolded singular values obtained via Gaussian broadening and the GOE logarithm law, $\propto \ln(2\pi l) + \text{const.}$. Here, Gaussian broadening refers to estimating the probability density from the vector of singular values as a sum of Gaussian functions centered around each singular value with width determined by the local level spacing (for details see Sec. 3.2.2). We find excellent agreement with the RMT predictions already for matrices with $n \sim 300$.

suggests a set of points – a population – in each iteration, on which one evaluates the function and provides the function values back to the algorithm. Using the best candidates from the population, the algorithm then learns directions and regions in the function’s input space to suggest better points for the next population.

In the following, we give a brief overview of how the algorithms work, problems to which they have been successfully applied, and how we are going to use them in the upcoming chapters.

1.2.1 Neural networks

In the following, we focus on neural networks for image classification that are trained using supervised learning. The idea is that we can give the network an image as input, and it will tell us as output what is shown in the image. To achieve this, the network is trained on a dataset containing many images and corresponding labels, so that the desired output of the network matches the labels. A popular dataset is CIFAR-10 [190], which consists of a total of 60000 colored images of size $32 \times 32 \times 3$ from ten different classes (airplane, automobile, bird, cat, deer, dog, frog, horse, ship, truck). Of this dataset, we use $N = 50000$ images $\mathbf{x}^{(k)}$ and their corresponding labels $\mathbf{y}^{(k)}$ as the training dataset, and the remaining 10000 images are only

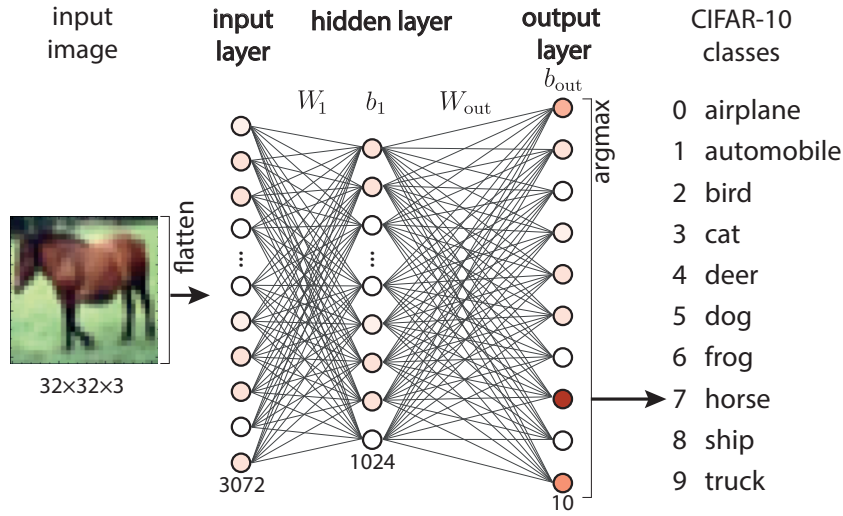


Figure 1.3: Example of a fully connected neural network with a single hidden layer. The network consists of an input layer with 3072 neurons (circles) that accepts images of size 32×32 with three color channels, an output layer of 10 neurons, each representing a class of the CIFAR-10 dataset, and the hidden layer with 1024 neurons. Every neuron in layer l is connected with each neuron in the previous layer $l - 1$ by weights $W^{(l)}$ and each layer (except for the input layer) has an associated bias vector $b^{(l)}$. For hidden layers we consider ReLu activations, and we use softmax as the output activation function. An image is flattened and then activates the input layer according to its pixel color values. Then these activations are propagated through the network, where the index of the neuron with the highest output activation defines the class predicted by the network.

shown to the network as a test dataset after training is fully completed to evaluate how well the network can generalize to new images. Thus, if the network learns an underlying rule in the training dataset, for example how to distinguish a horse from a deer, it is expected to be able to transfer this knowledge to the test dataset. On the other hand, if the network has only remembered the pixel values of the training images to distinguish them, this so-called overfitting prevents the network from generalizing to new, unseen images.

A simple feedforward network, such as shown for the case of three layers in Fig. 1.3, consists of a series of layers l containing n_l neurons whose activations can be described by a vector $\mathbf{a}^{(l)} \in \mathbb{R}^{n_l}$. Each layer after the input layer also has an $n_l \times n_{l-1}$ weight matrix $W^{(l)}$ and an $n_l \times 1$ bias vector $\mathbf{b}^{(l)}$, which contain the trainable parameters of the network. In addition, one assigns an activation function $f^{(l)} : \mathbb{R}^{n_l} \rightarrow \mathbb{R}^{n_l}$ to each layer $l > 1$, which allows the propagation of information to be strongly non-linear. One then chooses the input layer $l = 0$ and the output layer $l = L$ specifically for the problem at hand, so that we use the 3072 pixel values of a CIFAR-10 image directly as activations of the input layer, and the output layer indicates which of the ten classes the image belongs to. The layers in between are called hidden layers and can in principle be chosen arbitrarily – independent of the problem. If we now flatten a $32 \times 32 \times 3$ input image into a vector \mathbf{x} of length 3072, we can pass this to the network by activating the first layer with the pixel values

$$\mathbf{a}^{(0)} = \mathbf{x} . \quad (1.45)$$

By using the parameters of the network, these activations are then propagated through the network so that the activations in layer l are given by

$$\mathbf{a}^{(l)}(\mathbf{x}) = f^{(l)} \left(W^{(l)} \mathbf{a}^{(l-1)}(\mathbf{x}) + \mathbf{b}^{(l)} \right) . \quad (1.46)$$

In this way, activations are passed from layer to layer until they finally activate the output layer $\mathbf{a}^{(L)}(\mathbf{x})$. If we assign a class of the CIFAR-10 dataset to each of the ten output neurons and choose the labels so that they consist of vectors that have a 1 at the position of the class of the image and zeros in all other entries, we can define the prediction of the network for the class of the input image as the position of the neuron with the largest activation in the output layer.

At this point, of course, it is far from obvious how to choose the numerous parameters $W^{(l)}$ and $\mathbf{b}^{(l)}$ so that the network makes correct predictions. Therefore, one first selects the parameters randomly, for example by filling the weight matrices with i.i.d. normal random numbers and setting the bias vectors to zero. This, of course, makes the network before training perform no better than directly guessing the predictions randomly. So the question is how to align the parameters with the training data set such that the network can make accurate predictions. This is where training comes into play: First, we define a measure for the error that the network makes in the prediction of the training data, the so-called loss function, here exemplified by the mean squared error

$$\ell(\mathbf{x}, \mathbf{y}; \mathbf{W}, \mathbf{b}) = \frac{1}{2} \sum_{i=1}^{10} \left| a_i^{(L)}(\mathbf{x}) - y_i \right|^2 . \quad (1.47)$$

During training, one shows the network (sequentially or in batches) the images of the training dataset and minimizes the loss function using gradient descent (for details see Sec. 3.1.2) by adjusting weights and biases in each iteration according to the direction of the gradient $-\nabla_{\mathbf{W}} \ell$ and $-\nabla_{\mathbf{b}} \ell$, respectively. After training, one can use the test dataset to verify that the network has also learned the underlying rule to correctly predict the classes of the unknown images.

An archetype of such networks – the perceptron consisting of only a single output neuron and the input layer – was considered as early as 1958 with the goal of providing a highly simplified model for a neuron in the brain [191, 192]. Then in the 1960s, the first feedforward networks with multiple layers appeared [193–197] and again about ten years later training with gradient descent and backpropagation became popular [42, 198–200] as it is still used today. However, in this early period of neural networks, successes were limited and research had stalled by the 1990s: For one, a lack of large data sets and computer resources limited the training of sizable networks, but conceptual misunderstandings also prevailed which could not be resolved until later [42]. On the one hand, it was believed that training with gradient descent must necessarily get stuck in local minima [42]. On the other hand, one still expected the classical bias-variance tradeoff [201], and thus assumed that a large number of parameters would have to lead to extreme overfitting, which would prevent the networks from generalizing. Similar to the fit of polynomials to a set of points, it was assumed that considering too small orders would not allow the polynomial to capture all the features of the data, on the other hand, using an order as large as the number of points may perfectly capture all the points, but also fits the noise such that generalization to new points is not possible. It was not until after the turn of the millennium that the field of Deep Learning experienced a renaissance, where with the availability of larger datasets and computational resources, surprising successes were achieved in image [42, 202] and speech recognition [42, 203] - areas where traditional algorithms have always

struggled. Surprisingly, successful networks tend to have many more parameters than necessary to remember the training dataset [190, 204–214]. Now, instead of the classical bias-variance tradeoff, the networks were found to exhibit a double descent behavior of the generalization error with the number of parameters [215–217], i.e., the networks have the best performance when they are highly overparameterized.

In recent years, neural networks have become increasingly popular and have achieved tremendous success in many fields [42–47]: For example, modern networks surpass human abilities in image recognition [44], in board games such as chess [218] and Go [45], and even in complex real time strategy games such as Star Craft II [219]. They have even enabled significant progress in solving extremely hard problems such as protein folding [48]. Neural networks today have applications in many areas of everyday life, and yet there is no comprehensive theoretical understanding of their learning dynamics and how information is encoded in the weights [37, 38]. Most recently, some successes have been made in understanding overparameterized networks in the limit of infinitely wide layers using statistical analysis: Infinitely wide networks have been modeled with Gaussian processes [220–223] and insights into the dynamics for training with gradient flow [222–226] has been studied.

Another problem is obtaining accurate datasets for training. Although there is an abundance of data, it is extremely time and resource consuming to generate correct labels for the datasets, making mislabeled examples almost inevitable in large datasets [227]. Unfortunately, this label noise significantly reduces the generalization ability of the networks [40], so it is desirable to localize the noise in the networks, to make the networks more robust to mislabeling, or even filter out the memorized noise from the learned weights.

1.2.2 Applications of random matrix theory to neural networks

In Sec. 1.1.1, we mentioned several examples of successful applications of RMT in a variety of different fields. One use case we have omitted so far is the application of RMT to describe neural networks. The large number of parameters of deep neural networks (DNNs) with non-linear activation functions, makes it impossible to track the evolution of these parameters during training [190, 204–214]. Especially in the case of pattern and image recognition tasks, the problem is further complicated by the fact that the training data contains a lot of information that is irrelevant for the underlying rule that the network is supposed to learn in order to generalize to unseen data. On the other hand, neural networks are amazingly successful [42–48, 228], and many trained networks and promising architectures are available – so data is abundant. Thus, the situation is remarkably similar to that of nuclear physics in the 1950s, where RMT had already been proven to be a fruitful tool.

To make progress with a statistical analysis of networks, the generalization performance of single layer networks in the limit of an infinitely wide layer has been studied, where the asymptotic generalization error can be minimized by choosing weights based on a dataset by using RMT [35, 36]. Another approach is based on understanding the generalization dynamics in linear networks [229], i.e., the dynamics are made tractable by refraining from using nonlinear activation functions. This statistical analysis is based on a teacher-student setup instead of using an empirical training dataset. Here, the teacher is the network that generates the training data, and the goal is to choose the weights of the student network, which may have different dimensions than the teacher, so that it approximates the teacher as well as possible. These applications of RMT have in common that the weights are optimally chosen based on a statistical analysis rather than being trained on an empirical training dataset as is relevant for most modern applications.

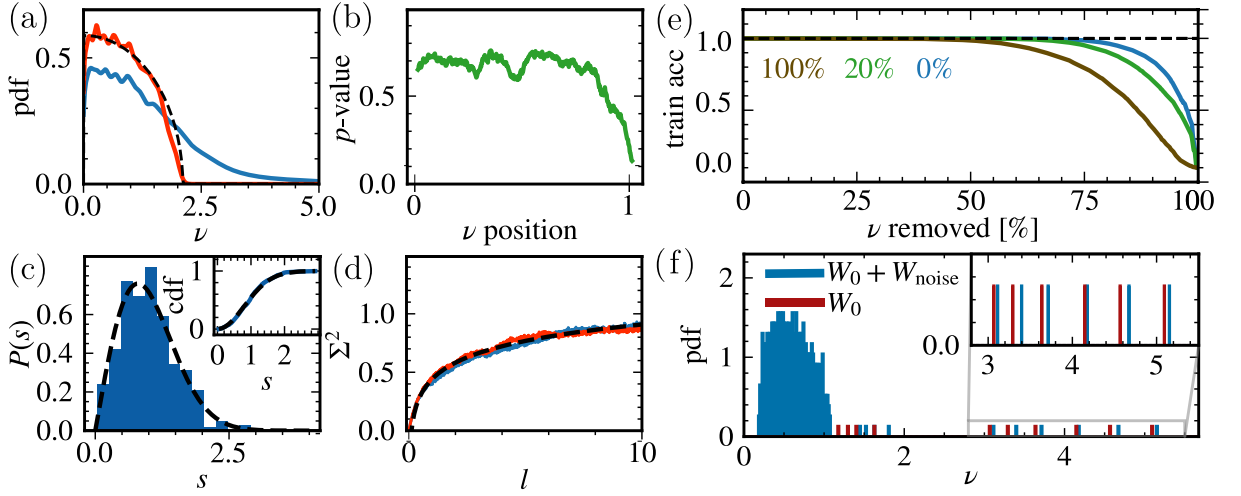


Figure 1.4: Example for a random matrix theory analysis of the weight matrix of a typical fully connected neural network (for details see Ch. 3). (a) Probability density of singular values obtained with Gaussian broadening (window size 15) for the trained weight (blue) and the for the random weight before training (red) together with the Marcenko-Pastur theory curve (black). After training one observes a Marcenko-Pastur bulk and a heavy tail with some large singular values. (b) Averaged p -values of Kolmogorov-Smirnov test of the singular vector entries against a Porter-Thomas distribution. One observes a significant drop of the p -values for singular vectors corresponding to large singular values which indicates deviations from the Porter-Thomas distribution for the entries of these vectors. (c) Level spacing distribution of unfolded singular values together with the Wigner surmise (main panel) and corresponding cumulative distributions (inset) showing excellent agreement. (d) Level number variance of unfolded singular values for the trained weight (blue) and random initial weight (red) together with the log-law, $\Sigma^2(l) \propto \ln(2\pi l) + \text{const.}$ (black). We again find excellent agreement with the universal RMT prediction. (e) Effect on the training accuracy of setting a percentage of the singular values (starting from the smallest ones) to zero in networks trained with various amounts of label noise (0% blue, 20% green, 100% brown). It becomes apparent that most of the singular values are not relevant for the accuracy, the largest singular values contain the information learned during training, and intermediate singular values and corresponding vectors are important in the case of training with label noise. (f) Example for the level repulsion of singular values, i.e. the upwards shift of singular values of a low rank matrix W_0 (red) due to an i.i.d. normal distributed dense bulk W_{noise} with $\mu = 0$ and $\sigma = 0.02$ (blue).

However, RMT has also been successfully applied to networks trained with backpropagation. For example, one application [230] is based on the observation that learning can be significantly sped up by a special initialization of weights so that the Jacobian between input and output has singular values tightly localized around one [231]. Here, RMT is used to find initializations that have Jacobians with such spectra [230]. Further analysis [232, 233] is based on the complicated loss surface, i.e., the weight dependence of the prediction error, which is accessible via the Hessian matrix of the loss function. It turns out that the spectra of the Hessian matrix for networks trained on real data agree with universal GOE predictions [232–234].

Furthermore, the evolution of the spectra of the weight matrices during training has been studied with RMT [235]: It turns out that the bulk of a typical spectrum of a weight matrix can

be described with the Marcenko-Pastur distribution. Martin et al. also report a heavy tailed self regularization, i.e. successfully trained modern DNNs had a spectral density with a powerlaw tail, as known from spectra of heavy-tailed random matrices [236, 237]. On the other hand, a clear separation between bulk and large singular values is observed for smaller, old networks. Based on these findings, Martin et al. [238] use spectral norms of the weights, as well as fit results for the exponents of the powerlaw tails to predict the generalization abilities of pre-trained networks.

In Chapter 3, we use RMT as a zero-information hypothesis to study the weights of trained neural networks. In Fig. 1.4, we show an example of such an analysis for the weight matrix of a feedforward network trained on the CIFAR-10 dataset (for more details see Ch. 3). We find excellent agreement between the universal RMT predictions (level spacing statistics: panel c, level number variance: panel d) and the trained weight matrix (blue), indicating that the bulk of the weights remains random during training. The non-universal predictions (singular value distribution: panel a; eigenvector entry distribution: panel b) show excellent agreement for the parts of the spectrum with small singular values and deviations for large singular values. This suggests that the learned information is stored in the largest singular values and corresponding vectors only. To validate this hypothesis, we set parts of the singular values to zero, reconstruct the weights, which then have lower rank, and evaluate the resulting network on the training data set (panel e). We repeat this analysis also for networks trained with a noisy training set, where we shuffled a certain percentage of the labels. It becomes apparent that indeed small singular values contain no information and large singular values and corresponding vectors encode the underlying rule of the training dataset. Interestingly, the intermediate singular values seem to store the memorized noisy labels. Hence, there are two factors that negatively influence the ability of the network to generalize when training with label noise: (i) the network additionally has to memorize the noise and (ii) under the assumption that the rule can be encoded in a low rank matrix, level repulsion with the random bulk causes an upwards shift of the few singular values encoding the rule (panel f). To mitigate these effects, we propose a filtering algorithm in Ch. 3: Assuming an additive noise term, we can revert this shift and by setting the majority of the singular values to zero, we can remove parts of the memorized noise. We will find that this can indeed significantly improve the generalization performance of networks trained with label noise.

1.2.3 The CMA-ES machine learning algorithm

Before proceeding with an introduction to topological superconductor nanowires, we would like to introduce another machine learning algorithm that we will use in the context of such wires with strong disorder. The Covariance Matrix Adaptation Evolution Strategy (CMA-ES) algorithm [239–248] is a population-based sampling algorithm for finding minima of non-continuous, non-monotonic functions $f(\mathbf{x}): \mathbb{R}^n \rightarrow \mathbb{R}$. In each iteration t , the algorithm generates a population, i.e., a set of candidates for the minimum, by randomly drawing from a multivariate normal distribution with covariance $(\sigma^{(t)})^2 \mathbf{C}^{(t)}$ and mean \mathbf{m}_t . Subsequently, these candidates are evaluated with the function, and the algorithm then determines a new covariance matrix and mean for the next population from the best candidates that yield the smallest function values. In doing so, the algorithm learns the region of the minimum by first expanding the search region to have a good chance of capturing points near the minimum, and then contracting the search region around the minimum (see Fig. 1.5). This mode of operation is commonly called a black-box optimizer – not because the optimization process is hidden, but because the algorithm does not need any information about the function other than the function values for the candidates.

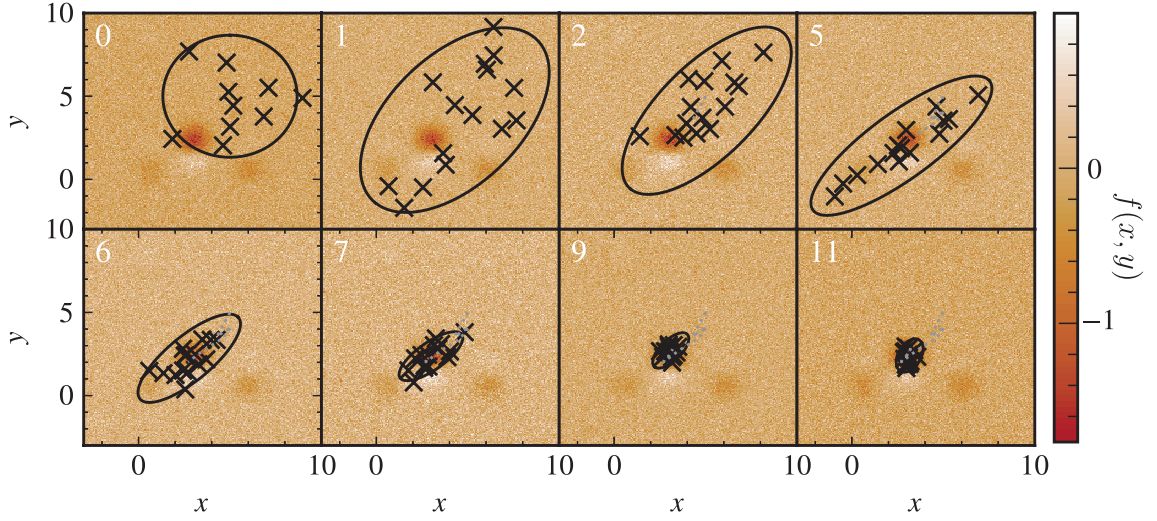


Figure 1.5: Example application of the CMA-ES algorithm for finding the minimum of a non-monotonic, noisy function $f(x, y): \mathbb{R}^2 \rightarrow \mathbb{R}$. The ellipses show the search region encoded in the covariance matrix $\mathbf{C}^{(t)}$ and the step size $\sigma^{(t)}$ in which most points of current population (black crosses) are located. The white numbers in the upper left-hand corners of the panels indicate the step t , and the gray dotted line indicates the path of the mean values during the steps. Initially the search region extends to explore a large region of the input space, and then it contracts around the region of the minimum. We choose an initial point $(5, 5)$, an initial step size 1.5, and a population size of 15.

Thus, the function is like a black box for the optimizer. This makes the CMA-ES machine learning algorithm particularly powerful for optimizing high-dimensional, non-monotonic, and non-continuous problems, and also, for example, for optimizing metrics in experiments where the functional dependence between input and output may be so complicated that it actually looks like a black box to the user.

Since the first proposal in 1995 by Hansen and Ostermaier [239], the algorithm has been steadily developed [240–248] and has since found numerous applications to minimizing high-dimensional problems [246, 249–252]. In the following, we summarize the algorithm [241, 253] for optimizing the function $f: \mathbb{R}^N \rightarrow \mathbb{R}$ (for a more detailed description of the steps of the algorithm, see Appendix A.5): The CMA-ES algorithm is first initialized with a starting step size $\sigma^{(0)}$, the starting correlation matrix $\mathbf{C}^{(0)} = \mathbf{I}$, and a gate configuration $\mathbf{V}_g^{(0)}$. In each iteration t , a population of n_{pop} candidate configurations $\mathbf{V}_{g,1}^{(t)}, \dots, \mathbf{V}_{g,n_{\text{pop}}}^{(t)}$ is drawn from a multivariate normal distribution with mean $\mathbf{V}_g^{(t)}$ and covariance matrix $(\sigma^{(t)})^2 \mathbf{C}^{(t)}$. Then, the function is evaluated for each candidate in the population and based on the \tilde{n} candidates with the smallest function values, the parameters for the next population are adjusted. Here, the new mean $\mathbf{V}_g^{(t+1)}$ is determined as a weighted average of the \tilde{n} best candidates. Both the step size $\sigma^{(t+1)}$ and the correlation matrix $\mathbf{C}^{(t+1)}$ are indirectly adjusted by first updating so-called evolution paths $\mathbf{s}_c^{(t)}$ and $\mathbf{s}_\sigma^{(t)}$ using weighted averages involving the path from the previous step and the difference of the means $\delta \mathbf{s}^{(t+1)} = (\mathbf{V}_g^{(t+1)} - \mathbf{V}_g^{(t)})/\sigma^{(t)}$. Then, the covariance matrix is updated as a weighted average including $\mathbf{s}_c^{(t+1)}[\mathbf{s}_c^{(t+1)}]^T$ and the step size is updated as a function of $\|\mathbf{s}_\sigma^{(t+1)}\|$ (for details see Appendix A.5). It turns out that separating $\sigma^{(t)}$ and $\mathbf{C}^{(t)}$ improves both robustness and speed of convergence of the algorithm by adjusting them on different time scales

[243]. During optimization, the search region from which the population is drawn (determined by $\mathbf{V}_g^{(t)}$ and $(\sigma^{(t)})^2 \mathbf{C}^{(t)}$) initially expands to find the region of the optimum, even if it is not in the starting region. Subsequently, the search region then contracts around the optimum. Convergence is achieved when either the difference of best function values $|A_{\text{best}}^{(t)} - A_{\text{best}}^{(t+1)}|$ or of the best candidates $\|\mathbf{V}_g^{(t)} - \mathbf{V}_g^{(t+1)}\|$ is repeatedly smaller than a tolerance value.

Conveniently, a mature Python implementation of the CMA-ES algorithm exists in the `pycma` [254] package, which we use in Sec. 2.5 to optimize topological superconductor wires with strong disorder, allowing us to stabilize Majorana zero modes and compensate for disorder effects. Here, a special requirement for an optimization algorithm is that it can also optimize non-continuous functions caused by the disorder and that the number of necessary function evaluations, i.e. measurements in an experiment, is kept as small as possible, e.g. by avoiding many measurement points for approximating gradients.

1.3 One-dimensional Majorana wires

In recent years, Majorana zero modes (MZMs) have received much attention as potential building blocks for qubits in a quantum computer because of their unique properties [56, 59, 79–84]. One characteristic that makes them particularly suitable for this purpose is their topological protection, i.e. that they are robust against small local perturbations: MZMs are strongly non-local zero-energy states that occur, for example, at the edges of topological superconductors [86, 255] and are therefore separated from higher energy levels by the superconducting gap [79, 255]. Moreover, MZMs – when moved around each other – exhibit interesting braiding statistics that enable protocols for partially protected operations on Majorana qubits [79, 256]. Particularly, scalable array structures of Majorana nanowires provide promising architectures for realizing quantum computers with many qubits [79, 83], but currently there are still numerous problems with their realization [102–106]. In the following, we start with a brief introduction into the properties of MZMs and their occurrence in topological superconductors. Furthermore, we describe the problems currently encountered in realizing MZMs in experiments and motivate possible solutions, which we discuss in detail later in Ch. 2. As the last part of this introductory section, we summarize some basic ideas to motivate how MZMs might be used as qubits in the future.

1.3.1 Majorana zero modes in topological superconductor nanowires

In 1928, P. Dirac formulated the Dirac equation

$$i\partial_t\psi(\mathbf{x}, t) = [-i\boldsymbol{\alpha} \cdot \nabla + m\beta]\psi(\mathbf{x}, t) \quad (1.48)$$

to describe relativistic fermions [255, 257, 258], where ψ is a spinor with four components, such that in Dirac’s representation the first two describe spin- \uparrow and spin- \downarrow electrons, and the other two components describe solutions with negative energy [259]. Here we use units in which $\hbar = c = 1$. Choosing the 4×4 matrices α_i and β such that $\{\alpha_i, \alpha_j\} = 2\delta_{ij}$, $\{\alpha_i, \beta\} = 0$, and $\beta^2 = \mathbb{1}$, the solutions of the Dirac equation also fulfill the Klein-Gordon equation, which is derived from the relativistic energy-momentum relation by the typical substitution $\mathbf{p} \rightarrow -i\nabla$ and $E \rightarrow i\partial_t$ [260]. It will become relevant later that with Dirac’s choice of matrices,

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} \quad \beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \quad (1.49)$$

by complex conjugation of the Dirac equation it is found that the complex conjugate wavefunctions ψ^* satisfies a different equation than ψ . The main criticism, however, which was raised against Dirac's theory was that the equation implies an infinite number of solutions with negative energy: If this was the case, a system could just keep lowering its energy by emitting photons [259, 261]. Dirac solved this problem by assuming that all negative states are filled by an electron and that removing an electron with energy $-\varepsilon$ from this *Dirac sea* costs an energy ε to create a hole with positive energy ε , equal mass m_e but opposite charge $-e$ [259, 261]. If ψ is a solution of the Dirac equation, then the solution of the corresponding antiparticle with opposite charge (the positron) is given by the charge conjugate spinor $C\psi^*$ [260, p. 503]. For $C\psi^*$ to actually be a solution, C must satisfy the following conditions: $C\beta^*C^{-1} = -\beta$ and $C(\beta\alpha_i)^*C^{-1} = -\beta\alpha_i$. In the representation chosen by Dirac, Eq. (1.49), this is satisfied by the choice $C = i\beta\alpha_2$, since the real matrices $\beta\alpha_1$, $\beta\alpha_3$, and β anticommute with C and the imaginary matrix $\beta\alpha_2$ commutes with C [260, p. 503]. Thus, the positron solution to the associated electron solution $\psi(\mathbf{x}, t)$ is given by $i\beta\alpha_2\psi^*(\mathbf{x}, t)$ [260, p. 504]. After initial skepticism about Dirac's solution to the problem of negative energy states [259, 261], C. D. Anderson was able to demonstrate the existence of positrons in 1933 when he observed cosmic rays creating positively charged particles with the mass of an electron in a cloud chamber [262]. Subsequently, the Dirac equation was accepted as the standard description for relativistic electrons [260, p. 500].

Motivated by the success story surrounding the prediction and discovery of the positron, in 1937 E. Majorana considered the Dirac equation with a focus on real solutions, which cannot be found with Dirac's choice of the matrices α_i and β [263]. However, if one chooses a different representation of the matrix algebra [255, 263]

$$\alpha_1 = \begin{pmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 0 & \sigma_3 \\ \sigma_3 & 0 \end{pmatrix}, \quad \alpha_3 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \quad \beta = \begin{pmatrix} 0 & \sigma_2 \\ \sigma_2 & 0 \end{pmatrix}, \quad (1.50)$$

then this choice also satisfies the necessary conditions $\{\alpha_i, \alpha_j\} = 2\delta_{ij}$, $\{\alpha_i, \beta\} = 0$, and $\beta^2 = \mathbb{1}$. In this representation, complex conjugation of the Dirac equation yields

$$-i\partial_t\psi^*(\mathbf{x}, t) = [i\boldsymbol{\alpha} \cdot \nabla - m\beta]\psi^*(\mathbf{x}, t), \quad (1.51)$$

since all terms pick up a minus sign [255]. Now if ψ is a solution of the Dirac equation, then ψ^* is also a solution and since the charge conjugation conditions $C\beta^*C^{-1} = -\beta$, $C(\beta\alpha_i)^*C^{-1} = -\beta\alpha_i$ are directly satisfied for $C = \mathbb{1}$, ψ^* describes the antiparticle to ψ [255]. This also opens the possibility for real solutions $\psi = \psi^*$ and thus for charge neutral particles which are their own antiparticle [263] – the so-called Majorana fermions [255]. Thus, in second quantization notation, such elementary particles are described by creation and annihilation operators that satisfy the Majorana condition

$$\gamma = \gamma^\dagger. \quad (1.52)$$

To this day, however, there is no conclusive evidence for the existence of Majorana fermions as elementary particles [255], so here we instead consider quasiparticles, i.e., emergent collective excitations in condensed matter systems. To satisfy the Majorana condition, it is promising to consider superconductors [264], since there, due to particle-hole symmetry, the annihilation operator for a quasiparticle with energy E is equal to the creation operator for a particle with energy $-E$, i.e. $\gamma_E^\dagger = \gamma_{-E}$. Therefore, a neutral quasiparticle with zero energy could satisfy the Majorana condition as an equal superposition of hole and electron. Unfortunately, however, this does not work for a conventional s-wave superconductor, where Cooper pairs are formed

from electrons with opposite spin, and therefore $\gamma = uc_{\uparrow} + v^*c_{\downarrow}^{\dagger}$ cannot be equal to γ^{\dagger} for any choice of u and v . Thus, one needs a more exotic superconductor, for example a spinless p-wave superconductor with the quasiparticle annihilation operator $\gamma = uc + v^*c^{\dagger}$, for which the Majorana condition for an equivalent superposition of hole and electron, $u = v$, could be satisfied [265]. Indeed, Majorana modes have been predicted in numerous systems: Among others, as quasiparticle excitations of the $\nu = 5/2$ quantum Hall state [266, 267], at the edges of two-dimensional p-wave superconductors [86, 268], in vortex cores of spinless $p_x + ip_y$ superconductors [256, 267, 269–272], in trapped cold atom systems [255, 273–276], in coupled quantum dots [277, 278], and for a chain of magnetic addatoms on an s-wave superconductor [279–282]. In the following, we will focus on Majorana modes in low-dimensional topological superconductors, since there are promising protocols for their application as qubits in a quantum computer [79, 283] (see also Sec. 1.3.2 below).

A toy model of such a topological p-wave superconductor on a lattice with N lattice sites is the Kitaev chain [86] with Hamiltonian

$$H = \sum_{j=1}^{N-1} \left[-tc_{j+1}^{\dagger}c_j + \Delta c_j^{\dagger}c_{j+1}^{\dagger} + \text{h.c.} \right] - \sum_{j=1}^N \mu c_j^{\dagger}c_j, \quad (1.53)$$

where c_j (c_j^{\dagger}) annihilates (creates) an electron at lattice site j , $-t$ is the energy for nearest neighbor hopping, μ is the chemical potential, and Δ describes the superconducting excitation gap for p-wave pairing. We first introduce two auxiliary operators [86]

$$\gamma_{2j-1} = c_j^{\dagger} + c_j \quad (1.54)$$

$$\gamma_{2j} = i(c_j^{\dagger} - c_j) \quad (1.55)$$

for each lattice site j satisfying the Majorana condition $\gamma_l^{\dagger} = \gamma_l$ and obeying the anti-commutation relations $\{\gamma_m, \gamma_l\} = 2\delta_{lm}$ thus squaring to one, $\gamma_l^2 = \mathbb{1}$. A fermion is then described by the combination of two such Majorana operators, $c_j = (\gamma_{2j-1} + i\gamma_{2j})/2$. We note, however, that we can always define these operators, in particular even when there are no Majorana quasiparticles in the system, since the operators in general do not create an eigenstate of the Hamiltonian. For the purposes of this section, we consider only two different examples of parameters [cf. 86] showing two distinct topological phases, and postpone a more detailed discussion to Sec. 2.2.1 after we have introduced the necessary methods for the topological classification of such systems. For the first example, the *topologically trivial* case $t = 0 = \Delta$, $\mu < 0$, the Hamiltonian in the Majorana operators reduces to

$$H_{\text{triv}} = -\frac{i}{2} \sum_{j=1}^N \mu \gamma_{2j-1} \gamma_{2j} + \text{const.}, \quad (1.56)$$

such that the Majorana operators belonging to the same lattice site are coupled (Fig. 1.6a). In contrast, in the more interesting *topological* case $t = \delta > 0$, $\mu = 0$, one finds the Hamiltonian

$$H_{\text{top}} = it \sum_{j=1}^{N-1} \gamma_{2j} \gamma_{2j+1}, \quad (1.57)$$

which now couples Majorana operators from adjacent lattice sites (Fig. 1.6b). However, since the wire has two ends at $j = 1$ and $j = N$, this leads to the Majorana operators γ_1 and γ_{2N}

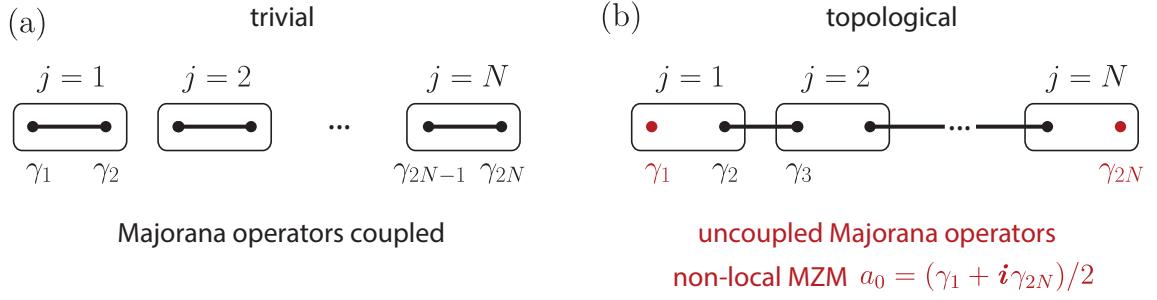


Figure 1.6: Special examples for the Kitaev chain [86] in the (a) topologically trivial regime for $t = 0 = \Delta$, $\mu < 0$ and (b) topological regime for $t = \Delta > 0$, $\mu = 0$. Each ellipse represents a lattice site j , the dots are Majorana operators, and the lines represent couplings in the Hamiltonian. In the topological case (b), there are two uncoupled Majorana operators (red) that can form a highly non-local Majorana zero mode.

remaining uncoupled and hence they do not appear in the Hamiltonian. This then allows the definition of a strongly non-local fermion quasiparticle operator [86]

$$a_0 = \frac{1}{2} (\gamma_1 + i\gamma_{2N}) , \quad (1.58)$$

which has contributions from the two Majorana operators at the wire ends and does not appear in the Hamiltonian, thus describing a zero-energy excitation. We discuss later in Sec. 2.2.1 that the Kitaev chain can be derived as a lattice approximation of a continuum p-wave superconductor. It turns out that the presence of Majorana excitations is linked to the topological phase, which exists for an extended parameter range. In fact, to change between topological phases, the excitation gap must be closed so that the Majorana modes exist as non-local zero-energy excitations at the boundaries between different topological domains. Thus, the wave function for a Majorana mode in a continuous topological superconducting wire is localized at the wire ends and decays exponentially into the wire. However, the Majorana states described here in the context of condensed matter – apart from the fact that their operators satisfy the Majorana condition – have nothing in common with the fermions predicted by E. Majorana, since they are not elementary particles with a dispersion, but bound states. Therefore, to distinguish between them, one uses the term Majorana zero modes (MZMs) instead of Majorana fermions [255].

As mentioned earlier, a conventional s-wave superconductor is not sufficient to realize a topologically nontrivial phase and MZMs, and it turns out that intrinsic p-wave superconductors are hard to find, making realization of MZMs in such systems unrealistic at the moment [255, 284–287]. To circumvent this problem, hybrid systems have been proposed that are effectively p-wave superconductors under certain conditions [87–89]. For this purpose, a conventional s-wave superconductor such as aluminum is placed on top of a semiconductor with strong spin-orbit coupling, for example, InAs [87–89]. Here, the proximity effect [288–292] induces superconductivity in the semiconductor, where the spin-orbit coupling rotates the spin directions at the Fermi level for certain choice of parameters such that effective p-wave pairing is induced. The third ingredient is a strong Zeeman field, which ensures that effectively only one spin component exists at the Fermi level. This makes it possible to engineer an effective spin-polarized p-wave superconductor that realizes topological superconductivity and thus MZMs for sufficiently large magnetic field. We call such hybrid wires *Majorana wires* or *Rashba wires* in the following, since the considered spin-orbit coupling is of the Rashba type.

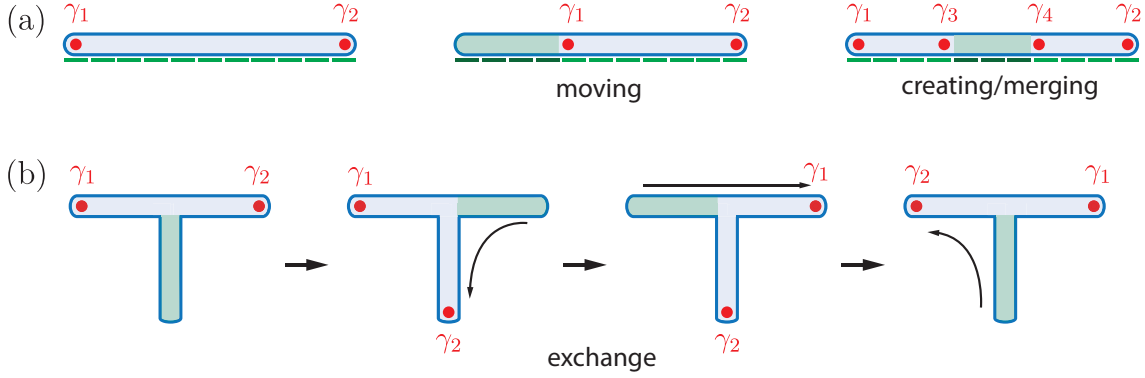


Figure 1.7: Moving and braiding Majorana zero modes (red) in topological superconductor wires by using arrays of gates (green) to adiabatically tune regions of the wire between topological (blue) and trivial phase (green) as described by Alicea et al. (a) By slowly increasing the gate voltage from one end towards the middle of the wire, a MZM can be moved with the edge of the topological region. By tuning the center region of the wire into the trivial regime, two new MZMs at the new edges of the two topological regions can be created. By slowly lowering the potential again, the MZMs can also be fused again. (b) Braiding operation to exchange the MZMs from both ends by using a T-junction. Figure adapted from [79].

1.3.2 Majorana zero modes for quantum computation

Since a pair of Majorana operators γ_1, γ_2 leads to a zero energy state $c_{12} = (\gamma_1 - i\gamma_2)/2$, it creates a degenerate ground state in which information can be encoded, e.g. one can assign the value 0 to the qubit if the level is empty, $|0\rangle$, and 1 if the level is occupied, $c_{12}^\dagger|0\rangle = |1\rangle$. The non-local character of the MZMs, as well as their energy in the middle of the excitation gap, makes them topologically protected, i.e. weak, local perturbations cannot affect the state, which protects it from decoherence. Since fermion parity is preserved in the superconductor, i.e. the condensate contains an even number of electrons, it is necessary to consider four Majorana operators as one qubit [293].

In addition to topological protection and the degenerate ground state, another property of Majoranas makes them particularly interesting as building blocks for qubits – their exotic braiding statistics. Here, braiding refers to moving Majoranas around each other, where the exchange of γ_i and γ_j can be described by the operator [255, 256]

$$B_{ij} = e^{-\pi\gamma_i\gamma_j/4} = \frac{1}{\sqrt{2}}(1 - \gamma_i\gamma_j) \quad (1.59)$$

for which $B_{ij}\gamma_i = \gamma_j B_{ij}$ and $B_{ij}\gamma_j = -\gamma_i B_{ij}$. We again use that, as for all Majorana operators, they square to one, $\gamma_n^2 = 1$, and they anti-commute among each other, $\{\gamma_i, \gamma_j\} = 2\delta_{ij}$. A twofold exchange, $B_{ij}^2 = -\gamma_i\gamma_j$, brings the Majoranas back to the initial position, and if we consider four Majorana operators $\gamma_1, \gamma_2, \gamma_3$, and γ_4 , it turns out that the braiding statistic is non-Abelian, since the order of the exchange matters. For example for braiding γ_1 around γ_2 and γ_2 around γ_3 , one finds $\gamma_1\gamma_3 = B_{12}^2 B_{23}^2 \neq B_{23}^2 B_{12}^2 = \gamma_3\gamma_1$. It is also interesting to note that moving Majorana γ_2 around Majorana γ_3 and then braiding γ_1 around γ_2 yields the same operator as if γ_1 had been moved around γ_3 . For these reasons, braiding can be used to implement a topologically protected NOT gate as described by Sato et al. [255]: Choosing the representation such that the fermion operators for the four Majoranas are given by $c_{12}^\dagger = (\gamma_1 + i\gamma_2)/2$ and

$c_{34}^\dagger = (\gamma_3 + i\gamma_4)/2$, braiding Majorana γ_2 around Majorana γ_3 , as described by B_{23}^2 , transforms the fermion operators according to: $B_{23}^2 c_{12} = c_{12}^\dagger B_{23}^2$ and $B_{23}^2 c_{34} = -c_{34}^\dagger B_{23}^2$, thus turning a qubit from the $|0\rangle$ to the $|1\rangle$ state and vice versa.

Although one can formally construct a description of the exotic non-Abelian braiding statistics, the question arises how to move these bound states in the quasi-one-dimensional systems in the first place. For this purpose, Alicea et al. [79] describe a protocol for performing braiding operations in Majorana wires: Since MZMs occur at the boundaries between topological and trivial domains, one can bring an array of gates into proximity with the wire and adjust which regions of the wire are in the topological phase by applying gate voltages. By adiabatically adjusting the voltages, i.e. slow adjustment such that the system cannot leave the ground state, the boundaries of the topological region and hence the Majoranas can be shifted (Fig. 1.7a). If one slowly creates a trivial region in the middle of the topological region, new MZMs can be created and if one reverses this process, it is also possible to merge Majoranas (Fig. 1.7a). By implementing a T-junction, Alicea et al. [79] also show a protocol for exchanging two Majoranas, which can also be extended to a full braiding process by repeating the procedure a second time with the same orientation [79]. The advantage of this implementation with Majorana wires is not only that the Majoranas can be manipulated easily and precisely using gates, but that these structures can also be extended for larger systems consisting of many qubits [79, 83, 84].

Although the exotic braiding statistics allow for a topologically protected implementation of most of the operations needed for a complete quantum computing protocol, there are a few operations (for example special phase gates and direct evaluation of the eigenvalue of a product of four Majorana operators) for which no topologically protected implementations are known [294–297]. Therefore, topological operations must be supplemented by conventional implementations that may induce decoherence [295].

1.3.3 Challenges for realization and detection of Majorana zero modes

Not long after the proposal of the Majorana wire, the construction of such hybrid systems could be realized in the laboratory [90]. This raises the question of how to experimentally demonstrate that MZMs are indeed present in a given wire. The first signatures that could be experimentally detected are based on the zero-energy nature of Majorana modes: If one connects a Majorana wire at one end to a lead and measures the conductance through the wire between the lead and the superconductor as a function of an applied bias voltage (Fig. 1.8a), one expects in the trivial case (red) that the conductance is zero for a small bias voltage and increases only when the voltage reaches the magnitude of the excitation gap, i.e., when the incoming electrons have enough energy to tunnel through the quasiparticle states above the excitation gap. In contrast, in the topological case (blue) with MZMs, there is a state at zero energy through which tunneling can occur even without bias voltage. A signature for MZMs is therefore a zero-bias conductance peak in the middle of the excitation gap, which theoretically has a height of $2e^2/h$ [89, 298]. While the appearance of the zero-bias conductance peak at the expected transition to the topological phase could actually be observed, its full height was not, which has been explained by disorder and imperfect couplings to the lead [90–93].

Another signature can be observed when the superconductor is not grounded and the wire is placed between two weakly coupled leads (Fig. 1.8b) so that it is in the Coulomb blockade regime, i.e. due to repulsion with the N_0 electrons in the wire, an additional charge energy E_c is necessary to bring another electron into the wire. If a gate is brought into proximity of the wire, the energy of the electrons in the wire can be lowered by $eV_g N_0$ by applying a gate voltage V_g . Whenever the contribution of the charging energy cancels the energy due to the gate, such that

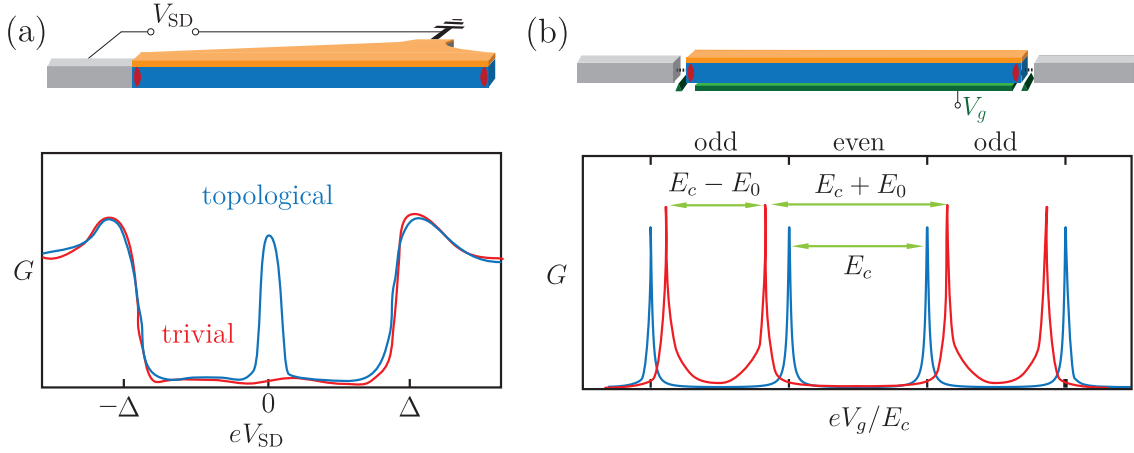


Figure 1.8: Signatures of Majorana zero modes based on the zero energy property (blue) and for comparison topologically trivial counterparts in the absence of a zero energy state (red): (a) When measuring the conductance through a grounded wire with Majorana zero modes as a function of the source-drain voltage V_{SD} , a zero bias conductance peak is found. (b) Measuring the conductance through the wire in Coulomb-blockade as a function of the external gate voltage, one observes a vanishing even-odd conductance resonance spacing difference if Majorana zero modes are present. The upper panels depict a sketch of the system with leads (gray) and Majorana wire consisting of semiconductor (blue) and superconductor (orange). If the wire is in the topological phase, Majorana zero modes (red ellipses) are present at both ends of the wire. For the wire in Coulomb blockade in panel (b), additionally a gate (green) is placed in proximity of the wire.

it is energetically equivalent whether there are N_0 or $N_0 + 1$ electrons in the wire, one observes a conductance resonance. If the voltage is increased beyond the resonance, $N_0 + 1$ electrons are henceforth in the ground state of the wire. However, the condensate of the superconductor consists of Cooper pairs, i.e., it always has an even number of electrons, which means that in case of an odd number of electrons in the wire, the lowest level E_0 must also be occupied in the ground state. This results in the conductance resonances having a spacing $E_c \pm E_0$, where the positive sign applies to an even number of particles, and the negative sign is valid for an odd number of particles. Thus, in the trivial case where E_0 corresponds to the size of the excitation gap, there are significant differences between the spacings of the conductance resonances for even and odd particle number parity (red). This even-odd spacing difference vanishes in the topological case with MZMs (blue) where $E_0 = 0$, which has also been observed experimentally [95, 130].

The problem with these signatures is that they occur whenever a zero energy state exists in the excitation gap, regardless of whether it is a topological MZM, or just a trivial Andreev bound state (ABS). These trivial ABS can also be pinned at zero energy for extended parameter ranges [111–128, 299] and occur, for example, when the boundary potential between wire and leads slowly decays into the wire [112]. Thus, these energy-based signatures cannot provide experimental evidence for the presence of MZMs, but can only rule them out. It is therefore desirable to consider signatures that are instead based on the non-locality of MZMs and are thus harder to be mimicked by trivial states [129, 132, 300, 301].

Currently, another major problem – if not the main problem – is the occurrence of disorder in the hybrid systems [102–105]. Here, strong disorder can not only destroy the topological

phase and the excitation gap [107], but it can also induce trivial zero modes and ABSs in the wire [102, 104, 108–111], giving the appearance of a topological phase in the trivial region. The extent of this problem is also evident from a recent study by the Microsoft Quantum group [106]: By using a sophisticated protocol of multi-terminal measurements to determine whether a topological phase exists in the device with high probability [302], no topological phase can be detected in most of the elaborately fabricated hybrid wires. Theoretical simulations of 50 disorder profiles have shown that only in ten cases the existence of a topological phase can be demonstrated⁴ [106]. This means that disorder alone causes about 80% of the fabricated wires to be unusable. This makes it seem hopeless to fabricate entire arrays of wires to enable quantum computing on a large scale, unless a way is found to control the disorder and increase the yield during fabrication. In addition, they observe large variations in the location of the topological phase and the size of the excitation gap, such that a method for stabilizing MZMs would be beneficial.

1.3.4 Amplitude of coherent transmission

To address the issues described in the previous section, we here consider the amplitude of coherent transmission through a Majorana wire, a metric that probes the non-locality of MZMs, is able to distinguish MZMs from trivial ABSs, and whose maximization can be used to counteract disorder effects. For this purpose, we consider the electron interferometer setup shown in Fig. 1.9a, where a Majorana hybrid wire is embedded in one arm of the interferometer, the reference arm contains a normal wire, and a flux Φ is applied through the interferometer loop. It is found that the current through the interferometer in first order of interference is given by [303]

$$I = \frac{e^2}{h} \sum_{\sigma\sigma'} |T_{\sigma\sigma'}^{\text{LR}}|^2 = \frac{e^2}{h} \left\{ \sum_{\sigma\sigma'} |T_{\sigma\sigma'}|^2 + 2|T^{\text{ref}}|^2 + I_{\text{intf}} \right\} \quad (1.60)$$

$$I_{\text{intf}} = \frac{2e^2}{h} |T^{\text{ref}}| |T_{\uparrow\uparrow} + T_{\downarrow\downarrow}| \cos(\Delta\varphi + \beta) . \quad (1.61)$$

One can therefore determine the coherent transmission amplitude $A = |T_{\uparrow\uparrow} + T_{\downarrow\downarrow}|$ from the amplitude of the current oscillations when the flux Φ is varied. Here $|T_{\sigma}|^2$ is the probability that a particle with spin σ tunnels through the Majorana wire. We assume that the wire is in the Coulomb blockade regime with a large charging energy E_c , such that it contains a fixed number of electrons in the ground state and additional electrons cannot enter away from a charge degeneracy point due to energetic reasons. This charging energy is indirectly proportional to the wire length L , i.e. $E_c \sim 1/L$. Away from the charge degeneracy points, transport is only possible via cotunneling processes, where an electron enters the wire only in a virtual state and then quickly leaves again – compatible with the energy-time uncertainty. The probability of an electron tunneling from the lower interferometer arm into the left end of the wire then depends on the weight of the wave function at the left end of the wire y_L , and analogously for the right end y_R . Thus, for tunneling through a dominant level with wave function φ , the transmission amplitude is approximately given by $A \sim |\varphi(y_L)\varphi(y_R)|/E_c$ [129] and $E_c \sim 1/L$ as described above. Hence, when tunneling through an extended trivial level $\varphi_{\text{triv}}(y_{R/L}) \sim 1/\sqrt{L}$, we expect a small transmission amplitude that is independent of the wire length. On the other hand, if the transmission occurs through a topological Majorana level exponentially localized at both wire ends, i.e. $\varphi_{\text{top}} \propto \chi_L + \chi_R$ with $\chi_L(y) \sim e^{-y/\xi}/\sqrt{\xi}$ and $\chi_R(y) \sim \chi_L(L - y)$, the transmission

⁴It may be present but only when computing the Pfaffian topological invariant (see also Sec. 2.1).

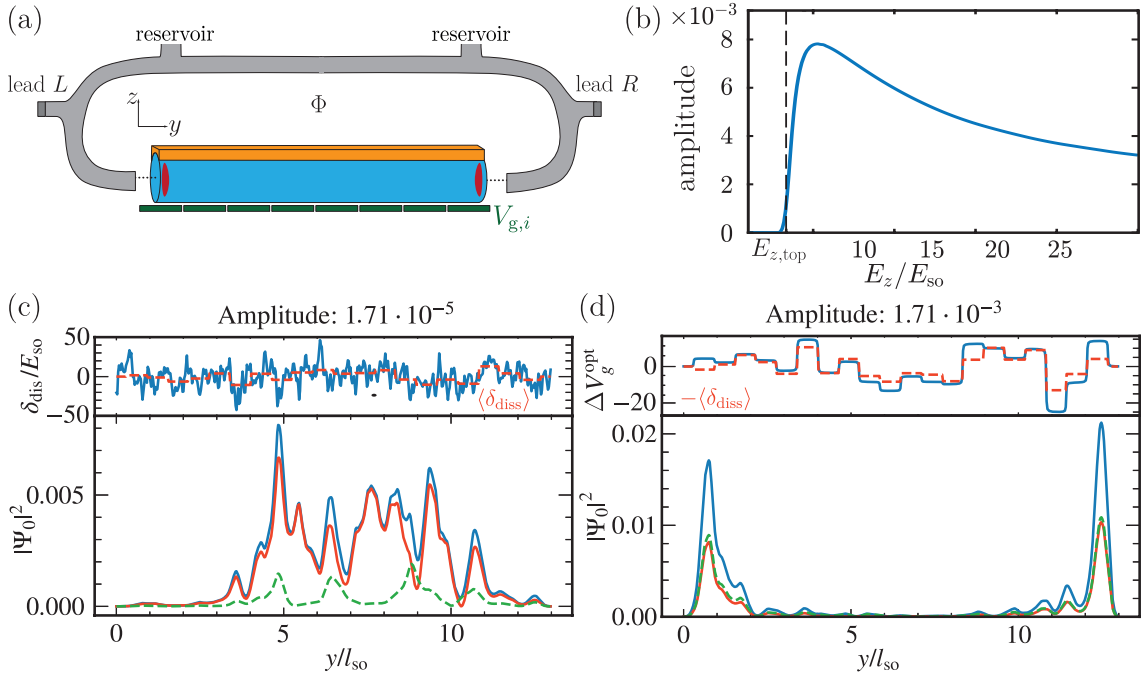


Figure 1.9: Example results of our study of Majorana wires (for details see Ch. 2). (a) Interferometer setup containing the Majorana wire (semiconductor: blue, superconductor: orange, array of gates: green) in one arm (adapted from [131]). (b) Magnetic field dependence of the amplitude of coherent transmission through the interferometer. One observes a maximum of the transmission amplitude soon after entering the topological phase by tuning the Zeeman field E_z above $E_{z,\text{top}}$. (c) Full probability function $|\Psi_0|^2 = |u_0|^2 + |v_0|^2$ (lower panel, blue), hole contribution $|v_0|^2$ (red), and electron contribution $|u_0|^2$ (green) of the lowest energy Bogoliubov eigenstate in a Majorana wire with strong disorder (upper panel, average disorder over the regions of the individual gates shown in red). The disorder completely destroyed the Majorana modes ($|v_0| \neq |u_0|$) such that the transmission amplitude has a very small value of $1.71 \cdot 10^{-5}$. (d) By using optimized gates V_g^{opt} (upper panel) determined by maximizing the transmission amplitude with the CMA-ES machine learning algorithm the Majorana zero modes are fully restored ($|u_0| = |v_0|$) and their localization at the wire ends is improved even compared to a clean wire. Optimization of only 20 gates yields full restoration of the Majorana modes and an improvement of the amplitude of two orders of magnitude. In the upper panel, we show the difference ΔV_g^{opt} of the optimized gate voltage V_g^{opt} and the gate voltage obtained for optimizing in case of a clean wire. This difference shows how the machine learning algorithm learns the disorder profile where the negative averaged disorder over the gates is shown in red, in very good agreement with ΔV_g^{opt} .

amplitude is proportional to the wire length $A_{\text{top}} \sim L/\xi$. Here, ξ is the coherence length of the MZMs, which determines the strength of localization and hence the transmission amplitude for transport through the Majorana level, and which is independent of the length of the wire. In Fig. 1.9b, the transmission amplitude is shown as a function of the external Zeeman field. If one starts with a small Zeeman field E_z so that the hybrid wire is initially in the trivial phase, one observes a small amplitude. If one then increases E_z , the transition to the topological regime takes place (vertical line), leading to the formation of the MZMs, for which the correlation length

can be extracted. Shortly after entering the topological phase, the correlation length becomes very short, and it then increases proportionally to the Zeeman field deeper in the topological phase [131]. The resulting maximum of the transmission amplitude as a function of the Zeeman field soon after entering the topological phase has recently been observed in an experiment by Whiticar et al. [130]. Following our reasoning, we also expect the height of this maximum to depend linearly on the wire length L .

For the same setup, Drukier et al. [304] studied the phase of the complex transmission $T_{\sigma\sigma}$ based on which they proposed another signature of MZMs: For transport through a trivial level, the phase is expected to jump by a value of π when the gate voltage V_g is changed from one conductance resonances to the next one – a so-called *phase laps* occurs. The authors of Ref. [304] were able to show theoretically that these jumps are forbidden by the properties of Majorana wave functions, and therefore phase lapses do not occur for transport through MZMs. Unfortunately, the absence of phase lapses is not limited to the topological regime [304, 305]. In the recent experiment by Whiticar et al. also some results for phase differences between resonances are shown, but they conclude that the results for the transmission phase in the Coulomb valleys are “not currently understood” [130]. The transmission phase is apparently more difficult to measure, and it is assumed that it might also be sensitive to the influence of higher levels.

We argued that the magnitude of the transmission amplitude depends strongly on the non-local character of the MZMs such that by maximizing the transmission amplitude also the desired features of the MZMs are improved. We exploit this by placing an array of gates under the Majorana wire (Fig. 1.9a, green) and optimizing the voltages on the gates using the CMA-ES machine learning algorithm to maximize the transmission amplitude [303]. Indeed, if we perform this optimization in a wire with strong disorder that destroyed the MZMs and the topological phase (Fig. 1.9c), we find that the optimized gate voltages are able to reliably restore the MZMs and cancel the disorder (Fig. 1.9d). Such optimization could significantly contribute to solving the problem of the poor yield in the fabrication of Majorana devices and thus be a step towards the realization of quantum computers with Majorana qubits.

1.4 Entanglement entropy in one dimensional fermion chains

In recent years, studies of entropy in quantum systems have gained much attention for quantum information processing [306, 307]. In classical systems, entropy measures the lack of information needed to infer the microstate of a system from a given macrostate [307]. In quantum systems, however, entropy can result from entanglement even if the microstate is completely known [307]. For instance, if we consider a quantum system in a pure state $|\psi\rangle$, such that its density matrix is given by $\rho = |\psi\rangle\langle\psi|$, and partition it into two regions A and B , we can calculate the entanglement entropy between the subregions from the reduced density matrix for region A , using the von Neumann entropy

$$S_1 = -\text{Tr}[\rho_A \ln \rho_A] . \quad (1.62)$$

Here, the reduced density matrix is obtained by tracing out region B : $\rho_A = \text{Tr}_B \rho$. In Fig. 1.10 this is shown for two different situations: For the case where there are no correlations between region A and B (panel a), such that we can write the full state as the product of the states of the regions, $|\psi\rangle = |\psi\rangle_A \otimes |\psi\rangle_B$, the reduced density matrix describes a pure state resulting in an entropy that is zero. On the other hand, if there are correlations between the subsystems (panel b), tracing out region B yields a mixed state that has nonzero entropy [306]. Since this

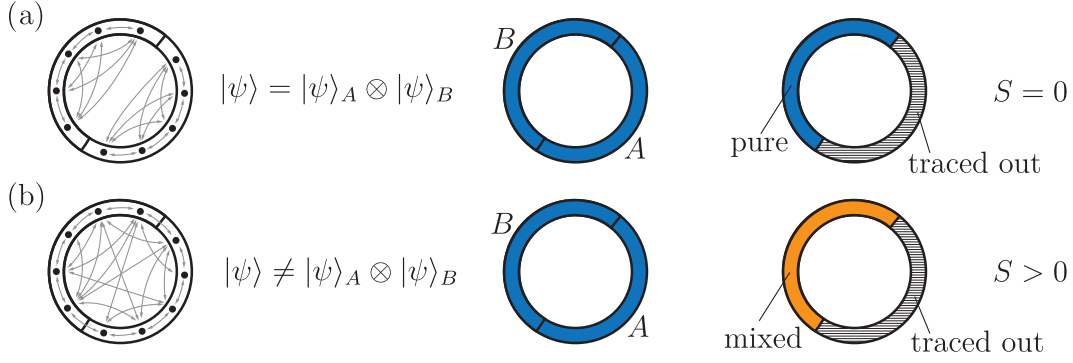


Figure 1.10: Entanglement entropy under a spatial bipartition of the system, a lattice on a ring, in subregions A and B . In the left panel arrows indicate correlations, in the center panel the system is divided into the subregions, and in the right panel, region A is traced out. In (a), there are no correlations between sites from different regions, such that the full state $|\psi\rangle$ is a product of pure states from the subregions $|\psi\rangle = |\psi\rangle_A \otimes |\psi\rangle_B$. Hence, after tracing out region B , the state in the remaining region is still a pure state with zero entropy. In case (b), the subregions are entangled, such that a mixed state and positive entropy are obtained after tracing out region B . A similar figure is shown in Ref. [306].

entropy is a measure of the correlations between subsystems, it determines how information about a sudden change in the system propagates and how the system equilibrates once it has been brought out of equilibrium [136–141, 308]. Being sensitive to correlations in the system, it is also interesting as a signature for phase transitions [309] in many-particle systems.

Since interactions in quantum systems are typically quite short ranged, the spatial entanglement entropy of the ground state usually scales with the size of the boundary of the subregion rather than its volume, which is commonly referred to as an *area law* and has been studied in numerous systems [307]. Instead of this well-studied spatial entanglement entropy, we focus in the following on the less well understood *particle entanglement entropy*, for which subregions are partitioned into subsets of particles rather than spatially.

1.4.1 Particle entanglement entropy

One problem for dividing a system into two partitions based on particles is that the fermions considered here are indistinguishable, and we cannot simply assign labels to the particles in the second quantization notation. It is however possible to consider the properly anti-symmetrized wave function in first quantization and then to trace out the particle positions of one of the partitions [143, 144, 153, 310–312]. If the N fermions are at positions (i_1, i_2, \dots, i_N) , the reduced density matrix for partitioning into n particles in partition A and $N - n$ particles B can be obtained via [161]

$$\rho_n^{(i_1 \dots i_n), (j_1, \dots, j_n)} = \sum_{i_{n+1}, \dots, i_N} \Psi_0^*(i_1, \dots, i_n, i_{n+1}, \dots, i_N) \Psi_0(j_1, \dots, j_n, i_{n+1}, \dots, i_N), \quad (1.63)$$

where the wave function of the state $|\psi\rangle$ is defined as $\Psi_0(i_1, \dots, i_N) = \langle i_1, \dots, i_N | \Psi_0 \rangle$. For a one-dimensional lattice, we will see that the density matrix can alternatively be computed in second quantized notation as an expectation value of field operators [145, 313]

$$\rho_n^{(i_1 \dots i_n), (j_1, \dots, j_n)} = \langle \Psi_0 | c_{i_1}^\dagger \dots c_{i_n}^\dagger c_{j_1} \dots c_{j_n} | \Psi_0 \rangle / \binom{N}{n}, \quad (1.64)$$

where c_i^\dagger (c_i) creates (annihilates) a particle at site i . The particle entanglement entropy is then calculated as $S_1(n) = -\text{Tr}[\rho_n \ln \rho_n]$, and it differs substantially from its spatial counterpart: Unlike spatial entanglement, the particle entanglement entropy is sensitive to interactions and particle statistics [146–157]. For instance, it is zero for free bosons, but scales like $S(n) = \ln \binom{N}{n}$ for non-interacting fermions [152]. Moreover, as the spectrum of the correlation matrix is invariant under unitary transformations, the density matrix is independent of the choice of modes (e.g. real space, momentum, or energy). Like spatial entanglement entropy, particle entanglement entropy is sensitive to many-body localization [314, 315] and telegraphs phase transitions [309, 316]. However, doubts have been raised that the particle entanglement entropy could actually be determined experimentally, since the particles are indistinguishable and thus cannot be labeled, and that it might therefore be useless for quantum information processing [317–319]. But, it has been shown that the entanglement of indistinguishable particles can be mapped onto physically accessible modes [158], as well as that the connection to the correlation function opens up possibilities for experimental measurement of particle entanglement [159, 160, 320].

In addition to obtaining the particle entanglement entropy from its spectrum, the n -RDM also provides access to any n -particle observable. It might also be possible to expand the n -RDM in more accessible reduced density matrices for lower n , similar to the classical case [321, 322]. Therefore, it is reasonable to focus on the 1-RDM as the first step of such an expansion in Ch. 4. Furthermore, in the case $n = 1$, we can perform an analytic bosonization calculation and compare with numerical results.

1.4.2 The J - V model

We focus in the following on N fermions on a one-dimensional lattice with L lattice sites, which are described by the J - V Hamiltonian

$$H = -J \sum_{i=1}^L (c_{i+1}^\dagger c_i + c_i^\dagger c_{i+1}) + V \sum_{i=1}^L n_i n_{i+1} . \quad (1.65)$$

Here J is the nearest-neighbor hopping parameter, V is the interaction energy for particles on neighboring lattice sites, and the operators c_i^\dagger and c_i create and annihilate a particle on lattice site i , respectively. We consider the case of a periodic lattice and measure length in units of the lattice constant. This J - V model is a generic integrable model for spinless fermions which can be mapped to a spinful XXZ model and then solved using a Bethe ansatz [323–326]. This allows us later to map the parameters J and V to those of a continuous Luttinger liquid model that can be solved analytically. Thus, the J - V model is an ideal platform to study particle entanglement entropy.

From mapping the J - V model onto an XXZ chain [323–325], one finds that phase transitions occur at interactions $V/J = \pm 2$. For large negative interactions $V/J < -2$, it is advantageous for the fermions to be on adjacent lattice sites, thus favoring clustering in the ground state in this *phase separated solid* phase. On the contrary, in the *charge density wave* phase for $V/J > +2$ it is energetically very unfavorable to have fermions on adjacent sites, such that an empty site between two fermions is preferred. We are later particularly interested in the case $-2 < V/J < 2$ in the so-called *Luttinger liquid* phase, where the weak interactions allow us to perform an analytic bosonization field theory calculation.

First, we consider an example without interaction term ($V = 0$) for a lattice with $L = 2N$ lattice sites, odd number of particles N , and periodic boundary conditions, $c_{L+i} = c_i$. For this

case, we express the Hamiltonian, Eq. (1.65), using the Fourier transform

$$c_j = \frac{1}{\sqrt{L}} \sum_k e^{ikj} c_k \quad (1.66)$$

$$c_k = \frac{1}{\sqrt{L}} \sum_j e^{-ikj} c_j \quad (1.67)$$

with $k = 2\pi m/L$, $m \in [-N, N)$ yielding the diagonal form of the Hamiltonian

$$H = -2J \sum_k \cos(k) c_k^\dagger c_k . \quad (1.68)$$

For odd N , the ground state is then given by $|\Psi\rangle = \prod_{|k| < k_F} c_k^\dagger |0\rangle$ with Fermi momentum $k_F = \pi(N-1)/L$. Thus, we find the momentum distribution

$$n_k = \langle c_k^\dagger c_k \rangle = \begin{cases} 1 & |k| < k_F \\ 0 & \text{otherwise} \end{cases} . \quad (1.69)$$

This also allows to calculate the 1-RDM

$$\rho_1^{ij} = \frac{1}{N} \langle c_k^\dagger c_k \rangle = \frac{1}{NL} \sum_{kk'} e^{-iki} e^{ik'j} \langle c_k^\dagger c_{k'} \rangle \quad (1.70)$$

$$= \frac{1}{NL} \sum_{|m| < (N-1)/2} e^{-i2\pi m|i-j|/L} \quad (1.71)$$

$$= \frac{1}{NL} \frac{\sin(\pi N|i-j|/L)}{\sin(\pi|i-j|/L)} , \quad (1.72)$$

for which we already know the eigenvalues from Eq. (1.69):

$$\lambda_m = \begin{cases} 1/N & \text{for } |m| \leq (N-1)/2 \\ 0 & \text{otherwise} \end{cases} , \quad m \in [-N, N) . \quad (1.73)$$

From this spectrum we can directly calculate the one-particle entanglement entropy, which is given by

$$S_1 = - \sum_{m=-N}^N \lambda_m \ln \lambda_m = -N \frac{1}{N} \ln \left(\frac{1}{N} \right) = \ln(N) \equiv S_{\text{ff}} . \quad (1.74)$$

We will later start from this case of free fermions and consider an *interaction quantum quench* [135], i.e. a sudden change of the interaction strength, $0 \rightarrow V$. In this case, the time evolution of the particle entanglement entropy contains information about the thermalization of the system after the quench [136–141, 308], where it could be shown that in the thermodynamic limit asymptotically – for long times after the quench – spatial and particle entanglement entropy are equivalent [308].

In Ch. 4 we compute the one-particle entanglement entropy numerically using exact diagonalization: For this purpose, a basis $|\varphi_i\rangle$ of all possible occupations of the N fermions on the L lattice sites is set up in which the Hamiltonian can be represented as a large matrix \mathcal{H} with entries $\mathcal{H}_{ij} = \langle \varphi_i | H | \varphi_j \rangle$. From this, the ground state can then be calculated by diagonalization with the Lanczos algorithm. The reduced density matrix is then calculated from the ground

state and its spectrum is determined to calculate the entanglement entropy. For example, for the case of $N = 2$ fermions on $L = 4$ lattice sites⁵, a basis in occupation number representation is given by $\{|0011\rangle, |0101\rangle, |0110\rangle, |1001\rangle, |1010\rangle, |1100\rangle\}$. The advantage of this approach is that the results are exact since no approximations are needed and the full Hilbert space is considered. However, this also poses a problem since the size of the Hilbert space grows exponentially as $\binom{L}{N}$ with the system size. Only by using efficient encoding of the basis and exploiting all symmetries of the Hamiltonian, we are able to consider systems up to $N = 19$ particles on a supercomputer with 1.5 TB of system memory. For the calculation of the time evolution after a quench, exact diagonalization is possible for systems at half filling with up to $N = 13$ fermions. Therefore, we later also perform approximate calculations which allows us to consider much larger systems and hence enable accurate finite-size scaling to the thermodynamic limit $L \rightarrow \infty$.

1.4.3 Luttinger liquid theory and bosonization

In the following, we introduce the basics for the bosonization calculation performed in Ch. 4, closely following parts of Chapter 2 and 3 of Ref. [162] by S. Eggert. Bosonization allows a model of interacting fermions with linear dispersion to be mapped to a bosonic model whose Hamiltonian is then quadratic in the boson operators and can therefore be solved analytically, unlike the original fourth-order fermionic model. The requirement for this to be a good approximation is that the dynamics of the one-dimensional system are determined by small energy excitations for which we can assume the dispersion of the fermions to be approximately linear around the Fermi points. We therefore begin this introduction with the linearization of the dispersion. This allows to express excitations in terms of small density fluctuations around an average density background which are bosonic in nature, and can be expressed by operators with bosonic commutation relations. After introducing these operators, we show how the Hamiltonian for interacting Fermions with an interaction similar to that of the J - V lattice model can be expressed in the boson operators. As the last part of this section, we briefly discuss the *refermionization*, i.e. the inverse transformation that allows to express fermion field operators in terms of the boson operators, which will be necessary for the computations of correlation functions in Ch. 4.

Linearization of the dispersion

As a first step we assume that there are only low energy excitations, such that we can limit the considerations to momenta close to the Fermi points $k \in [\pm k_F - \Lambda, \pm k_F + \Lambda]$, where $\xi_{k_F} = 0$

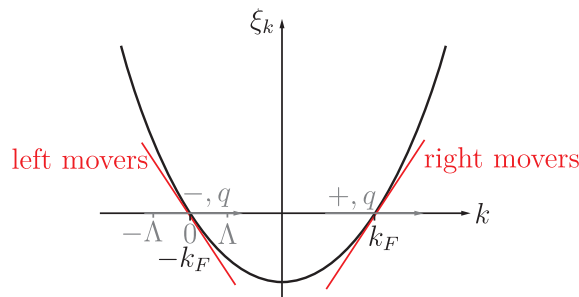


Figure 1.11: Linearization of the dispersion ξ_k around the Fermi points $\pm k_F$ as an approximation for low energy excitations. Momenta \pm, q for right (+) and left (-) movers are measured relative to $\pm k_F$ with cutoffs at $q = \pm \Lambda$.

⁵We discuss this example of $N = 2$ fermions on $L = 4$ sites in detail later in Sec. 4.2.2.

and Λ is a cutoff, for now assuming a generic non-interacting Hamiltonian for spinless fermions

$$H = \sum_k \xi_k c_k^\dagger c_k \approx \sum_{k=-k_F-\Lambda}^{-k_F+\Lambda} \xi_k c_k^\dagger c_k + \sum_{k=k_F-\Lambda}^{k_F+\Lambda} \xi_k c_k^\dagger c_k . \quad (1.75)$$

Here expansion of the dispersion around the Fermi points to linear order (Fig. 1.11) yields

$$\xi_k = \xi_{\pm k_F} + (k \mp k_F) \left. \frac{\partial \xi_k}{\partial k} \right|_{\pm k_F} + \mathcal{O}\left((k - k_F)^2\right) , \quad (1.76)$$

where the Fermi point $-k_F$ with a negative slope gives rise to so-called *left movers* and excitations near $+k_F$ are denoted as *right movers*. The slope of this linear dispersion is given by the Fermi velocity

$$\pm v_F = \left. \frac{\partial \xi_k}{\partial k} \right|_{\pm k_F} . \quad (1.77)$$

One then defines the fermion operators for left ($\alpha = -$) and right ($\alpha = +$) movers $\tilde{c}_{\alpha,q} \equiv c_{\alpha k_F + q}$ where q is the momentum relative to the Fermi points. In these operators, the Hamiltonian, Eq. (1.75), is given by

$$H \approx \sum_{q=-\Lambda}^{\Lambda} v_F q \left(\tilde{c}_{+,q}^\dagger \tilde{c}_{+,q} - \tilde{c}_{-,q}^\dagger \tilde{c}_{-,q} \right) . \quad (1.78)$$

Similarly, we expand the field operators

$$\psi(x) = \frac{1}{\sqrt{L}} \sum_k e^{ikx} c_k \quad (1.79)$$

$$\approx \frac{1}{\sqrt{L}} \sum_q \left(e^{ik_F x} e^{iqx} \tilde{c}_{+,q} + e^{-ik_F x} e^{iqx} \tilde{c}_{-,q} \right) \quad (1.80)$$

$$\equiv e^{ik_F x} \psi_+(x) + e^{-ik_F x} \psi_-(x) \quad (1.81)$$

where the sums over q are again extended from $-\infty$ to ∞ as we assume that there are no high energy excitations such that states with $|q| > \Lambda$ are not occupied anyway [162]. This also ensures that the operators

$$\psi_\alpha(x) = \frac{1}{\sqrt{L}} \sum_q e^{iqx} \tilde{c}_{\alpha,q} \quad (1.82)$$

fulfil the fermionic commutation relations $\{\psi_\alpha^\dagger(x), \psi_{\alpha'}(y)\} = \delta_{\alpha\alpha'} \delta(x-y)$ and $\{\psi_\alpha(x), \psi_{\alpha'}(y)\} = 0$.

Bosonic operators

The concept behind bosonization is to describe an excited low energy state in terms of boson operators. As explained in Sec. 2.2 of Ref. [162], an excited state can be constructed as follows: Starting from the ground state, where all states with $|k| < k_F$ are occupied and the states above k_F are empty, denoted by $|\cdots 11111|000000\cdots\rangle$, we can construct an excited state by first

adding the excess electrons – in case the final state has more electrons above k_F then holes below – and then by shifting the occupied states upwards. To understand what is meant by this, we consider an example for the excited state $|\cdots 111110|0110010\cdots\rangle$, which we can construct from the ground state by first adding the two excess electrons $|\cdots 111111|0000000\cdots\rangle \rightarrow |\cdots 111111|1100000\cdots\rangle$ and then shifting the highest electron by 4 places, and the next two electrons by 2 places, $|\cdots 111111|1100000\cdots\rangle \rightarrow |\cdots 111111|1000010\cdots\rangle \rightarrow |\cdots 111110|0110010\cdots\rangle$. To shift a fermion from l to $l+p$, one can use the operator $c_{l+p}^\dagger c_l$ that consists of a pair of fermionic operators. Hence, the trick is, to express the excited states in terms of these shifting operators, which for the example above means the excited state $|\cdots 111110|0110010\cdots\rangle$ is expressed as one boson in state $p = 4$ and two bosons in state $p = 2$. In addition, the energy for adding the two excess electron also has to be considered. This is of course a simplification meant to motivate the idea for constructing the boson operators. In the following, we will formally introduce the operators and show that the fermionic Hamiltonian can indeed be mapped onto the bosonic problem.

Based on this idea, one considers shifting operators [162]

$$\begin{aligned}\rho_{\alpha,q} &= \sum_{q'} \tilde{c}_{\alpha,q'+q}^\dagger \tilde{c}_{\alpha,q'} \\ \rho_{\alpha,q}^\dagger &= \rho_{\alpha,-q} ,\end{aligned}\tag{1.83}$$

which are, however, not yet the boson operators we are looking for as they do not fulfill boson commutation relations. One needs to be careful when computing the commutator between the shifting operators

$$[\rho_{\alpha,q_1}, \rho_{\alpha',q_2}] = \delta_{\alpha,\alpha'} \sum_{q'} \sum_{q''} \left(\delta_{q',q_2+q''} \tilde{c}_{\alpha,q_1+q'}^\dagger \tilde{c}_{\alpha,q''} - \delta_{q'',q_1+q'} \tilde{c}_{\alpha,q_2+q''}^\dagger \tilde{c}_{\alpha,q'} \right) \tag{1.84}$$

$$= \delta_{\alpha,\alpha'} \sum_{q'} \left(\tilde{c}_{\alpha,q'+q_1}^\dagger \tilde{c}_{\alpha,q'-q_2} - \tilde{c}_{\alpha,q'+q_2+q_1}^\dagger \tilde{c}_{\alpha,q'} \right) , \tag{1.85}$$

because the q' and q'' sums are over infinitely many values such that they do not commute in general. The problem is that after linearizing the dispersion, there are infinitely many occupied states with negative energy. We can resolve this by using that there are only low energy excitations and hence all states with $q < -\Lambda$ are occupied, $n_{q<-\Lambda} = 1$ [162]. Then clearly if $q_1 \neq -q_2$ the sums are converging, and we can simply shift the sum for the second term $q' \rightarrow q' - q_2$ such that the commutator vanishes. If $q_1 = -q_2$, we use $n_{\alpha,q} = \tilde{c}_{\alpha,q}^\dagger \tilde{c}_{\alpha,q} = 1$ for $q < -\Lambda$ to cancel the divergence from the two terms. We can thus compute the commutator for $q_2 \geq 0$ as [162]

$$[\rho_{\alpha,q_1}, \rho_{\alpha',q_2}] = \delta_{\alpha,\alpha'} \delta_{q_1,-q_2} \sum_{q'} (n_{\alpha,q'-q_2} - n_{\alpha,q'}) \tag{1.86}$$

$$= \delta_{\alpha,\alpha'} \delta_{q_1,-q_2} \sum_{q' \geq -\Lambda} (n_{\alpha,q'-q_2} - n_{\alpha,q'}) \tag{1.87}$$

$$= \delta_{\alpha,\alpha'} \delta_{q_1,-q_2} \left(\sum_{q' \geq -\Lambda-q_2} n_{\alpha,q'} - \sum_{q' \geq -\Lambda} n_{\alpha,q'} \right) \tag{1.88}$$

$$= \delta_{\alpha,\alpha'} \delta_{q_1,-q_2} \sum_{-\Lambda-q_2 \leq q' \leq -\Lambda} 1 \tag{1.89}$$

$$= \delta_{\alpha,\alpha'} \delta_{q_1,-q_2} \frac{qL}{2\pi} , \tag{1.90}$$

where we used that $q = 2\pi m/L$, $m \in \mathbb{Z}$. This allows to define the actual boson operators (for $q > 0$) as

$$b_{\alpha,q}^\dagger = \alpha i \sqrt{\frac{2\pi}{qL}} \rho_{\alpha,\alpha q} \quad (1.91)$$

$$b_{\alpha,q} = -\alpha i \sqrt{\frac{2\pi}{qL}} \rho_{\alpha,-\alpha q} , \quad (1.92)$$

which then fulfil the correct commutation relations $[b_{\alpha,q}, b_{\alpha',q'}] = 0$, $[b_{\alpha,q}, b_{\alpha',q'}^\dagger] = \delta_{\alpha\alpha'} \delta_{qq'}$. Here, the complex phase of the prefactor is to some extent arbitrary, but we will see later that the definition chosen here is useful [162].

Hamiltonian expressed in boson operators

In order to express the Hamiltonian in terms of these operators, we consider the commutator of the non-interacting Hamiltonian with the creation operator and find [162]:

$$[H, b_{\alpha,q}^\dagger] = v_F \sum_{\alpha'} \sum_{q'q''} \alpha' q' \alpha i \sqrt{\frac{2\pi}{q'L}} \left[\tilde{c}_{\alpha',q'}^\dagger \tilde{c}_{\alpha',q'} \tilde{c}_{\alpha,\alpha(q''+q)}^\dagger \tilde{c}_{\alpha,q''} \right] \quad (1.93)$$

$$= v_F \sum_{q'q''} \alpha i \sqrt{\frac{2\pi}{q'L}} \left(\delta_{q',\alpha(q''+q)} \alpha q' \tilde{c}_{\alpha,q'}^\dagger \tilde{c}_{\alpha,\alpha q''} - \delta_{q',\alpha q''} \alpha q' \tilde{c}_{\alpha,\alpha(q''+q)}^\dagger \tilde{c}_{\alpha,q'} \right) \quad (1.94)$$

$$= v_F q b_{\alpha,q}^\dagger . \quad (1.95)$$

This commutator together with $[b_{\alpha,q}, b_{\alpha',q'}^\dagger] = \delta_{\alpha\alpha'} \delta_{qq'}$ defines the harmonic oscillator ladder [162]. Therefore, the non-interacting Hamiltonian in the boson operators is given by [162]

$$H = v_F \sum_{q>0} q \left(b_{+,q}^\dagger b_{+,q} + b_{-,q}^\dagger b_{-,q} \right) + \frac{\pi v_F}{L} \left(N_+^2 + N_-^2 \right) , \quad (1.96)$$

where the first term directly follows from the commutation relations, in analogy to a harmonic oscillator, and the last term is the zero point energy with the number operator for excess fermions⁶

$$N_\alpha = \sum_{q>0} \tilde{c}_{\alpha,q}^\dagger \tilde{c}_{\alpha,q} - \sum_{q\leq 0} \tilde{c}_{\alpha,q} \tilde{c}_{\alpha,q}^\dagger . \quad (1.97)$$

It is crucial that we can express the trivial diagonal Hamiltonian for free fermions in terms of the new boson operators, but the true power of the bosonization transformation is revealed for systems involving interactions. As an example, we consider the density-density interaction [162]

$$H_{\text{int}} = \frac{1}{2} \int dx \int d\Delta x \psi^\dagger(x) \psi(x) U(\Delta x) \psi^\dagger(x + \Delta x) \psi(x + \Delta x) \quad (1.98)$$

$$= \frac{1}{2L} \sum_{kk'q} c_k^\dagger c_{k-q} U_q c_{k'}^\dagger c_{k'+q} \quad (1.99)$$

⁶As in the simple example discussed before, this number of additional fermions that need to be added in the first step before the particles can be shifted, is given by the difference of the number of fermions above the Fermi level and the number of holes below. For the purpose of this section, we do not give a detailed derivation for this term and refer the reader to section 2 of Ref. [162].

with $U_q = \int d\Delta x e^{iq\Delta x} U(\Delta x)$. If we again only consider low energy excitations and, in addition, only small momentum transfer such that the interaction never scatters left-movers into right-movers or vice versa, we can write

$$\sum_k c_k^\dagger c_{k\pm q} \approx \rho_{+,\mp q} + \rho_{-,\mp q} . \quad (1.100)$$

For short range interactions, one often neglects the momentum dependence of the interaction U_q for terms involving ρ operators with the same α , and denotes the interaction strength by g_4 . Similarly, for terms with ρ operators from different α denoted as g_2 , which are both independent of q [162]. The interaction Hamiltonian can then be expressed in the shifting operators, Eq. (1.83), as

$$H_{\text{int}} \approx \frac{1}{2L} \sum_q [g_4 (\rho_{+,q} \rho_{+,-q} + \rho_{-,q} \rho_{-,-q}) + g_2 (\rho_{+,q} \rho_{-,-q} + \rho_{-,q} \rho_{+,-q})] . \quad (1.101)$$

Using the correspondence between density and boson operators,

$$\rho_{\alpha,\alpha q} = \begin{cases} -\alpha i \sqrt{|q|L/2\pi} b_{\alpha,|q|}^\dagger & \alpha q > 0 \\ \alpha i \sqrt{|q|L/2\pi} b_{\alpha,|q|} & \alpha q < 0 \end{cases} \quad (1.102)$$

$$N_\alpha = \rho_{\alpha,0} - \langle \rho_{\alpha,0} \rangle , \quad (1.103)$$

the Hamiltonian for the density-density interactions can be finally written as [162]

$$H_{\text{int}} \approx \frac{1}{2\pi} \sum_{q>0} [qg_2 (b_{+,q} b_{-,q} + b_{+,q}^\dagger b_{-,q}^\dagger) + qg_4 (b_{+,q}^\dagger b_{+,q} + b_{-,q}^\dagger b_{-,q})] \quad (1.104)$$

$$+ \frac{g_4}{L} (N_+^2 + N_-^2) + \frac{2g_2}{L} N_+ N_- + \text{const.} \quad (1.105)$$

The amazing aspect of this result is that we managed to map the Hamiltonian $H + H_{\text{int}}$, which is of fourth order in the fermion operators, to a bosonic Hamiltonian, which is only quadratic in the boson operators and can thus be diagonalized analytically. For this reason, bosonization is a very powerful tool for solving one-dimensional, interacting systems, which we will make use of in Ch. 4 to study a continuum approximation of the J - V model.

Refermionization

To evaluate correlation functions, which we use for the calculation of the RDM, we also need an expression for the fermion field operators in terms of the boson operators. For this purpose, one first defines bosonic field operators [162]

$$\phi_\alpha(x) = - \sum_{q>0} \sqrt{\frac{2\pi}{qL}} e^{-q\eta/2} [e^{i\alpha qx} b_{\alpha,q} + e^{-i\alpha qx} b_{\alpha,q}^\dagger] . \quad (1.106)$$

To motivate this expression, we consider the inverse Fourier transform of the normal ordered fermionic density operator for α -movers, Eq. (1.82),

$$\int dx e^{ikx} : \psi_\alpha^\dagger(x) \psi_\alpha(x) : = \frac{1}{L} \int dx \sum_{k'k''} e^{ikx} e^{-ik''x} e^{ik'x} : \tilde{c}_{\alpha,k''}^\dagger c_{\alpha,k'} : \quad (1.107)$$

$$= \sum_{k'} : c_{\alpha,k+k'}^\dagger c_{\alpha,k'} : \quad (1.108)$$

$$= \begin{cases} -i\alpha \sqrt{\frac{|k|L}{2\pi}} b_{\alpha,|k|}^\dagger & \alpha k > 0 \\ N_\alpha & \alpha k = 0 \\ i\alpha \sqrt{\frac{|k|L}{2\pi}} b_{\alpha,|k|} & \alpha k < 0 \end{cases} . \quad (1.109)$$

Here, normal ordering means subtracting its ground state expectation value from the operator to remove possible divergencies (which can occur when removing the cutoff Λ as discussed earlier), $: f(x) : = f(x) - \langle f(x) \rangle$. Using the inverse transform of Eq. (1.109), one finds [162]

$$: \psi_\alpha^\dagger(x) \psi_\alpha(x) : = \frac{1}{L} \sum_{q>0} \sqrt{\frac{qL}{2\pi}} \left(\alpha i b_{\alpha,q} e^{i\alpha qx} - i\alpha b_{\alpha,q}^\dagger e^{-i\alpha qx} \right) + \frac{N_\alpha}{L} . \quad (1.110)$$

Therefore, the spatial derivative of the boson field yields

$$\frac{-\partial_x \phi_\alpha(x)}{2\pi} + \frac{N_\alpha}{L} = : \psi_\alpha^\dagger(x) \psi_\alpha(x) : , \quad (1.111)$$

such that the free Hamiltonian can be conveniently written as

$$H = v_F \int dx \sum_\alpha \pi \left(\frac{-\partial_x \phi_\alpha(x)}{2\pi} + \frac{N_\alpha}{L} \right)^2 \quad (1.112)$$

$$= v_F \sum_{q>0} q \left(b_{+,q}^\dagger b_{+,q} + b_{-,q}^\dagger b_{-,q} \right) + \frac{\pi v_F}{L} \left(N_+^2 + N_-^2 \right) . \quad (1.113)$$

We can then further use the boson fields to express the fermion field operators as an exponential of ϕ_α [162]

$$\psi_\alpha(x) = \frac{\chi_\alpha}{\sqrt{2\pi\eta}} e^{i(\varphi_{0,\alpha} + \alpha \frac{2\pi x}{L} N_\alpha)} e^{-i\alpha \phi_\alpha(x)} , \quad (1.114)$$

where the first factor contains the so-called Klein factor χ_α that ensures fermionic commutation relations and for which $\chi_\alpha^\dagger \chi_\alpha = 1$, and the second factor contains the zero modes with $[\varphi_{0,\alpha}, N_\alpha] = -\delta_{\alpha\alpha'} i$.

One can check (for details, see Appendix C.1) that this definition is indeed consistent, by using $\psi_\alpha(x)$ from Eq. (1.114) and the expression for $\psi_\alpha^\dagger(x') \psi_\alpha(x') = \rho_\alpha(x')$ from Eq. (1.110)⁷ to show that $\psi_\alpha(x)$ indeed affects the density at position $x' = x$ according to [162]

$$[\psi_\alpha(x), \psi_\alpha^\dagger(x') \psi_\alpha(x')] = \delta(x - x') \psi_\alpha(x) . \quad (1.115)$$

One could further also check the validity of the normalization factor by checking the fermion commutation relations between the ψ_α operators, which is shown in Ref. [162].

Equipped with the ability to bosonize a fermion Hamiltonian with density-density interactions, to diagonalize it, and then also to compute correlation functions by refermionization, we are able to analytically compute the 1-RDM for the J - V Model with weak interactions in the Luttinger liquid phase in Ch. 4.

⁷Normal ordering does not affect the commutator, as a ground state expectation value commutes with any operator.

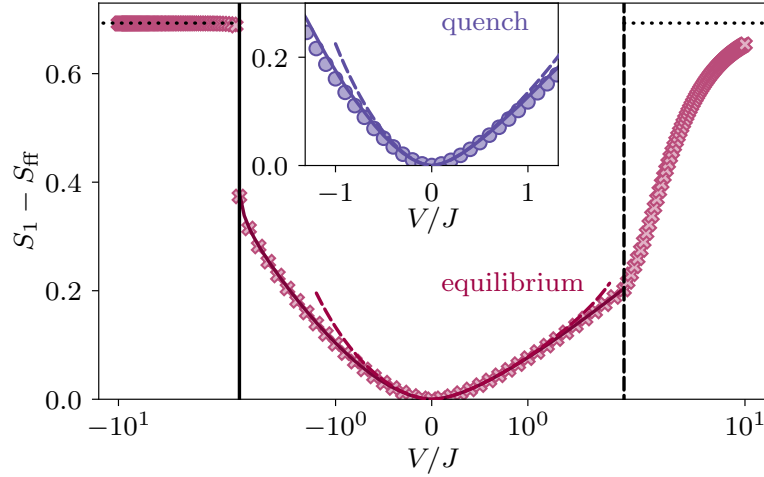


Figure 1.12: Interaction dependence of the one particle von Neumann entanglement entropy S_1 obtained numerically from the J - V model and from an effective low energy Luttinger liquid calculation (dashed lines) with a fixed interaction cutoff. Here, S_{ff} is the entropy for free fermions. The main panel depicts the equilibrium ground state entropy with numerical data from DMRG for a system of $L = 102$ lattice sites at half filling (crosses). The solid line represents finite size scaling of numerical data to the thermodynamic limit. The excellent agreement with the finite size DMRG data shows that the system with $N = 51$ fermions is large enough to describe the thermodynamic limit accurately in the whole LL phase. The inset depicts finite size exact diagonalization results for $N = 12$ fermions on $L = 24$ sites after an interaction quantum quench (circles) in the asymptotic steady state. The solid line shows the thermodynamic limit of the numerical data obtained from finite size scaling the time averaged one particle entropy (circles) after the interaction quench. The dashed line is the result of non-equilibrium bosonization using the same value of the interaction cutoff as in the main panel. (Figure taken from Ref. [327])

Comparison to J - V model

We will see later that to describe the very short-ranged interaction of the J - V model, we need to introduce an interaction cutoff in the bosonization calculation [163, 327], which cannot be determined by the analytical derivation. However, we can determine the cutoff by comparing with numerical solutions. In Fig. 1.12 we show as an example some results for the von Neumann one-particle entanglement entropy, $S_1 - S_{\text{ff}}$, as a function of the interaction strength V/J , where S_{ff} is the entropy for free fermions calculated earlier. In the main panel, the entropy for $N = 51$ particles on $L = 102$ lattice sites (crosses), obtained numerically by approximate methods, is shown together with the finite-size scaled entropy for the thermodynamic limit (solid line). The excellent agreement shows that we are able to numerically consider systems that are close to the thermodynamic limit. The dashed line in the Luttinger liquid phase (between the black vertical lines) is obtained by the analytical bosonization calculation with fixed cutoff and is in excellent agreement with the numerical results for small interactions. In addition, the entropy clearly indicates the phase transitions of the J - V model (vertical lines) by a change in slope. Further, the inset shows the asymptotic entropy long after an interaction quantum quench from free fermions to interactions V/J . The numerical results here are obtained using exact diagonalization for $N = 12$ particles, and we find similar agreement with the analytical results for small interactions.

1.5 Structure of this thesis and publications

In Chapter 2, we study Majorana zero modes in topological superconductors. We start with an introduction to topological superconductivity (Sec. 2.1), followed by a detailed introduction to Majorana zero modes and their proposed realizations (Sec. 2.2). In Sec. 2.3, we describe the setup and the scattering matrix formalism that is used for determining the transmission amplitude through a Coulomb blockaded Majorana wire embedded into one arm of an Aharonov-Bohm interferometer. We use this scattering matrix formalism in Sec. 2.4 to study the magnetic field and wire length dependence of the coherent transmission amplitude, where we explain the occurrence of an amplitude maximum at the onset of the topological regime based on the localization properties of the Majorana zero modes. We published the results of this section in *Transmission amplitude through a Coulomb blockaded Majorana wire*, Matthias Thamm and Bernd Rosenow, Phys. Rev. Research **3**, 023221 (2021) [131]. Based on this, in Sec. 2.5, we use the transmission amplitude as a metric for optimizing an array of gates using the CMA-ES machine learning algorithm for Majorana wires with strong disorder. We show that the algorithm is capable of learning disorder profiles, stabilizing Majorana zero modes, and even restoring Majorana modes that were completely destroyed by disorder. The results of this section are contained in the preprint *Machine learning optimization of Majorana hybrid nanowires*, Matthias Thamm and Bernd Rosenow, arXiv:2208.02182 [303] (accepted for publication in Physical Review Letters).

In Chapter 3, we study deep neural networks with the help of random matrix theory. We start with an introduction into neural network architectures and training in Sec. 3.1. In Sec. 3.2, we use random matrix theory as a zero-information hypothesis to locate stored information in the weights of deep neural networks, where we show that the weights after training are predominately random, and that the information is encoded in the largest singular values and corresponding vectors of the weight matrices only. These results are published in *Random matrix analysis of deep neural network weight matrices*, Matthias Thamm, Max Staats, and Bernd Rosenow, Phys. Rev. E **106**, 054124 (2022) [328]. We extend the analysis in Sec. 3.3 to networks trained with label noise, identify a separation between noise and information in the weights of suitably trained networks, and propose a filtering algorithm to partially remove the influence of the noise from the weights. These findings are published as the preprint *Boundary between noise and information applied to filtering neural network weight matrices*, Max Staats, Matthias Thamm, and Bernd Rosenow, arXiv:2206.03927 (2022) [329].

Chapter 4 discusses the particle entanglement entropy in one-dimensional chains of interacting fermions and is structured as follows: We start with a short introduction and a discussion of the Schmidt decomposition used for efficiently computing entanglement entropy in Sec. 4.1. We then introduce the J - V model and discuss the phase diagram in Sec. 4.2. In the Luttinger liquid phase, we perform a bosonization calculation to analytically determine the reduced one-particle density matrix in Sec. 4.3. To make a connection to the lattice model, we perform large scale numerical computations for which we describe the details in sections 4.4 (exact diagonalization) and 4.5 (approximate DMRG). Finally, in Sec. 4.6, we compare the analytical field theory calculation to the numerical results scaled to the thermodynamic limit. The results of this chapter are published in *One-particle entanglement for one dimensional spinless fermions after an interaction quantum quench*, Matthias Thamm, Harini Radhakrishnan, Hatem Barghathi, Bernd Rosenow, and Adrian Del Maestro, Phys. Rev. B **106**, 165116 (2022) [327].

2. Majorana zero modes in topological superconductors

2.1 Topological superconductors: BdG Hamiltonian, topological invariants, and zero modes

By cooling a superconductor below the critical temperature, a phase transition to the superconducting phase occurs, in which electrons form Cooper pairs [330] by a weakly attractive interaction. Since these pairs of electrons have an integer spin, they are no longer subject to the Pauli principle and can thus, similar to Bose-Einstein condensation [331, 332], form a condensate in which they are in the same quantum mechanical state and can be described by a common wave function. This ensures that scattering at local scattering centers is energetically unfavorable and thus causes the vanishing of the electrical resistance [333]. For the excitation of a quasi-particle it is necessary to break up a Cooper pair, for which the binding energy Δ is needed, such that superconductors are characterized by an excitation gap [333]. Under certain conditions, however, zero-energy excitations can be present in the middle of the gap, which can be caused by disorder [108], multiband effects [109, 110], the Condo effect [334], and Andreev-bound states [112, 335] (see section 2.2.4). In addition, there are so-called topological superconductors which, under certain conditions, host Majorana zero modes in the topologically non-trivial phase. These are robust against external disturbances, since they are connected to the presence of the topological phase, which exists over an extended parameter range.

Following the discussion by Sato et al. [255], we describe superconductivity starting from a generic Hamiltonian, which contains an attractive pair interaction. Using a mean-field approximation, we can draw conclusions about ground states and excited states. Furthermore, we consider how trivial phases can be distinguished from topological phases and how they can be formally classified.

We consider a general one-band effective Hamiltonian [255, 284, 336]

$$H = \sum_{\mathbf{k}} \sum_{\sigma\sigma'} h_{\sigma\sigma'}(\mathbf{k}) c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma'} + \frac{1}{2} \sum_{\mathbf{k}\mathbf{k}'} \sum_{\substack{\sigma_1\sigma_2 \\ \sigma_3\sigma_4}} V_{\sigma_1\sigma_2\sigma_3\sigma_4}(\mathbf{k}, \mathbf{k}') c_{-\mathbf{k}\sigma_1}^\dagger c_{\mathbf{k}\sigma_2}^\dagger c_{\mathbf{k}'\sigma_3} c_{-\mathbf{k}'\sigma_4} \equiv H_0 + H_{\text{int}} , \quad (2.1)$$

where $c_{\mathbf{k}\sigma}^\dagger$ ($c_{\mathbf{k}\sigma}$) is the creation (annihilation) operator for an electron with momentum \mathbf{k} and spin σ , $h_{\sigma\sigma'}(\mathbf{k})$ is the band Hamiltonian without the interaction¹, and energy is measured relative to the chemical potential. Here, $V_{\sigma_1\sigma_2\sigma_3\sigma_4}$ is the pair interaction which causes the weakly attractive interaction between electrons, and we focus on the case of pairing of electrons with opposite momenta, so that the total momentum of the Cooper pair vanishes. Because of the fermionic commutation relations $\{c_{\mathbf{k}\sigma}, c_{\mathbf{k}'\sigma'}\} = 0 = \{c_{\mathbf{k}\sigma}^\dagger, c_{\mathbf{k}'\sigma'}^\dagger\}$, $\{c_{\mathbf{k}\sigma}^\dagger, c_{\mathbf{k}'\sigma'}\} = \delta_{\mathbf{k}\mathbf{k}'}\delta_{\sigma\sigma'}$, the pair interaction satisfies [255, 336]

$$V_{\sigma_1\sigma_2\sigma_3\sigma_4}(\mathbf{k}, \mathbf{k}') = -V_{\sigma_2\sigma_1\sigma_3\sigma_4}(-\mathbf{k}, \mathbf{k}') = -V_{\sigma_1\sigma_2\sigma_4\sigma_3}(\mathbf{k}, -\mathbf{k}') = V_{\sigma_4\sigma_3\sigma_2\sigma_1}(\mathbf{k}', \mathbf{k}) . \quad (2.2)$$

We now assume that Cooper pairs are formed by electrons with (\mathbf{k}, σ_1) and $(-\mathbf{k}, \sigma_2)$, so that operators $c_{\mathbf{k}\sigma_1} c_{-\mathbf{k}\sigma_2}$ have a finite ground state expectation value [284]. To treat the Hamiltonian in a mean-field approximation, we define operators $(\delta c^2)_{\mathbf{k}, \sigma_1\sigma_2} = c_{\mathbf{k}\sigma_1} c_{-\mathbf{k}\sigma_2} - \langle c_{\mathbf{k}\sigma_1} c_{-\mathbf{k}\sigma_2} \rangle$ and

¹The spin dependence can for example be caused by spin-orbit coupling.

neglect quadratic fluctuations around this expectation value, i.e. terms $\mathcal{O}[(\delta c^2)^2]$. Inserting the operators into the interaction term of the Hamiltonian yields

$$H_{\text{int}} = \frac{1}{2} \sum_{\mathbf{k}\mathbf{k}'} \sum_{\substack{\sigma_1\sigma_2 \\ \sigma_3\sigma_4}} V_{\sigma_1\sigma_2\sigma_3\sigma_4} \left[(\delta c^2)_{\mathbf{k}\sigma_2\sigma_1}^\dagger \langle c_{\mathbf{k}'\sigma_3} c_{-\mathbf{k}'\sigma_4} \rangle + (\delta c^2)_{\mathbf{k}'\sigma_3\sigma_4} \langle c_{-\mathbf{k}\sigma_1}^\dagger c_{\mathbf{k}\sigma_2}^\dagger \rangle \right. \\ \left. + \langle c_{-\mathbf{k}\sigma_1}^\dagger c_{\mathbf{k}\sigma_2}^\dagger \rangle \langle c_{\mathbf{k}'\sigma_3} c_{-\mathbf{k}'\sigma_4} \rangle \right] + \mathcal{O}[(\delta c^2)^2] . \quad (2.3)$$

We reinsert the auxiliary operators in the linear terms and use Eq. (2.2) (for full calculation see appendix App. A.1) to express the interaction Hamiltonian as

$$H_{\text{int}} = \frac{1}{2} \sum_{\mathbf{k}} \sum_{\sigma\sigma'} \left[- \sum_{\tilde{\mathbf{k}}} \sum_{\tilde{\sigma}\tilde{\sigma}'} V_{\sigma'\sigma\tilde{\sigma}\tilde{\sigma}'}(\mathbf{k}, \tilde{\mathbf{k}}) \langle c_{\tilde{\mathbf{k}}\tilde{\sigma}} c_{-\tilde{\mathbf{k}}\tilde{\sigma}'} \rangle c_{\mathbf{k}\sigma}^\dagger c_{-\mathbf{k}\sigma'}^\dagger \right. \\ \left. - \sum_{\tilde{\mathbf{k}}} \sum_{\tilde{\sigma}\tilde{\sigma}'} V_{\sigma'\sigma\tilde{\sigma}\tilde{\sigma}'}(\mathbf{k}, \tilde{\mathbf{k}}) \langle c_{-\tilde{\mathbf{k}}\tilde{\sigma}'}^\dagger c_{\tilde{\mathbf{k}}\tilde{\sigma}}^\dagger \rangle c_{-\mathbf{k}\sigma'} c_{\mathbf{k}\sigma} \right] \\ - \text{const.} + \mathcal{O}[(\delta c^2)^2] . \quad (2.4)$$

It will turn out to be helpful to define the pair potential [255, 336]

$$\Delta_{\sigma\sigma'}(\mathbf{k}) = - \sum_{\tilde{\mathbf{k}}} \sum_{\tilde{\sigma}\tilde{\sigma}'} V_{\sigma'\sigma\tilde{\sigma}\tilde{\sigma}'}(\mathbf{k}, \tilde{\mathbf{k}}) \langle c_{\tilde{\mathbf{k}}\tilde{\sigma}} c_{-\tilde{\mathbf{k}}\tilde{\sigma}'} \rangle , \quad (2.5)$$

with $\Delta_{\sigma\sigma'}(\mathbf{k}) = -\Delta_{\sigma'\sigma}(-\mathbf{k})$ [255] such that the Hamiltonian in mean-field approximation is given by

$$H_{\text{MF}} = \sum_{\mathbf{k}\sigma\sigma'} h_{\sigma\sigma'}(\mathbf{k}) c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma'} + \frac{1}{2} \sum_{\mathbf{k}\sigma\sigma'} \left[\Delta_{\sigma\sigma'}(\mathbf{k}) c_{\mathbf{k}\sigma}^\dagger c_{-\mathbf{k}\sigma'}^\dagger + \text{h.c.} \right] . \quad (2.6)$$

Here we neglected the quadratic terms $\mathcal{O}[(\delta c^2)^2]$ and the constant term. It is advantageous to express this Hamiltonian in matrix form, for which we define the Nambu spinor $\Psi_{\mathbf{k}}^\dagger = (c_{\mathbf{k}\uparrow}^\dagger, c_{\mathbf{k}\downarrow}^\dagger, c_{-\mathbf{k}\uparrow}, c_{-\mathbf{k}\downarrow})$ such that [255]

$$H_{\text{MF}} = \frac{1}{2} \sum_{\mathbf{k}} \Psi_{\mathbf{k}}^\dagger \mathcal{H}_{\text{BdG}}(\mathbf{k}) \Psi_{\mathbf{k}} + \frac{1}{2} \sum_{\mathbf{k}\sigma} h_{\sigma\sigma}(\mathbf{k}) \quad (2.7)$$

$$\mathcal{H}_{\text{BdG}}(\mathbf{k}) = \begin{pmatrix} h(\mathbf{k}) & \Delta(\mathbf{k}) \\ \Delta^\dagger(\mathbf{k}) & -h^T(-\mathbf{k}) \end{pmatrix} . \quad (2.8)$$

Here every entry in $\mathcal{H}_{\text{BdG}}(\mathbf{k})$ is itself a 2×2 matrix in the spins, i.e.

$$h(\mathbf{k}) = \begin{pmatrix} h_{\uparrow\uparrow}(\mathbf{k}) & h_{\uparrow\downarrow}(\mathbf{k}) \\ h_{\downarrow\uparrow}(\mathbf{k}) & h_{\downarrow\downarrow}(\mathbf{k}) \end{pmatrix} \text{ and } \Delta(\mathbf{k}) = \begin{pmatrix} \Delta_{\uparrow\uparrow}(\mathbf{k}) & \Delta_{\uparrow\downarrow}(\mathbf{k}) \\ \Delta_{\downarrow\uparrow}(\mathbf{k}) & \Delta_{\downarrow\downarrow}(\mathbf{k}) \end{pmatrix} .$$

In this context, \mathcal{H}_{BdG} is referred to as the Bogoliubov-de Gennes Hamiltonian. Here, energies are measured relative to the Fermi level. Considering the form of the Nambu-Spinor, one recognizes that $\mathcal{H}_{\text{BdG}}(\mathbf{k})$ contains a particle sector with momentum \mathbf{k} and a hole sector with momentum $-\mathbf{k}$. Furthermore, the off-diagonal blocks contain the mean-field term $\Delta(\mathbf{k})$ ($\Delta^\dagger(\mathbf{k})$), which describes creation (annihilation) of a Cooper pair. In addition, the form of the Hamiltonian

shows that electrons and holes are not independent: When an electron is excited from the Fermi sea, a hole is left behind. This fact is formally described by the particle-hole symmetry, with the particle-hole operator [255]

$$\mathcal{P} = \tau_x \otimes \sigma_0 \mathcal{K} = \begin{pmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{pmatrix} \mathcal{K} \text{ with } \mathcal{P}^2 = \mathbb{1}, \quad (2.9)$$

where \mathcal{K} is the operator for complex conjugation and τ_i (σ_i) are Pauli matrices in particle-hole (spin) space. The operator for particle-hole conjugation transforms the Hamiltonian as

$$\mathcal{P} \mathcal{H}_{\text{BdG}}(\mathbf{k}) \mathcal{P}^{-1} = \mathcal{K} \tau_x \mathcal{H}_{\text{BdG}}(\mathbf{k}) \tau_x \mathcal{K}^{-1} = \mathcal{K} (-\mathcal{H}_{\text{BdG}}^T(-\mathbf{k})) \mathcal{K}^{-1} = -\mathcal{H}_{\text{BdG}}(-\mathbf{k}). \quad (2.10)$$

It follows that for each eigenstate $\xi_E(\mathbf{k})$ of $\mathcal{H}_{\text{BdG}}(\mathbf{k})$ with energy eigenvalue E , i.e. $\mathcal{H}_{\text{BdG}}(\mathbf{k}) \xi_E(\mathbf{k}) = E \xi_E(\mathbf{k})$, there exists an eigenstate $\mathcal{P} \xi_E(-\mathbf{k}) = \xi_{-E}(-\mathbf{k})$ with eigenvalue $-E$, i.e. $\mathcal{H}_{\text{BdG}}(\mathbf{k}) \mathcal{P} \xi_E(-\mathbf{k}) = -E_{-\mathbf{k}} \mathcal{P} \xi_E(-\mathbf{k})$. The particle-hole operator thus transforms between these eigenstates

$$\mathcal{P} \xi_E(-\mathbf{k}) = \xi_{-E}(-\mathbf{k}). \quad (2.11)$$

The Hermitian 4×4 matrix $\mathcal{H}_{\text{BdG}}(\mathbf{k})$ thus has eigenvalue-eigenvector pairs $\{(E_1(\mathbf{k}), \xi_{E_1}(\mathbf{k})), (E_2(\mathbf{k}), \xi_{E_2}(\mathbf{k})), (-E_1(-\mathbf{k}), \xi_{-E_1}(-\mathbf{k})), (-E_2(-\mathbf{k}), \xi_{-E_2}(-\mathbf{k}))\}$. To calculate the eigenstates and energy eigenvalues of the Hamiltonian Eq. (2.8), we perform a Bogoliubov transformation [337, 338], for which we define new operators $\tilde{\Psi}_{\mathbf{k}}^\dagger = (a_{\mathbf{k}1}^\dagger, a_{\mathbf{k}2}^\dagger, a_{-\mathbf{k},1}, a_{-\mathbf{k},2})$ via the unitary transformation

$$\Psi_{\mathbf{k}} = (\xi_{E_1}(\mathbf{k}), \xi_{E_2}(\mathbf{k}), \xi_{-E_1}(-\mathbf{k}), \xi_{-E_2}(-\mathbf{k})) \tilde{\Psi}_{\mathbf{k}}^\dagger \equiv U_{\mathbf{k}} \tilde{\Psi}_{\mathbf{k}}^\dagger, \quad (2.12)$$

where $\xi_{\pm E_i}(\mathbf{k}) = (u_{\pm i, \uparrow}(\mathbf{k}), u_{\pm i, \downarrow}(\mathbf{k}), v_{\pm i, \uparrow}(-\mathbf{k}), v_{\pm i, \downarrow}(-\mathbf{k}))^T$ is a vector with four components, and U is a unitary 4×4 matrix. This definition already takes into account the restrictions of the eigenvectors and eigenvalues, which result from the particle-hole symmetry. If we further use the relation

$$\xi_{-E_i}(-\mathbf{k}) = \mathcal{P} \xi_{E_i}(-\mathbf{k}) = \left(v_{i\uparrow}^*(\mathbf{k}), v_{i\downarrow}^*(\mathbf{k}), u_{i\uparrow}^*(-\mathbf{k}), u_{i\downarrow}^*(-\mathbf{k}) \right)^T, \quad (2.13)$$

we obtain the transformation matrix

$$U_{\mathbf{k}} = \begin{pmatrix} u_{1\uparrow}(\mathbf{k}) & u_{2\uparrow}(\mathbf{k}) & v_{1\uparrow}^*(\mathbf{k}) & v_{2\uparrow}^*(\mathbf{k}) \\ u_{1\downarrow}(\mathbf{k}) & u_{2\downarrow}(\mathbf{k}) & v_{1\downarrow}^*(\mathbf{k}) & v_{2\downarrow}^*(\mathbf{k}) \\ v_{1\uparrow}(-\mathbf{k}) & v_{2\uparrow}(-\mathbf{k}) & u_{1\uparrow}^*(-\mathbf{k}) & u_{2\uparrow}^*(-\mathbf{k}) \\ v_{1\downarrow}(-\mathbf{k}) & v_{2\downarrow}(-\mathbf{k}) & u_{1\downarrow}^*(-\mathbf{k}) & u_{2\downarrow}^*(-\mathbf{k}) \end{pmatrix}. \quad (2.14)$$

For unitarity, i.e. $U_{\mathbf{k}}^\dagger U_{\mathbf{k}} = \mathbb{1} = U_{\mathbf{k}} U_{\mathbf{k}}^\dagger$, we need

$$\begin{aligned} \sum_{\sigma} [u_{i\sigma}^*(\mathbf{k}) u_{j\sigma}(\mathbf{k}) + v_{i\sigma}^*(\pm\mathbf{k}) v_{j\sigma}(\pm\mathbf{k})] &= \delta_{ij}, \\ \sum_{\sigma} [u_{i\sigma}(\mathbf{k}) v_{j\sigma}(\mathbf{k}) + v_{i\sigma}(-\mathbf{k}) u_{j\sigma}(-\mathbf{k})] &= 0, \\ \sum_{\sigma} [u_{i\sigma}^*(\mathbf{k}) v_{j\sigma}(-\mathbf{k}) + v_{i\sigma}^*(\mathbf{k}) u_{j\sigma}(-\mathbf{k})] &= 0. \end{aligned} \quad (2.15)$$

We assume now that the matrix $U_{\mathbf{k}}$ is chosen such that it diagonalizes the Hamiltonian which yields

$$H = \frac{1}{2} \sum_{\mathbf{k}} \tilde{\Psi}_{\mathbf{k}}^{\dagger} U_{\mathbf{k}}^{\dagger} \mathcal{H}_{\text{BdG}}(\mathbf{k}) U_{\mathbf{k}} \tilde{\Psi}_{\mathbf{k}} + \frac{1}{2} \sum_{\mathbf{k}\sigma} h_{\sigma\sigma}(\mathbf{k}) = \frac{1}{2} \sum_{\mathbf{k}} \tilde{\Psi}_{\mathbf{k}}^{\dagger} D_{\mathbf{k}} \tilde{\Psi}_{\mathbf{k}} + \frac{1}{2} \sum_{\mathbf{k}\sigma} h_{\sigma\sigma}(\mathbf{k}) . \quad (2.16)$$

By the above choice of the order of the eigenvectors in the transformation matrix, the diagonal matrix has the form

$$D_{\mathbf{k}} = \begin{pmatrix} E_1(\mathbf{k}) & 0 & 0 & 0 \\ 0 & E_2(\mathbf{k}) & 0 & 0 \\ 0 & 0 & -E_1(-\mathbf{k}) & 0 \\ 0 & 0 & 0 & -E_2(-\mathbf{k}) \end{pmatrix} . \quad (2.17)$$

We assume without loss of generality that all $E_i(\mathbf{k}) \geq 0$ [255]. Using U , Eq. (2.14), we obtain the new operators in which H_{MF} is diagonal as

$$a_{\mathbf{k}i} = \sum_{\sigma} \left[u_{i\sigma}^*(\mathbf{k}) c_{\mathbf{k}\sigma} + v_{i\sigma}^*(-\mathbf{k}) c_{-\mathbf{k}\sigma}^{\dagger} \right] . \quad (2.18)$$

Because of the unitarity requirement, Eq. (2.15), these operators also satisfy fermionic commutator relations

$$\{a_{\mathbf{k}i}, a_{\mathbf{k}'j}\} = 0 = \{a_{\mathbf{k}i}^{\dagger}, a_{\mathbf{k}'j}^{\dagger}\} \text{ and } \{a_{\mathbf{k}i}^{\dagger}, a_{\mathbf{k}'j}\} = \delta_{\mathbf{k}\mathbf{k}'} \delta_{ij} . \quad (2.19)$$

The quasi-particle excitations with energy $E_i(\mathbf{k})$, which are generated by $a_{\mathbf{k},i}^{\dagger}$, are called Bogoliubons. As one can see from the definition of $a_{\mathbf{k},i}^{\dagger}$, these are superpositions of electrons and holes. In these new operators the Hamiltonian has the diagonal form

$$\begin{aligned} H &= \frac{1}{2} \sum_{\mathbf{k}i} E_i(\mathbf{k}) a_{\mathbf{k}i}^{\dagger} a_{\mathbf{k}i} - \frac{1}{2} \sum_{\mathbf{k}i} E_i(-\mathbf{k}) a_{-\mathbf{k}i} a_{-\mathbf{k}i}^{\dagger} + \frac{1}{2} \sum_{\mathbf{k}\sigma} h_{\sigma\sigma}(\mathbf{k}) \\ &= \sum_{\mathbf{k}i} E_i(\mathbf{k}) a_{\mathbf{k}i}^{\dagger} a_{\mathbf{k}i} - \frac{1}{2} \sum_{\mathbf{k}i} E_i(\mathbf{k}) + \frac{1}{2} \sum_{\mathbf{k}\sigma} h_{\sigma\sigma}(\mathbf{k}) , \end{aligned} \quad (2.20)$$

where we again assumed in the last step that for every momentum \mathbf{k} in the sum the momentum $-\mathbf{k}$ is also present.

We define the ground state $|\text{GS}\rangle$ as the state which is annihilated by all operators $a_{\mathbf{k}i}$, i.e. which contains no quasi-particles

$$a_{\mathbf{k}i} |\text{GS}\rangle = 0 . \quad (2.21)$$

Since the pairing term in the mean-field Hamiltonian Eq. (2.6) does not commute with the particle number operator $\hat{N} = \sum_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma}^{\dagger} c_{\mathbf{k}\sigma}$, states can have an indefinite particle number. For the ground state the mean particle number can be calculated using the operators $a_{\mathbf{k}i}$, Eq. (2.18), and the relation Eq. (2.21) as

$$\begin{aligned} N &= \langle \text{GS} | \hat{N} | \text{GS} \rangle = \sum_{\mathbf{k}\sigma} \langle \text{GS} | c_{\mathbf{k}\sigma}^{\dagger} c_{\mathbf{k}\sigma} | \text{GS} \rangle \\ &= \sum_{\mathbf{k}\sigma ij} \langle \text{GS} | \left(u_{i\sigma}^*(\mathbf{k}) a_{\mathbf{k}i}^{\dagger} + v_{i\sigma}(\mathbf{k}) a_{-\mathbf{k}i} \right) \left(u_{j\sigma}(\mathbf{k}) a_{\mathbf{k}j} + v_{j\sigma}^*(\mathbf{k}) a_{-\mathbf{k}j}^{\dagger} \right) | \text{GS} \rangle \\ &= \sum_{\mathbf{k}\sigma i} |v_{i\sigma}(\mathbf{k})|^2 . \end{aligned} \quad (2.22)$$

Here, u are the particle-like components of the wave function, and the hole-like components are given by v . On the other hand, since the Hamiltonian contains only pairs of fermion operators, the number parity is uniquely defined. Here, the corresponding number parity operator is defined as $(-1)^{\hat{N}}$ [86, 265], and we denote the parity of the ground state of the Hamiltonian H as $P[H] = \langle \text{GS} | (-1)^{\hat{N}} | \text{GS} \rangle$ for even and odd ground state number parity, respectively. Before we discuss the importance of the ground state number parity further, we consider the energy of the ground state of the mean field Hamiltonian, which is given by

$$E = \langle \text{GS} | H | \text{GS} \rangle = -\frac{1}{2} \sum_{\mathbf{k}i} E_i(\mathbf{k}) + \frac{1}{2} \sum_{\mathbf{k}\sigma} h_{\sigma\sigma}(\mathbf{k}) . \quad (2.23)$$

From this we see that all states with negative energies are occupied, as one would expect from a zero temperature ground state. Since the ground state is the state with the lowest energy, it is uniquely determined as long as no zero energy states exist, i.e. $|E_i(\mathbf{k})| > 0$. However, if there is a zero-energy level, $E_j(\mathbf{q}) = 0$, another ground state with the same energy is given by $a_{\mathbf{q}j}^\dagger |\text{GS}\rangle$, which has the opposite number parity than $|\text{GS}\rangle$. In the case of n zero energy states there are 2^n such ground states, which differ in the occupation numbers of the zero energy states. In this case, one speaks of a 2^n -fold degenerate ground state, which can for instance be exploited for the construction of a qubit in the case $n = 1$. We will show in the following that such degenerate ground states can occur in the non-trivial phase of topological superconductors. However, the presence of the topological phase alone is not sufficient for the presence of such states. In addition, one needs defects, such as a transition from positive to negative chemical potential, or a system with a boundary, so that two topologically different areas come into contact. In this case, zero modes located at these boundary regions appear. In the following, we will therefore first consider how phases with different topology can be distinguished in the non-degenerate case, i.e. for $|E_i(\mathbf{k})| > 0$, and in later sections we then add defects to create zero modes.

It turns out that quasi-non-interacting Hamiltonians with an excitation gap in the spectrum – for example Eq. (2.8) – can be divided according to their symmetry properties into 10 different Altland-Zirnbauer symmetry classes, for which systems have different types of topological invariants, depending on the number of spatial dimension [339–342]. Topological invariants are specific numbers defined for systems with periodic boundary conditions, which distinguish topological states and cannot be influenced by external perturbations without breaking the underlying symmetries or without closing the excitation gap. They therefore classify different topological phases, which may also be related to physical properties such as the occurrence of zero modes on defects. Such properties are then called topologically protected, and a phase which has such properties is called topologically non-trivial. Since the topological invariant cannot be changed without lifting the above assumptions, a phase transition between two topological phases is associated with closing the excitation gap (or breaking a symmetry).

Three basic symmetries are crucial for the classification of Hamiltonians into the Altland-Zirnbauer classes:

- (i) Time reversal $\mathcal{T} = U_{\mathcal{T}}\mathcal{K}$ with $\mathcal{T} = \pm 1$ is an anti-unitary operator, which can be constructed from a unitary transformation $U_{\mathcal{T}}$ and complex conjugation \mathcal{K} . One calls a Hamiltonian H time reversal symmetric, if it commutes with \mathcal{T} , i.e. $[H, \mathcal{T}] = 0$. For the case of a \mathcal{T} -symmetric BdG Hamiltonian Eq. (2.8), this is equivalent to $\mathcal{T}\mathcal{H}_{\text{BdG}}(\mathbf{k})\mathcal{T}^{-1} = \mathcal{H}_{\text{BdG}}(-\mathbf{k})$.
- (ii) The particle-hole operator $\mathcal{P} = U_{\mathcal{P}}\mathcal{K}$ with $\mathcal{P}^2 = \pm 1$ is another anti-unitary operator, and one says that H has \mathcal{P} -symmetry if it anti-commutes with \mathcal{P} , i.e. $\{H, \mathcal{P}\} = 0$. We already

Table 2.1: Altland-Zirnbauer [339–342] symmetry classes and topological invariants.

class	symmetry			dimension d			
	\mathcal{T}	\mathcal{P}	\mathcal{C}	0	1	2	3
A	0	0	0	\mathbb{Z}	0	\mathbb{Z}	0
AIII	0	0	1	0	\mathbb{Z}	0	\mathbb{Z}
AI	+1	0	0	\mathbb{Z}	0	0	0
BDI	+1	+1	1	\mathbb{Z}_2	\mathbb{Z}	0	0
D	0	+1	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}	0
DIII	-1	+1	1	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}
AII	-1	0	0	\mathbb{Z}	0	\mathbb{Z}_2	\mathbb{Z}_2
CII	-1	-1	1	0	\mathbb{Z}	0	\mathbb{Z}_2
C	0	-1	0	0	0	\mathbb{Z}	0
CI	+1	-1	0	0	0	0	\mathbb{Z}

considered \mathcal{P} with $\mathcal{P}^2 = 1$ in the case of the BdG Hamiltonian in Eqs. (2.8) and (2.9), where we found that particle-hole symmetry implies that $\mathcal{P}\mathcal{H}_{\text{BdG}}(\mathbf{k})\mathcal{P}^{-1} = -\mathcal{H}_{\text{BdG}}(-\mathbf{k})$.

- (iii) The chiral operator \mathcal{C} is a unitary operator that anti-commutes with H in case of chiral symmetry. If the Hamiltonian has \mathcal{P} and \mathcal{T} symmetry, it automatically has a \mathcal{C} symmetry with $\mathcal{C} = \mathcal{T}\mathcal{P} = U_{\mathcal{T}}U_{\mathcal{P}}$. In the case of the BdG Hamiltonian, chiral symmetry is equivalent to $\mathcal{C}\mathcal{H}_{\text{BdG}}(\mathbf{k})\mathcal{C}^{-1} = -\mathcal{H}_{\text{BdG}}(\mathbf{k})$.

Table 2.1 depicts the classification into the Altland-Zirnbauer symmetry classes according to symmetries and the value of the square of the symmetry operator. Here, an entry 0 means that the symmetry is not present, and ± 1 means that the symmetry is there, and the symmetry operator squares to ± 1 . Of particular interest for the superconductors in the BdG formalism considered so far are the classes that exhibit particle-hole symmetry with $\mathcal{P}^2 = 1$, i.e. BDI, D, and DIII. The different classes can be characterized by various topological invariants, depending on the spatial dimension d , where a zero indicates that only the trivial phase exists. For \mathbb{Z}_2 , there is a topological invariant, which can take two different values, e.g. ± 1 , such that two topologically different phases exist. If one computes the topological \mathbb{Z}_2 invariant for a given Hamiltonian, one can determine in which topological phase the system is in. However, a system can also be in the same topological phase for all parameters (e.g. BCS s-wave superconductors) or, depending on the parameters, a topological phase transition can occur at which the excitation gap closes (e.g. spin-polarized p-wave superconductors [86]). In case of a \mathbb{Z} topological invariant, there are infinitely many topologically different phases. The invariant can be computed as a winding number in one spatial dimension (e.g. SSH model [343]).

For the moment, we restrict the discussion to the BdG Hamiltonian Eq. (2.8) for a system in symmetry class D. As described above, the ground state (we assume for now that it is not degenerate) has a defined number parity $\mathcal{Q} = P[\mathcal{H}] = \langle \text{GS} | (-1)^{\hat{N}} | \text{GS} \rangle = \pm 1$. For this case Kitaev [86, 341, 344] showed that \mathcal{Q} is also the topological invariant. Using the arguments of Kitaev, we consider here how \mathcal{Q} can be calculated for the BdG Hamiltonian Eq. (2.8) in class D.

After the Hamiltonian $\mathcal{H} = \text{diag}_{\mathbf{k}}(\mathcal{H}_{\text{BdG}}(\mathbf{k}))$ is diagonalized by the unitary transformation

$U = \text{diag}_{\mathbf{k}}(U_{\mathbf{k}})$ which transforms the fermion operators $c_{\mathbf{k}} \rightarrow a_{\mathbf{k}}$, the ground state $|\text{GS}\rangle$ contains zero quasi-particles. Therefore, the ground state is the vacuum for the $a_{\mathbf{k}}$ operators, so that the transformed Hamiltonian has even particle number parity $P[U^\dagger H U] = (-1)^0 = +1$ with respect to the operators $(a_{\mathbf{k}}, a_{\mathbf{k}}^\dagger) = U_{\mathbf{k}}(c_{\mathbf{k}}, c_{\mathbf{k}}^\dagger)$ [86]. But if $\det(U) = +1$, the transformation preserves the number parity such that also $P[H] = +1$. On the other hand if $\det(U) = -1$ then U reverses the parity and $P[H] = -1$ [86]. Here, the notation $P[H]$ describes the parity related to $\hat{N} = \sum_{\mathbf{k}} c_{\mathbf{k}}^\dagger c_{\mathbf{k}}$, while $P[U^\dagger H U]$ is the parity with respect to the quasi-particle number operator $\sum_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}}$. Thus, the particle number parity is given by²

$$\mathcal{Q} = \langle \text{GS} | (-1)^{\hat{N}} | \text{GS} \rangle = \text{sgn}(\det(U)) . \quad (2.24)$$

Here one could omit the sgn , since we find below that $\det(U) = \pm 1$ holds, but it will turn out to be helpful in the calculation below. Since the determinant of a block diagonal matrix is given by the product of the determinants of the blocks, we first obtain [86]

$$\mathcal{Q} = \prod_{\mathbf{k}} \text{sgn}(\det(U_{\mathbf{k}})) . \quad (2.25)$$

From the discussion about particle-hole symmetry \mathcal{P} we already know that for all momenta $\mathbf{k} \neq -\mathbf{k}$ one finds $\mathcal{P}^{-1} U_{\mathbf{k}} \mathcal{P} = U_{-\mathbf{k}}$, where $\mathcal{P} = U_{\mathcal{P}} \mathcal{K}$ with $\mathcal{P}^2 = 1$ and $U_{\mathcal{P}} = \tau_x \otimes \sigma_0$. Because of $\det(U_{\mathcal{P}}) = 1$ and properties of the determinant

$$\text{sgn}[\det(U_{-\mathbf{k}})] = \text{sgn}[\det(U_{\mathcal{P}} U_{\mathbf{k}}^* U_{\mathcal{P}})] = \text{sgn}[\det(U_{\mathcal{P}}) \det(U_{\mathbf{k}}^*) \det(U_{\mathcal{P}})] = \text{sgn}[\det(U_{\mathbf{k}})]^{-1} , \quad (2.26)$$

which is why the factors for \mathbf{k} and $-\mathbf{k}$ pairwise multiply together to $+1$. This reduces the expression of the number parity to [86]

$$\mathcal{Q} = \prod_{\mathbf{k}=-\mathbf{k}} \text{sgn}(\det(U_{\mathbf{k}})) , \quad (2.27)$$

where $\mathbf{k} = -\mathbf{k}$ has to be understood as an equivalence under the addition of reciprocal lattice vectors. In the following we denote momenta satisfying $\mathbf{k} = -\mathbf{k}$ by \mathbf{K} . For the example of a one-dimensional system with Brillouin zone $(-\pi, \pi]$, these momenta are $K = 0$ and $K = \pi$. In addition, Eq. (2.26) ensures that $\det U_{\mathbf{K}} = \pm 1$. To find an expression for \mathcal{Q} in terms of the Hamiltonian, we consider the diagonalized blocks corresponding to these momenta

$$U_{\mathbf{K}}^\dagger \mathcal{H}_{\text{BdG}}(\mathbf{K}) U_{\mathbf{K}} = D_{\mathbf{K}} = \begin{pmatrix} E_1(\mathbf{K}) & 0 & 0 & 0 \\ 0 & E_2(\mathbf{K}) & 0 & 0 \\ 0 & 0 & -E_1(\mathbf{K}) & 0 \\ 0 & 0 & 0 & -E_2(\mathbf{K}) \end{pmatrix} . \quad (2.28)$$

Application of particle-hole symmetry using $\mathcal{P} U_{\mathbf{K}} = U_{\mathbf{K}} \mathcal{P}$ and $\mathcal{P} \mathcal{H}_{\text{BdG}}(\mathbf{K}) \mathcal{P} = -\mathcal{H}_{\text{BdG}}(\mathbf{K})$ yields

$$U_{\mathbf{K}}^\dagger \mathcal{P} \mathcal{H}_{\text{BdG}}(\mathbf{K}) \mathcal{P} U_{\mathbf{K}} = -U_{\mathbf{K}}^\dagger \mathcal{H}_{\text{BdG}}(\mathbf{K}) U_{\mathbf{K}} = \mathcal{P} D_{\mathbf{K}} \mathcal{P} \quad (2.29)$$

$$\Rightarrow -U_{\mathbf{K}}^\dagger \mathcal{H}_{\text{BdG}}(\mathbf{K}) \mathcal{P} U_{\mathbf{K}} = \mathcal{P} D_{\mathbf{K}} , \quad (2.30)$$

$$-U_{\mathbf{K}}^\dagger \mathcal{H}_{\text{BdG}}(\mathbf{K}) U_{\mathcal{P}} U_{\mathbf{K}}^* = U_{\mathcal{P}} D_{\mathbf{K}} . \quad (2.31)$$

²In fact, the determinant of a unitary matrix is a complex number with modulus 1 in general. However, we argue in the following that the particle-hole symmetry ensures that $\det(U) = \pm 1$ for a gapped Hamiltonian.

In the last step we used that $D_{\mathbf{K}}$ is real as $\mathcal{H}_{\text{BdG}}(\mathbf{K})$ is Hermitian. We performed these transformations in order to express \mathcal{Q} as a Pfaffian [265]

$$\text{Pf}[A] = \frac{1}{2^n n!} \sum_{\sigma \in S_{2n}} \text{sgn}(\sigma) \prod_{i=1}^n a_{\sigma(2i-1), \sigma(2i)} \quad (2.32)$$

of a skew-symmetric $2n \times 2n$ matrix $A^T = -A$, where S_{2n} are permutations and $\text{sgn}(\sigma)$ is the sign of the permutation σ . The following properties [345] hold for skew-symmetric $2n \times 2n$ matrices A , B and $\lambda \in \mathbb{C}$:

$$\text{Pf}[\lambda A] = \lambda^n \text{Pf}[A] , \quad (2.33)$$

$$\text{Pf}[BAB^T] = \det(B) \text{Pf}[A] , \quad (2.34)$$

$$(\text{Pf}[A])^2 = \det(A) . \quad (2.35)$$

In our case the particle-hole symmetry guarantees that $\mathcal{P}\mathcal{H}_{\text{BdG}}(\mathbf{K})\mathcal{P} = -\mathcal{H}_{\text{BdG}}(\mathbf{K})$ and $[\mathcal{H}_{\text{BdG}}(\mathbf{K})]^\dagger = \mathcal{H}_{\text{BdG}}(\mathbf{K}) \Rightarrow U_{\mathcal{P}}[\mathcal{H}_{\text{BdG}}(\mathbf{K})]^T = -\mathcal{H}_{\text{BdG}}(\mathbf{K})U_{\mathcal{P}}$, so that the 4×4 matrix $\mathcal{H}_{\text{BdG}}(\mathbf{K})U_{\mathcal{P}}$ is skew-symmetric, i.e.

$$(\mathcal{H}_{\text{BdG}}(\mathbf{K})U_{\mathcal{P}})^T = U_{\mathcal{P}}[\mathcal{H}_{\text{BdG}}(\mathbf{K})]^T = -\mathcal{H}_{\text{BdG}}(\mathbf{K})U_{\mathcal{P}} . \quad (2.36)$$

If we now evaluate the Pfaffian of Eq. (2.31), using Eq. (2.34) and Eq. (2.33) with $(-1)^2 = 1$, we obtain

$$\det(U_{\mathbf{K}}^\dagger) \text{Pf}[\mathcal{H}_{\text{BdG}}(\mathbf{K})U_{\mathcal{P}}] = \text{Pf}[U_{\mathbf{K}}^\dagger \mathcal{H}_{\text{BdG}}(\mathbf{K})U_{\mathcal{P}}U_{\mathbf{K}}^*] \quad (2.37)$$

$$= \text{Pf}[U_{\mathcal{P}}D_{\mathbf{K}}] \quad (2.38)$$

$$= \text{Pf} \begin{pmatrix} 0 & 0 & -E_1(\mathbf{K}) & 0 \\ 0 & 0 & 0 & -E_2(\mathbf{K}) \\ E_1(\mathbf{K}) & 0 & 0 & 0 \\ 0 & E_2(\mathbf{K}) & 0 & 0 \end{pmatrix} \\ = -(-E_1(\mathbf{K}))(-E_2(\mathbf{K})) . \quad (2.39)$$

Since $E_i(\mathbf{K}) > 0$ and $\det U_{\mathbf{K}}^\dagger = 1/\det U_{\mathbf{K}}$ as U is unitary, we obtain the topological invariant

$$\mathcal{Q} = \prod_{\mathbf{K}} \text{sgn}(\det[U_{\mathbf{K}}]) = - \prod_{\mathbf{K}} \text{sgn}(\text{Pf}[\mathcal{H}_{\text{BdG}}(\mathbf{K})U_{\mathcal{P}}]) . \quad (2.40)$$

Using Eq. (2.33) and $-1 = i^2$, we obtain the final expression [255, 265, 283]

$$\mathcal{Q} = \prod_{\mathbf{k}=-\mathbf{k}} \text{sgn}(\text{Pf}[i\mathcal{H}_{\text{BdG}}(\mathbf{k})U_{\mathcal{P}}]) . \quad (2.41)$$

We also see that a continuous change of the sign of $\text{Pf}[i\mathcal{H}_{\text{BdG}}(\mathbf{K})U_{\mathcal{P}}]$ can only occur if $0 = (\text{Pf}[i\mathcal{H}_{\text{BdG}}(\mathbf{K})U_{\mathcal{P}}])^2 = \det(\mathcal{H}_{\text{BdG}}(\mathbf{k})) = E_1^2(\mathbf{K})E_2^2(\mathbf{K})$ for any \mathbf{K} , i.e. if the gap closes. For $\mathcal{Q} = +1$ the system is in the topologically trivial phase³ and for $\mathcal{Q} = -1$ in the topological phase, where the ground state has odd number parity.

So far we did not consider any other symmetry besides the particle-hole symmetry. As shown in Tab. 2.1, an additional pseudo-time reversal symmetry \mathcal{T} with $\mathcal{T}^2 = +1$ brings the system

³In the trivial phase the ground state number parity is even as it is also the case for the vacuum.

into the BDI symmetry class. In one dimension, a Hamiltonian in this case is associated with a \mathbb{Z} topological invariant, which can be calculated as a winding number. However, the Pfaffian invariant $P[H]$, Eq. (2.41), still describes the ground state number parity, and it can be shown that this corresponds to the parity of the winding number [346, 347]. Since it will later become apparent that only an odd winding number, i.e. $P[H] = -1$, is relevant for the occurrence of Majorana zero modes [347, 348], we restrict ourselves to the calculation of \mathcal{Q} using Eq. (2.41) and do not elaborate further on the \mathbb{Z} topological invariant.

As we saw, the topological invariant cannot change as long as the excitation gap is not closed. If two systems with different topological invariants are brought into contact, the topological invariant at the boundary changes. If both systems have an excitation gap, this apparent problem is solved by the occurrence of a zero energy mode localized at the boundary [255, 349]. This connection between the bulk property \mathcal{Q} and the edge mode is called bulk-edge correspondence. Furthermore, in a topological superconductor, particle-hole symmetry guarantees the relation $\gamma_E^\dagger = \gamma_{-E}$ for the creators and annihilators of quasi-particles. At energy $E = 0$, this localized zero mode thus satisfies the Majorana condition $\gamma^\dagger = \gamma$. The aim of this chapter is to describe an experimentally realizable topological superconductor and to investigate the occurrence of Majorana zero modes and their properties.

As shown above, the points $\mathbf{k} = -\mathbf{k}$ in the Brillouin zone determine the topological invariant. Because of the property $\Delta(\mathbf{k}) = -\Delta^T(-\mathbf{k})$, we find in particular

$$\Delta_{\sigma\sigma}(\mathbf{K}) = -\Delta_{\sigma\sigma}(\mathbf{K}) = 0, \quad (2.42)$$

$$\Delta_{\uparrow\downarrow}(\mathbf{K}) = -\Delta_{\downarrow\uparrow}(\mathbf{K}). \quad (2.43)$$

If we consider a free electron system described by $h(\mathbf{k}) = \text{diag}[\hbar^2 k^2/(2m) - \mu, \hbar^2 k^2/(2m) - \mu]$ together with the pairing $\Delta(\mathbf{k})$, the possibility of entering the topological phase, i.e. closing the gap, depends on the pair potential $\Delta(\mathbf{k})$. In general, the pair potential can be written as [255, 284, 336]

$$\Delta(\mathbf{k}) = i\psi(\mathbf{k})\sigma_y + i(\mathbf{d}(\mathbf{k}) \cdot \boldsymbol{\sigma})\sigma_y = \begin{pmatrix} -d_x(\mathbf{k}) + id_y(\mathbf{k}) & \psi(\mathbf{k}) + d_z(\mathbf{k}) \\ -\psi(\mathbf{k}) + d_z(\mathbf{k}) & d_x(\mathbf{k}) + id_y(\mathbf{k}) \end{pmatrix}, \quad (2.44)$$

where $\boldsymbol{\sigma}$ is the vector of the Pauli matrices. Here, the condition $\Delta(\mathbf{k}) = -\Delta^T(-\mathbf{k})$ yields

$$\psi(\mathbf{k}) = \psi(-\mathbf{k}) \text{ and } \mathbf{d}(\mathbf{k}) = -\mathbf{d}(-\mathbf{k}). \quad (2.45)$$

Since the first term $\Delta_{\sigma\sigma'}^{\text{singlet}} = i\psi(\mathbf{k})[\sigma_y]_{\sigma\sigma'}$ is antisymmetric under spin reversal, i.e. $\Delta_{\sigma\sigma'}^{\text{singlet}} = -\Delta_{\sigma'\sigma}^{\text{singlet}}$, it is called spin-singlet pairing term. In contrast, the second term with $\Delta_{\sigma\sigma'}^{\text{triplet}}(\mathbf{k}) = i\mathbf{d}(\mathbf{k})[\boldsymbol{\sigma}\sigma_y]_{\sigma\sigma'} = \Delta_{\sigma'\sigma}^{\text{triplet}}(\mathbf{k})$ is denoted as the triplet pairing term. In analogy to atomic orbitals, singlet pairings are called s,d,...-wave depending on angular momentum and triplet pairings are called p,f,...-wave.

Many superconductors have s-wave pairing symmetry, which corresponds to the case $\Delta(\mathbf{k}) = i\Delta\sigma_y$ with isotropic $\psi(\mathbf{k}) = \Delta \in \mathbb{R}$ and $\mathbf{d}(\mathbf{k}) = 0$. For free electrons subject to such pairing, we obtain the conventional BCS superconductor [350] with Hamiltonian

$$\mathcal{H}_{\text{BdG}}(\mathbf{k}) = \begin{pmatrix} h(\mathbf{k}) & \Delta(\mathbf{k}) \\ \Delta^\dagger(\mathbf{k}) & -h^T(-\mathbf{k}) \end{pmatrix}, \quad h(\mathbf{k}) = \begin{pmatrix} \frac{\hbar^2 k^2}{2m} - \mu & 0 \\ 0 & \frac{\hbar^2 k^2}{2m} - \mu \end{pmatrix}, \text{ and } \Delta(\mathbf{k}) = \begin{pmatrix} 0 & \Delta \\ -\Delta & 0 \end{pmatrix}. \quad (2.46)$$

Diagonalization yields the spectrum $E(\mathbf{k}) = \sqrt{\xi_{\mathbf{k}}^2 + \Delta^2}$ with $\xi_{\mathbf{k}} = \hbar^2 k^2 / (2m)$. The excitation gap is minimal at the Fermi level and given by 2Δ . Since $[\hbar^2 k^2 / (2m)]^2 \geq 0$, the gap can never close for the s-wave case and one remains in the same topological phase independent of the parameters. The topological invariant, Eq. (2.41), is given by

$$\mathcal{Q} = \prod_{\mathbf{k}=-\mathbf{k}} \text{sgn}(\text{Pf}[\mathbf{i}\mathcal{H}_{\text{BdG}}(\mathbf{k})U_{\mathcal{P}}]) = \prod_{\mathbf{k}=-\mathbf{k}} \text{sgn}(\xi_{\mathbf{k}}^2 + \Delta^2) = +1. \quad (2.47)$$

This means that a conventional s-wave superconductor is always topologically trivial.

In the following sections we also show examples of topologically non-trivial superconductors with $\mathcal{Q} = -1$. Among other things, we will see that the combination of an s-wave superconductors and spin-orbit coupling can mimic the physics of p-wave pairing and can thus give rise to topological superconductivity.

2.2 Realization of Majorana zero modes

2.2.1 Spin-polarized p-wave superconductor and the Kitaev chain

As we have seen in the previous section, a conventional BCS s-wave superconductor is always topologically trivial. There we have already shown that the spin dependence of the pair potential $\Delta_{\uparrow\downarrow}(\mathbf{k}) = -\Delta_{\downarrow\uparrow}(-\mathbf{k})$ ensures that the excitation gap never closes. On the other hand, for a p-wave superconductor $\Delta_{\sigma\sigma'}(\mathbf{k}) = -\Delta_{\sigma\sigma'}(-\mathbf{k})$ holds, so that the gap can vanish at a point $\mathbf{k} = -\mathbf{k}$. However, if we consider a one-dimensional superconductor with $h_{\sigma\sigma'}(\mathbf{k}) = \xi_{\mathbf{k}}\delta_{\sigma\sigma'}$ and $\Delta_{\sigma\sigma'} = \mathbf{i}k\Delta$, the spin degeneracy is a problem: Even if we were able to achieve an odd ground state number parity for \uparrow -electrons, the overall system including the \downarrow -electrons would have an even ground state number parity. To circumvent this problem, we consider a spin-polarized p-wave superconductor, and first show that it can acquire a topologically non-trivial state. Afterwards we consider a contact between topologically different phases, first in a continuum model with a domain wall and then in a lattice model for a one-dimensional wire with two ends, where we will find localized Majorana modes. The simpler lattice model provides a good insight into the formation of these zero modes and their basic properties. For the moment we ignore the experimental feasibility of spin-polarized p-wave pairing. We will see later that we can make use of all concepts of this section again when we look at realizable but more complicated systems.

To describe the spin-polarized superconductor, we start from Eq. (2.6) with $h_{\sigma\sigma'}(k) = \xi_k\delta_{\sigma\uparrow}\delta_{\sigma'\uparrow}$ and $\Delta_{\sigma\sigma'}(k) = -\mathbf{i}k\Delta\delta_{\sigma\uparrow}\delta_{\sigma'\uparrow}$, thus obtaining the Hamiltonian

$$H = \sum_k \xi_k c_k^\dagger c_k - \frac{1}{2} \sum_k \Delta (\mathbf{i}k c_k^\dagger c_{-k}^\dagger + \mathbf{i}k c_k c_{-k}) , \quad (2.48)$$

with $\xi_k = \hbar^2 k^2 / (2m) - \mu$, where we omitted the spin index. Again we introduce the Nambu spinor $\psi_k^\dagger = (c_k^\dagger, c_{-k})$ to obtain the BdG Hamiltonian [351]

$$H = \frac{1}{2} \sum_k \psi_k^\dagger \mathcal{H}_{\text{BdG}}(k) \psi_k + \frac{1}{2} \sum_k \xi_k \quad (2.49)$$

$$\mathcal{H}_{\text{BdG}}(k) = \begin{pmatrix} \xi_k & -\mathbf{i}k\Delta \\ \mathbf{i}k\Delta & -\xi_{-k} \end{pmatrix} .$$

The eigenvalue spectrum of $\mathcal{H}_{\text{BdG}}(k)$ is given by $E_k = \pm\sqrt{\xi_k^2 + k^2\Delta^2}$, where we used that $\xi_k = \xi_{-k}$. Therefore, the excitation gap at $k = 0$ is given by $2|E_{k=0}| = 2|\mu|$, and it closes for

$\mu = 0$. Due to the absence of a spin degree of freedom, the Hamiltonian is only a 2×2 matrix, so that the Pfaffian in Eq. (2.39) differs by a minus sign and therefore the topological invariant is given by Eq. (2.40) with a positive sign for this case. Hence, for 2×2 systems Eq. (2.41) is modified to

$$\mathcal{Q} = \prod_{k=-k} \text{sgn}(\text{Pf}[\mathcal{H}_{2 \times 2}(k)U_{\mathcal{P}}]) . \quad (2.50)$$

Computing the topological invariant using particle-hole symmetry $\mathcal{P} = U_{\mathcal{P}}\mathcal{K} = \tau_x\mathcal{K}$ yields

$$\mathcal{Q} = \text{sgn} \text{Pf}[\mathcal{H}_{\text{BdG}}(k=0)\tau_x] = \text{sgn} \text{Pf} \left[\begin{pmatrix} 0 & -\mu \\ \mu & 0 \end{pmatrix} \right] = -\text{sgn}(\mu) . \quad (2.51)$$

Thus the system is in the topological phase with $\mathcal{Q} = -1$ for $\mu > 0$ and in the trivial phase for $\mu < 0$. In addition, we see that for $\mu < 0$ the vacuum limit $\mu \rightarrow -\infty$ [283] can be taken without closing the gap, i.e. the trivial phase is topologically equivalent to the vacuum. In fact, the gap is also closing for $\Delta = 0$, so that for $\mu > 0$ two topological phases with winding numbers ± 1 [283] occur. As we discussed in the previous section, the reason for this is that the system belongs to the higher symmetry class BDI. The associated pseudo-time reversal symmetry is given by $\mathcal{T} = \mathcal{K}$ with $\mathcal{T}^2 = +1$. Here, $\mathcal{T}\mathcal{H}_{\text{BdG}}(k)\mathcal{T}^{-1} = \mathcal{H}_{\text{BdG}}(-k)$ becomes apparent when the Hamiltonian $\mathcal{H}_{\text{BdG}}(k) = \xi_k\tau_z + \Delta k\tau_y$ is expressed in terms of Pauli matrices τ_i . The parity of the winding number is then given by the class D Pfaffian invariant [346, 347].

Having shown that the system is topologically non-trivial for $\mu > 0$, we next consider the contact of two topologically distinct regions. In position space the Hamiltonian Eq. (2.48) has the form

$$H = \int_0^L dx \left[\psi^\dagger(x) \left(-\frac{\hbar^2 \partial_x^2}{2m} - \mu \right) \psi(x) - \frac{\Delta}{2} \psi^\dagger(x) \partial_x \psi^\dagger(x) + \frac{\Delta}{2} \psi(x) \partial_x \psi(x) \right] , \quad (2.52)$$

where $\psi^\dagger(x)$ ($\psi(x)$) creates (annihilates) an electron at position x . In order to create a transition between the topological phase for $\mu > 0$ and the trivial phase for $\mu < 0$, we consider a domain wall

$$\mu(x) = \begin{cases} -\mu & \text{for } x < 0 \\ +\mu & \text{for } x > 0 \end{cases} = \mu \text{sgn}(x) , \quad (2.53)$$

with $\mu > 0$. According to our considerations in the last section, we expect a zero mode localized at the interface for this case, which we obtain by solving the BdG equation

$$\begin{pmatrix} -\hbar^2/(2m) \partial_x^2 - \mu(x) & -\Delta \partial_x \\ \Delta \partial_x & \hbar^2/(2m) \partial_x^2 + \mu(x) \end{pmatrix} \begin{pmatrix} u_0(x) \\ v_0(x) \end{pmatrix} = 0 \quad (2.54)$$

for $E = 0$. Here, the particle-hole symmetry $\mathcal{P} = \tau_x\mathcal{K}$ with $\mathcal{P}\mathcal{H}_{\text{BdG}}(x)\mathcal{P}^{-1} = -\mathcal{H}_{\text{BdG}}(x)$ implies that we can choose the solution so that $\mathcal{P}(u_0(x), v_0(x))^T = \pm(u_0(x), v_0(x))^T$, i.e. $u_0(x) = \pm v_0^*(x)$ (see also Appendix A.2). This yields a damped harmonic oscillator in x which is given by

$$-\frac{\hbar^2}{2m} u_0''(x) - \mu(x)u_0(x) \mp \Delta u_0'(x) = 0 , \quad (2.55)$$

with solution (see Appendix A.2 for details)

$$\begin{pmatrix} u_0(x) \\ v_0(x) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tilde{N} \begin{cases} e^{-x/\xi} [\cos(Kx) + \kappa/K \sin(Kx)] & \text{for } x > 0 \\ e^{-x/\xi} e^{\kappa x} & \text{for } x < 0 \end{cases} \quad (2.56)$$

in the underdamped case $\mu > m\Delta^2/(2\hbar^2)$. Here, $K = \sqrt{2m\mu/\hbar^2 - \xi^{-2}}$, $\kappa = \sqrt{2m\mu/\hbar^2 + \xi^{-2}}$ and $\xi = \hbar^2/(m\Delta)$. The correlation length can be written as the well known result $\xi = \hbar v_F/(\Delta k_F)$ [352] with $v_F = k_F/m$. There is only a normalizable solution for $u_0(x) = v_0(x)$ with normalization constant \tilde{N} , and Fig. 2.1 depicts the wave functions in the two regions.

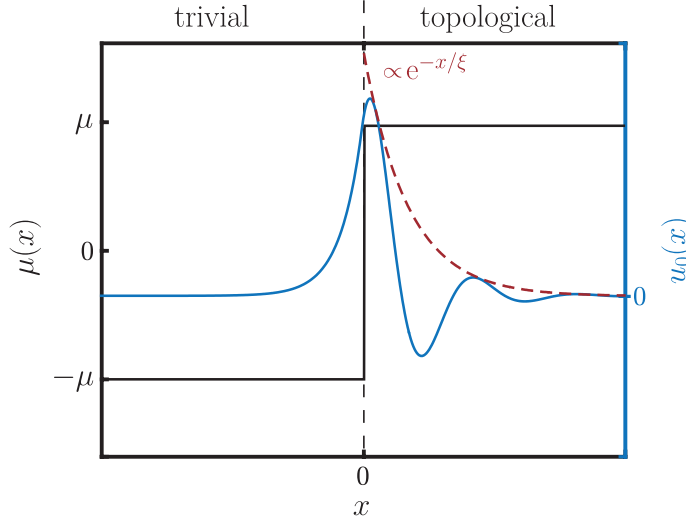


Figure 2.1: Zero-energy wave function (blue) for the domain wall zero mode in a one dimensional spinless p-wave superconductor in the underdamped case Eq. (2.56). The wave function is exponentially localized on the domain wall, with the dashed red line showing the decay $\propto e^{-x/\xi}$ in the topological region. The black line depicts the chemical potential that is given by a step function with a jump at $x = 0$.

As shown in Appendix A.2, there is also a zero-energy solution in the overdamped case $\mu < m\Delta^2/(2\hbar^2)$. Both solutions have in common that they decay in the topological region proportionally to $e^{-x/\xi}$, i.e. exponentially with correlation length $\xi = \hbar^2/(m\Delta)$. Therefore, they are called localized zero modes. In addition, we can define the corresponding quasi-particle operator

$$\gamma = \int dx \left[u_0(x)\psi(x) + v_0(x)\psi^\dagger(x) \right] = \int dx u_0(x) \left[\psi(x) + \psi^\dagger(x) \right], \quad (2.57)$$

which in fact satisfies the Majorana condition $\gamma^\dagger = \gamma$, since $u_0(x) = v_0(x)$.

As a second example, we revisit the spinless p-wave superconductor on a lattice of length L – the Kitaev chain described in the introduction (Sec. 1.3.1). We show how the model is related to the continuum p-wave superconductor and provide more details on its topological classification that we motivated in the introduction. The simple lattice model allows us to get an intuition for the occurrence of the Majorana modes, as discussed by Kitaev in Ref. [86]. To discretize the Hamiltonian Eq. (2.52) on a lattice with a lattice constant a and $N = L/a$ sites, we use

$\int_0^L dx f(x) = \sum_{j=1}^N \int_{a(j-1)}^{aj} dx f(x) \approx \sum_{j=1}^N af(aj)$ and the Taylor expansion

$$\psi(aj + a) = \psi(aj) \pm a\psi'(aj) + \frac{a^2}{2}\psi''(aj) + \dots \quad (2.58)$$

$$\psi''(aj) \approx \frac{1}{a^2} [\psi(aj + a) + \psi(aj - a) - 2\psi(aj)] \quad (2.59)$$

$$\psi'(aj) \approx \frac{1}{2a} [\psi(aj + a) - \psi(aj - a)] . \quad (2.60)$$

Furthermore we define fermion operators at the lattice sites via $c_j = a^{-1/2} \int_{a(j-1)}^{aj} \psi(x) dx \approx \sqrt{a}(\psi(aj) + \psi(aj - a))/2 \approx \sqrt{a}\psi(aj)$. Here all approximations are exact in the limit $a \rightarrow 0$, $N \rightarrow \infty$, $L = Na = \text{const.}$ and we obtain the discretized Hamiltonian

$$\begin{aligned} H = & -t \left[\sum_{j=1}^{N-1} c_j^\dagger c_{j+1} + \sum_{j=2}^N c_j^\dagger c_{j-1} \right] + (2t - \mu) \sum_{j=1}^N c_j^\dagger c_j \\ & - \frac{1}{2} \tilde{\Delta} \left[\sum_{j=1}^{N-1} c_j^\dagger c_{j+1}^\dagger - \sum_{j=2}^N c_j^\dagger c_{j-1}^\dagger \right] + \frac{1}{2} \tilde{\Delta} \left[\sum_{j=1}^{N-1} c_j c_{j+1} - \sum_{j=2}^N c_j c_{j-1} \right] , \end{aligned} \quad (2.61)$$

with $t = \hbar^2/(2ma^2)$ and $\tilde{\Delta} = \Delta/a$. To bring the Hamiltonian to the form used by Kitaev [86], we define $\tilde{\mu} = -(2t - \mu)$ and shift the sums so that they range from $j = 1$ to $j = N - 1$. We also add a constant term $\tilde{\mu}/2$, which does not change the topology described, since it does not depend on the fermion operators. Thus, we obtain the Kitaev chain [86], discussed in the introduction,

$$H = \sum_{j=1}^{N-1} \left[-t (c_j^\dagger c_{j+1} + c_{j+1}^\dagger c_j) - \tilde{\mu} \left(c_j^\dagger c_j + \frac{1}{N+1} c_N^\dagger c_N - \frac{1}{2} \right) + \tilde{\Delta} c_{j+1}^\dagger c_j^\dagger + \tilde{\Delta} c_j c_{j+1} \right] . \quad (2.62)$$

After the substitution, the chemical potential is now $\mu = \tilde{\mu} - 2t$, but our previous calculations were performed for the continuum, which corresponds to the limit value $a \rightarrow 0$ with $t \rightarrow \infty$. We therefore look again at the topological invariant in the discrete case. Here, we assume periodic boundary conditions for Eq. (2.62) for which we obtain the momentum space representation

$$H = \sum_k [-2t \cos(ka) - \tilde{\mu}] c_k^\dagger c_k - \frac{1}{2} \sum_k 2i\tilde{\Delta} \sin(ka) (c_k^\dagger c_{-k}^\dagger + c_k c_{-k}) + \text{const.} \quad (2.63)$$

$$= \frac{1}{2} \sum_k \psi_k^\dagger \mathcal{H}_{\text{BdG}}(k) \psi_k + \text{const.} \quad (2.64)$$

$$\mathcal{H}_{\text{BdG}}(k) = \begin{pmatrix} -2t \cos(ka) - \tilde{\mu} & -2i\tilde{\Delta} \sin(ka) \\ 2i\tilde{\Delta} \sin(ka) & 2t \cos(ka) + \tilde{\mu} \end{pmatrix} , \quad (2.65)$$

with Nambu basis $\psi_k^\dagger = (c_k^\dagger, c_{-k})$. This yields the spectrum $E_k = \pm \sqrt{(2t \cos(ka) + \tilde{\mu})^2 + 4\tilde{\Delta}^2 \sin^2(ka)}$ where the excitation gap closes at $k = 0$ for $\mu = -2t$ and at $k = \pi/a$ for $\mu = 2t$. The difference to before is that there is now another point $k = -k$ given by $k = \pi/a$ in the Brillouin zone $(-\pi/a, \pi/a]$. Thus we obtain the topological invariant

$$\mathcal{Q} = \prod_{k=-k} \text{sgn Pf}[\mathcal{H}_{\text{BdG}}(k) \tau_x] = \prod_{k=-k} \text{sgn} (-2t \cos(ka) - \tilde{\mu}) \quad (2.66)$$

$$= \text{sgn} (-2t - \tilde{\mu}) \text{sgn} (2t - \tilde{\mu}) . \quad (2.67)$$

As $t > 0$, this means that the system is in the topological phase with $\mathcal{Q} = -1$ for $|\tilde{\mu}| < 2t$. On the other hand, it is topologically trivial, i.e. $\mathcal{Q} = +1$, for $|\tilde{\mu}| > 2t$.

To better understand what makes the $\mathcal{Q} = -1$ state special, we consider again the chain with two ends and length L given by Eq. (2.62). As discussed before, Kitaev's trick [86] is to divide each fermion operator c_j into two Majorana operators Eq. (1.54), $\gamma_{2j-1} = c_j^\dagger + c_j$, and Eq. (1.55), $\gamma_{2j} = i(c_j^\dagger - c_j)$, with $\gamma_l^\dagger = \gamma_l$. The Majorana operators fulfill the commutator relations $\{\gamma_m, \gamma_l\} = 2\delta_{lm}$, and thus they are not fermion operators as $\gamma_l^2 = 1$. Instead of N lattice sites there are now $2N$ Majorana sites, and the Majorana operators form the fermion operators by the inverse transformation

$$c_j = \frac{1}{2}(\gamma_{2j-1} + i\gamma_{2j}) \quad (2.68)$$

$$c_j^\dagger = \frac{1}{2}(\gamma_{2j-1} - i\gamma_{2j}) . \quad (2.69)$$

Fig. 2.2 shows such a lattice where ellipses represent lattice sites and dots depict Majorana sites.

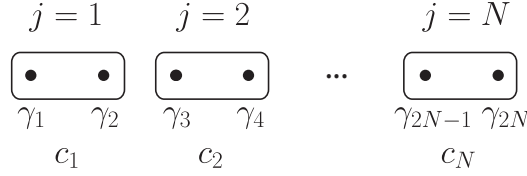


Figure 2.2: Division of the fermion operators into Majorana operators. An ellipse corresponds to one of the N lattice points and the dots represent the $2N$ Majorana sites.

We can always define these Majorana operators regardless of the topological phase, so they do not automatically give rise to Majorana zero modes, as we will see in the following. To study the qualitative differences between topological and trivial phase, we express the Hamiltonian Eq. (2.62) using the Majorana operators and obtain

$$H = \frac{i}{2} \sum_{j=1}^{N-1} [(\tilde{\Delta} - t)\gamma_{2j-1}\gamma_{2j} + (\tilde{\Delta} + t)\gamma_{2j}\gamma_{2j+1}] + \frac{i}{2} \sum_{j=1}^N (-\tilde{\mu})\gamma_{2j-1}\gamma_{2j} . \quad (2.70)$$

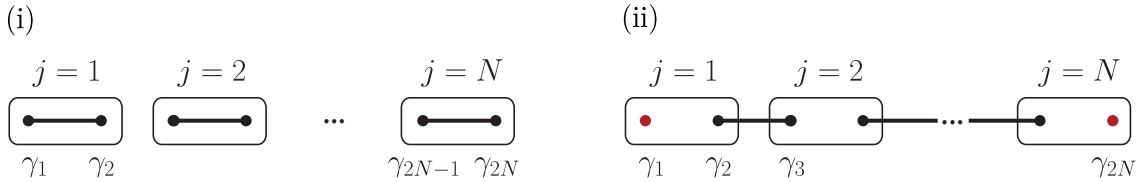


Figure 2.3: Illustration of the two special cases from the introduction. Ellipses represent lattice sites, dots Majorana sites, and black lines are couplings of operators in the Hamiltonian. In the trivial case (i) $\tilde{\Delta} = t = 0$, $\mu < 0$ Majorana operators belonging to the same lattice site are coupled in the Hamiltonian. In contrast, in the topological case (ii) with $\tilde{\Delta} = t > 0$, $\mu = 0$, Majorana operators of consecutive lattice sites are coupled. Thus, the two Majorana operators at the ends of the wire remain uncoupled, i.e. they do not appear in the Hamiltonian and give rise to a non-local zero mode.

We next revisit the two special cases of parameters [86] from the introduction, which lie in the different topological phases according to the topological invariant \mathcal{Q} .

- (i) In the special trivial case $t = \tilde{\Delta} = 0$, $\tilde{\mu} < 0$ with $\mathcal{Q} = +1$ (as $|\tilde{\mu}| > 2t$), the Hamiltonian reduces to

$$H_{(i)} = -\frac{i}{2} \sum_{j=1}^N \tilde{\mu} \gamma_{2j-1} \gamma_{2j} + \text{const.} . \quad (2.71)$$

As shown in Fig. 2.3(i), this on-site term describes an interaction between the Majoranas γ_{2j-1} and γ_{2j} , which belong to the same fermion c_j . Therefore, the ground state, i.e. the state without quasi-particles, is unique.

- (ii) In the special topological case $t = \tilde{\Delta} > 0$, $\tilde{\mu} = 0$ with $\mathcal{Q} = -1$ (as $|\tilde{\mu}| < 2t$), we find the Hamiltonian

$$H_{(ii)} = it \sum_{j=1}^{N-1} \gamma_{2j} \gamma_{2j+1} . \quad (2.72)$$

In this case, as shown in Fig. 2.3(ii), the Majoranas from adjacent lattice sites are coupled. Importantly, the two Majoranas at the ends of the wire γ_1 and γ_{2N} remain uncoupled, and thus they do not appear in the Hamiltonian. Therefore, by reconstructing the fermion operators from the Majoranas [86]

$$a_j = \frac{1}{2}(\gamma_{2j} + i\gamma_{2j+1}) \quad (2.73)$$

$$a_0 = \frac{1}{2}(\gamma_1 + i\gamma_{2N}) \quad (2.74)$$

for $j = 1 \dots N-1$, one obtains

$$H_{(ii)} = 2t \sum_{j=1}^{N-1} \left(a_j^\dagger a_j - \frac{1}{2} \right) + 0 a_0^\dagger a_0 . \quad (2.75)$$

The operator a_0 describes a fermionic zero energy state, so that the states $|0\rangle$ and $a_0^\dagger|0\rangle$ have the same energy. Here, the ground state $|0\rangle$ has even particle number parity and the ground state $a_0^\dagger|0\rangle$ odd parity. In addition, we see that a_0 describes a highly non-local state, since the fermion consists of the Majorana operators, which are localized at the two distant ends of the wire. Most importantly, we see that the Majorana zero modes emerge due to the interplay of the topological phase and the ends of the wire.

We have picked out two examples of parameters from the different topological phases, as they illustrate the formation of the Majorana zero modes particularly well. In addition, in Sec. 2.2.2 we present an analytical solution for an effective spin-polarized p-wave superconducting wire with two ends, consisting of an s-wave superconductor and a semiconductor with strong spin orbit coupling.

2.2.2 Majorana zero modes in Rashba wires

As we have seen in the last section, a spin-polarized p-wave superconductor may exhibit Majorana zero modes, which we are interested in. However, it is a problem to produce a suitable superconductor [336]. While triplet, p-wave superfluidity in ^3He has long been known to exist [353–356], it is difficult to find an intrinsic p-wave superconductor at all [357]. A promising candidate for intrinsic p-wave pairing is Sr_2Ru_4 [285], but there are still doubts about this [286, 287, 358]. Another approach is induced superconductivity in graphene systems [359]. However, it becomes apparent that the experimental realization of Majorana zero modes in intrinsic p-wave superconductors is currently out of reach [286, 360]. Therefore, we pursue a different approach here: We show that effective p-wave pairing can be achieved from the interplay of spin-orbit coupling and s-wave pairing [87, 88, 361]. In addition, a magnetic Zeeman field can be used to achieve that only one spin species is present at the Fermi level, so that the low-energy physics effectively resembles a spin-polarized p-wave superconductor. Fortunately, the required spin-orbit coupling occurs in semiconductor structures such as InAs and GaAs, which are experimentally accessible [95, 130, 362–364]. However, these structures are not superconductors themselves, but superconductivity can be induced using the proximity effect [288–290, 292] by bringing a conventional superconductor in contact with the semiconductor. In this section, we consider a Hamiltonian for such a hybrid wire with spin-orbit coupling and Zeeman term as well as an s-wave pairing term, and describe the pairing in a mean-field approximation similar to Sec. 2.1.

Hamiltonian

We consider a one-dimensional Rashba wire with the momentum space Hamiltonian [87, 88]

$$H = \sum_k \psi_k^\dagger \mathcal{H}(k) \psi_k \quad (2.76)$$

$$\mathcal{H}(k) = \tau_z \otimes [\xi_k \sigma_0 + \hbar \alpha_R k \sigma_x] - \tau_0 \otimes \sigma_z E_z + \tau_x \otimes \sigma_0 \Delta \quad (2.77)$$

in the basis $\psi_k^\dagger = (c_{k\uparrow}^\dagger, c_{k\downarrow}^\dagger, c_{-k\downarrow}, c_{-k\uparrow})$. For reasons of convention, we have chosen a different basis than before, which does not change the previous considerations. Here $\xi_k = \hbar^2 k^2 / (2m^*) - \mu$, the second term describes the Rashba spin-orbit coupling with strength α_R , the third term contains the Zeeman energy E_z due to a magnetic field \mathbf{B}_z , and the last term is the s-wave mean-field pairing with gap $\Delta \in \mathbb{R}$. As we will see, it is important that the spin-orbit coupling direction (here x) is orthogonal to the Zeeman field (here z). In our case the wire extends in the y -direction, so that $k \equiv k_y$. Diagonalizing the BdG Hamiltonian, Eq. (2.77), we find the energy eigenvalues $\{E_\pm(k), -E_\pm(k)\}$ with

$$E_\pm(k) = \sqrt{\xi_k^2 + \hbar^2 k^2 \alpha_R^2 + E_z^2 + \Delta^2 \pm 2\sqrt{E_z^2(\Delta^2 + \xi_k^2) + \hbar^2 k^2 \alpha_R^2 \xi_k^2}}. \quad (2.78)$$

In the normal-conducting case, $\Delta = 0$, the spectrum reduces to

$$E_\pm(k, \Delta = 0) = \xi_k \pm \sqrt{E_z^2 + \hbar^2 k^2 \alpha_R^2}. \quad (2.79)$$

The dispersion for $\Delta = 0$ is shown in Fig. 2.4 for $E_z = 0$ and $E_z > 0$. In the case of $E_z = 0$ the influence of the spin-orbit coupling on the dispersion can be observed. Here the parabolas $\hbar^2 k^2 / (2m^*)$ for the two spin directions are shifted to the side by a momentum l_{so}^{-1} and downwards by the energy E_{so} . If we further add a Zeeman field E_z , we observe that a gap of size $2E_z$ is

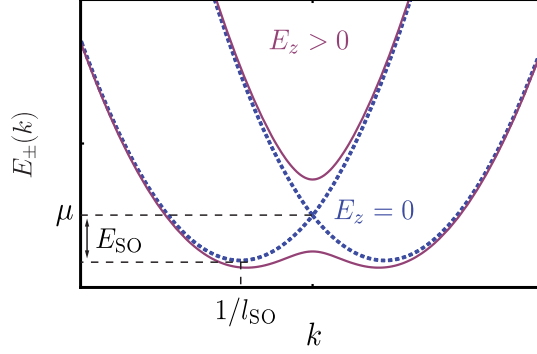


Figure 2.4: Dispersion of the Rashba wire Hamiltonian for $\Delta = 0$, Eq. (2.79). The dotted line depicts the case $E_z = 0$ and the solid line is for $E_z > 0$. The shift of the bands due to the spin-orbit coupling term defines a characteristic energy scale E_{so} and length scale l_{so} . Adapted from [305].

formed at $k = 0$. If the chemical potential μ lies within this gap, the energies at the Fermi level are only twofold degenerate. We will find later that this is crucial for the occurrence of the topological superconducting state.

The quantities

$$l_{\text{so}} = \frac{\hbar}{\alpha_{\text{R}} m^*} \quad (2.80)$$

$$E_{\text{so}} = \frac{\alpha_{\text{R}}^2 m^*}{2} \quad (2.81)$$

define a characteristic length and energy scale for the Rashba wire [365]. It is therefore useful to measure lengths in units of l_{so} and energies in E_{so} , by defining reduced quantities $\tilde{k} = kl_{\text{so}}$, $\tilde{\Delta} = \Delta/E_{\text{so}}$, $\tilde{E}_z = E_z/E_{\text{so}}$, and $\tilde{\mu} = \mu/E_{\text{so}}$. In the reduced units, the Rashba wire Hamiltonian is given by

$$\frac{\mathcal{H}(k)}{E_{\text{so}}} = \tau_z \otimes [(\tilde{k}^2 - \tilde{\mu})\sigma_0 + 2\tilde{k}\sigma_x] - \tau_0 \otimes \sigma_z \tilde{E}_z + \tau_x \otimes \sigma_0 \tilde{\Delta} \quad (2.82)$$

and the dispersion is

$$\frac{E_{\pm}(k)}{E_{\text{so}}} = \sqrt{(\tilde{k}^2 - \tilde{\mu})^2 + 4\tilde{k}^2 + \tilde{E}_z^2 + \tilde{\Delta}^2 \pm 2\sqrt{\tilde{E}_z^2 (\tilde{\Delta}^2 + (\tilde{k}^2 - \tilde{\mu})^2) + 4\tilde{k}^2(\tilde{k}^2 - \tilde{\mu})^2}}. \quad (2.83)$$

In addition to the dispersion for $\Delta = 0$, we next study the influence of spin-orbit coupling on the spin structure by calculating the expectation values of the spin operators

$$\langle S_i \rangle = \frac{\hbar}{2} \langle \sigma_i \rangle. \quad (2.84)$$

We compute the expectation values for the normal-conducting eigenstates \mathbf{v}_{\pm} , which we obtain from the eigenvectors of the upper 2×2 block of $\mathcal{H}/E_{\text{so}}$. For the eigenvalues $E_{\pm}(k)/E_{\text{so}} = \tilde{k}^2 - \tilde{\mu} \pm \sqrt{\tilde{E}_z^2 + 4\tilde{k}^2}$, we find the corresponding eigenvectors (see Appendix A.3)

$$\mathbf{v}_+(k) = \frac{1}{\sqrt{2}} \begin{pmatrix} \text{sgn}(k) \sqrt{1 + \frac{\tilde{E}_z}{\sqrt{\tilde{E}_z^2 + 4\tilde{k}^2}}} \\ \sqrt{1 - \frac{\tilde{E}_z}{\sqrt{\tilde{E}_z^2 + 4\tilde{k}^2}}} \end{pmatrix} \quad \text{and} \quad \mathbf{v}_-(k) = \frac{1}{\sqrt{2}} \begin{pmatrix} -\sqrt{1 - \frac{\tilde{E}_z}{\sqrt{\tilde{E}_z^2 + 4\tilde{k}^2}}} \\ \text{sgn}(k) \sqrt{1 + \frac{\tilde{E}_z}{\sqrt{\tilde{E}_z^2 + 4\tilde{k}^2}}} \end{pmatrix}. \quad (2.85)$$

Thus we obtain the spin expectation values $\langle S_i \rangle_{\pm}(k) = \hbar/2 \mathbf{v}_{\pm}^T(k) \sigma_i \mathbf{v}_{\pm}(k)$ as

$$\langle S_x \rangle_{\pm}(k) = \pm \frac{\hbar}{2} \frac{2\tilde{k}}{\sqrt{\tilde{E}_z^2 + 4\tilde{k}^2}} = \pm \frac{\hbar}{2} \frac{\hbar\alpha_R k}{\sqrt{E_z^2 + \alpha_R^2 \hbar^2 k^2}} \quad (2.86)$$

$$\langle S_y \rangle_{\pm}(k) = 0 \quad (2.87)$$

$$\langle S_z \rangle_{\pm}(k) = \mp \frac{\hbar}{2} \frac{\tilde{E}_z}{\sqrt{\tilde{E}_z^2 + 4\tilde{k}^2}} = \mp \frac{\hbar}{2} \frac{E_z}{\sqrt{E_z^2 + \alpha_R^2 \hbar^2 k^2}}. \quad (2.88)$$

In the limit $\alpha_R \rightarrow 0$, i.e. without spin-orbit coupling, we find $\lim_{\alpha_R \rightarrow 0} \langle S_x \rangle_{\pm}(k) = 0$ and $\lim_{\alpha_R \rightarrow 0} \langle S_z \rangle_{\pm}(k) = \mp \hbar/2$, so that the upper band contains the spin- \uparrow and the lower band the spin- \downarrow states. By switching on spin-orbit coupling, the spins are rotated in x -direction depending on k . In Fig. 2.5 we show the calculated spin structure as a function of momentum k for $\alpha_R \rightarrow 0$ and with spin-orbit coupling $\alpha_R > 0$. As shown in Fig. 2.5(a), without spin orbit coupling, each band can be assigned exactly one spin direction, and since we have chosen the Zeeman field in the z -direction, these are the directions along the z -axis with $\langle S_z \rangle = \pm \hbar/2$. If we switch on only spin-orbit coupling but no Zeeman field, as in Fig. 2.5(b), the spins are polarized in the x -direction, i.e. $\langle S_x \rangle = \pm \hbar/2$, so that spin-orbit coupling effectively acts like a magnetic field in x -direction. Furthermore, the effects of the combination of Zeeman field and spin-orbit coupling on the spin structure are shown in Fig. 2.5(c) and (d) for Zeeman fields of different strengths. One finds that for large Zeeman fields $E_z > E_{\text{so}}$ in the region of the spectrum that is two-fold degenerate ($\mu < |E_z|$), the spin- \uparrow component dominates. In addition, the lower band has only a single minimum if E_z is sufficiently large. If we now add an s-wave pairing term between particles with momentum k and $-k$, the interplay of spin-orbit coupling and the Zeeman field leads to a mixing between effective s-wave and p-wave pairing due to the spin rotation. We can calculate this induced p-wave gap by applying the transformation for diagonalizing the electron and hole sector of the Hamiltonian (as for $\Delta = 0$) to the total Hamiltonian with $\Delta > 0$. We have already calculated the corresponding transformation matrix in Eq. (2.85) as

$$T(k) = \begin{pmatrix} [\mathbf{v}_+(k)]_1 & [\mathbf{v}_-(k)]_1 & 0 & 0 \\ [\mathbf{v}_+(k)]_2 & [\mathbf{v}_-(k)]_2 & 0 & 0 \\ 0 & 0 & [\mathbf{v}_+(-k)]_1 & [\mathbf{v}_-(-k)]_1 \\ 0 & 0 & [\mathbf{v}_+(-k)]_2 & [\mathbf{v}_-(-k)]_2 \end{pmatrix}, \quad (2.89)$$

where the eigenvectors of the hole sector are constructed taking into account the structure of the basis spinor $(c_{k\uparrow}^\dagger, c_{k\downarrow}^\dagger, c_{-k\downarrow}, -c_{-k\uparrow}) = (c_{k,+}^\dagger, c_{k,-}^\dagger, c_{-k,-}, -c_{-k,+})T^\dagger$. Thus, we find the transformed Hamiltonian

$$T(k)\mathcal{H}(k)T^\dagger(k) = \begin{pmatrix} E_+(k) & 0 & \Delta_{+-}(k) & \Delta_{++}(k) \\ 0 & E_-(k) & \Delta_{--}(k) & \Delta_{-+}(k) \\ \Delta_{+-}^\dagger(k) & \Delta_{--}^\dagger(k) & -E_-(-k) & 0 \\ \Delta_{++}^\dagger(k) & \Delta_{-+}^\dagger(k) & 0 & -E_+(-k) \end{pmatrix}, \quad (2.90)$$

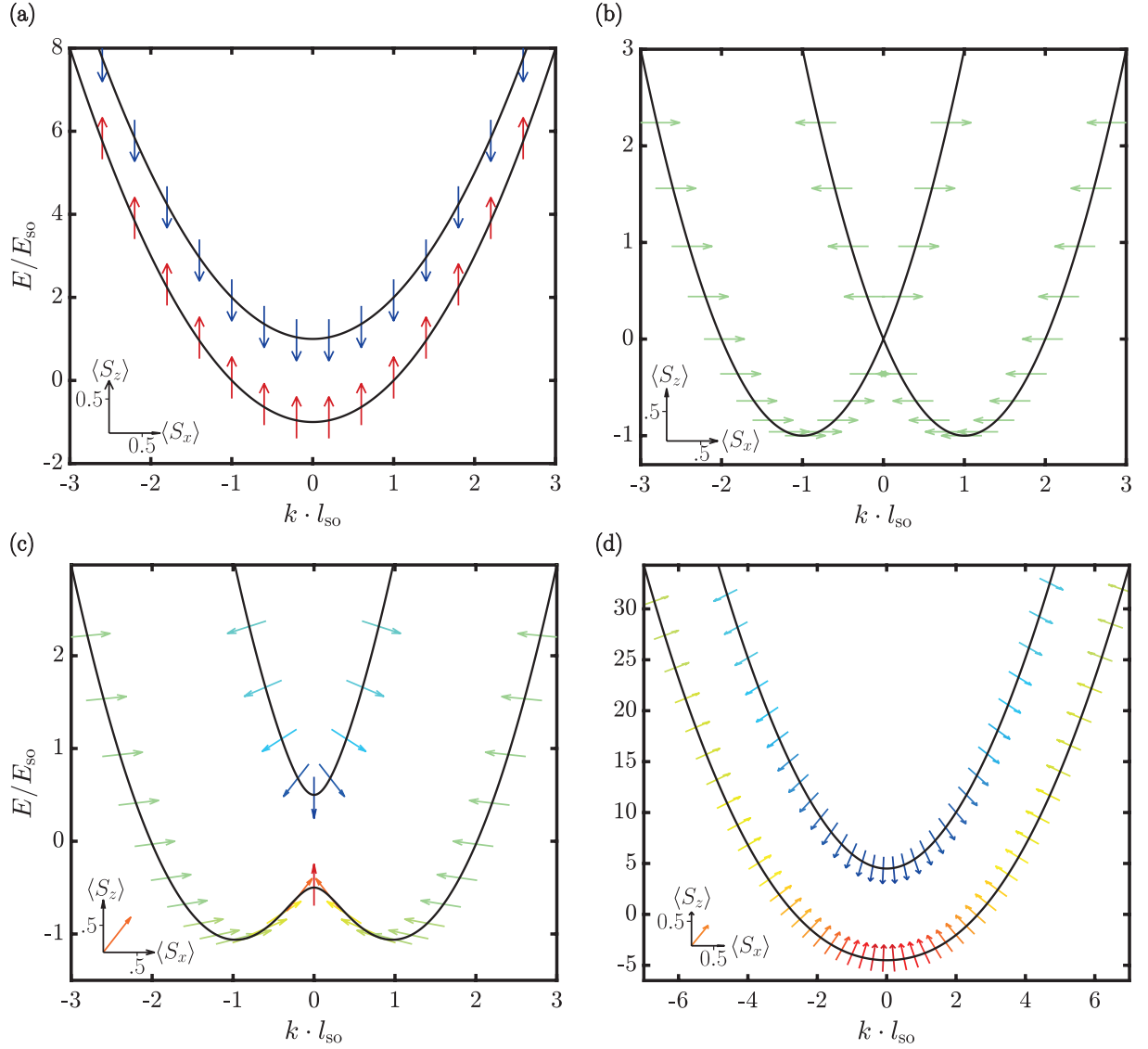


Figure 2.5: Spin structure in the bands of the Rashba wire eigenstates (a) for $\alpha_R = 0$ without spin-orbit coupling with $E_z > 0$, (b) for $\alpha_R > 0$, i.e. with spin-orbit coupling but without Zeeman term $E_z = 0$, (c) with spin-orbit coupling and small Zeeman term $E_z = 0.5 E_{so}$, and (d) with spin-orbit coupling and larger Zeeman field of $E_z = 4.5 E_{so}$. Here, the solid lines are the energy bands and the arrows are the spin vectors in the S_x - S_z -plane. The colors indicate the rotation angle of the spins from the \uparrow -direction (angle 0, red) to the \downarrow direction (180° , blue).

with $E_{\pm}/E_{so} = \tilde{k}^2 - \tilde{\mu} \pm \sqrt{\tilde{E}_z^2 + 4\tilde{k}^2}$ and [366]

$$\begin{aligned} \Delta_{+-}(k) &= \Delta([v_-(k)]_1[v_-(-k)]_1 - [v_+(k)]_1[v_+(-k)]_1) \\ &= \frac{-\Delta\tilde{E}_z}{\sqrt{\tilde{E}_z^2 + 4\tilde{k}^2}} = \Delta_{-+}(k) \end{aligned} \quad (2.91)$$

$$\begin{aligned} \Delta_{++}(k) &= \Delta([v_-(k)]_1[v_-(-k)]_2 + [v_+(k)]_1[v_+(-k)]_2) \\ &= \frac{\Delta 2\tilde{k}}{\sqrt{\tilde{E}_z^2 + 4\tilde{k}^2}} = -\Delta_{--}(k) . \end{aligned} \quad (2.92)$$

It becomes apparent that both effective p-wave interband pairing Δ_{++}, Δ_{--} and s-wave intra-band pairing Δ_{+-}, Δ_{-+} occur. If we consider the topologically non-trivial case from Sec. 2.2.1, i.e. where E_z and μ are chosen such that effectively only the lower band is occupied, only the effective interband p-wave pairing $\Delta_{--}(k)$ is relevant and given by

$$\Delta_p(k) = \frac{\Delta 2\tilde{k}}{\sqrt{\tilde{E}_z^2 + 4\tilde{k}^2}} = \frac{\Delta \hbar \alpha_R k}{\sqrt{E_z^2 + \hbar^2 \alpha_R^2 k^2}}. \quad (2.93)$$

As expected, this shows that the Rashba spin-orbit coupling yields effective triplet pairing for a suitable choice of the parameters. In addition, the Zeeman field lifts the degeneracy at Fermi level for suitable chemical potentials. However, we also see that the p-wave gap, $\Delta_p \sim E_z^{-1}$, vanishes in the limit $E_z \rightarrow \infty$. It is therefore crucial that E_z is large enough such that effectively only the lower band is occupied, but not so large that the induced triplet pairing vanishes or superconductivity in the parent superconductor is destroyed [82].

Gap and topological classification

After showing that the Rashba wire is effectively a spin-polarized p-wave superconductor for an appropriate choice of Zeeman field and chemical potential, we next examine its topological classification. For this we first consider the excitation gap for the case $\Delta > 0$, which we can derive from the dispersion $E_{\pm}(k)$, Eq. (2.83), as the energy of the lower band at either the Fermi level k_F or at $k = 0$ depending on at which of these two points the energy is smaller. Here, the Fermi momentum for the lower band can be obtained from the point $E_{-}(k_F) = 0$ and is given by

$$l_{\text{so}} k_F = \sqrt{\tilde{\mu} + 2 + \sqrt{4 + \tilde{E}_z^2 + 4\tilde{\mu}}} . \quad (2.94)$$

Thus, the excitation gap is determined via

$$\begin{aligned} E_{\text{gap}} = \min_k(E_{-}(k)) &\approx \min \left\{ |E_z - \sqrt{\Delta^2 + \mu^2}|, E_{-}(k_F) \right\} \\ &\approx \min \left\{ |E_z - \sqrt{\Delta^2 + \mu^2}|, |\Delta_p(k_F)| \right\} , \end{aligned} \quad (2.95)$$

where the last approximate equality only holds if interband pairing is negligible.

Fig. 2.6(a) depicts the excitation gap as a function of the Zeeman field, where it becomes apparent that the gap is closing at $E_z = \sqrt{\Delta^2 + \mu^2}$. In order to show that the closing of the gap is again associated with a topological phase transition and to calculate which topological phase exists in which parameter region, we calculate the topological invariant in the following. For this we first consider the symmetries of the Hamiltonian, Eq. (2.82):

- (i) The Hamiltonian has pseudo time reversal symmetry $\mathcal{T} = \tau_0 \otimes \sigma_z \mathcal{K}$ with $\mathcal{T}^2 = +1$, since $\mathcal{T}\mathcal{H}(k)\mathcal{T}^{-1} = \mathcal{H}(-k)$.
- (ii) The Hamiltonian also has particle-hole symmetry $\mathcal{P} = \tau_y \otimes \sigma_y \mathcal{K}$ with $\mathcal{P}^2 = +1$ and $\mathcal{P}\mathcal{H}(k)\mathcal{P}^{-1} = -\mathcal{H}(-k)$.
- (iii) Because of the presence of \mathcal{P} and \mathcal{T} symmetry, we can also construct $\mathcal{C} = \tau_y \otimes \sigma_x$, where the chiral operator satisfies $\mathcal{C}\mathcal{H}(k)\mathcal{C}^{-1} = -\mathcal{H}(k)$.

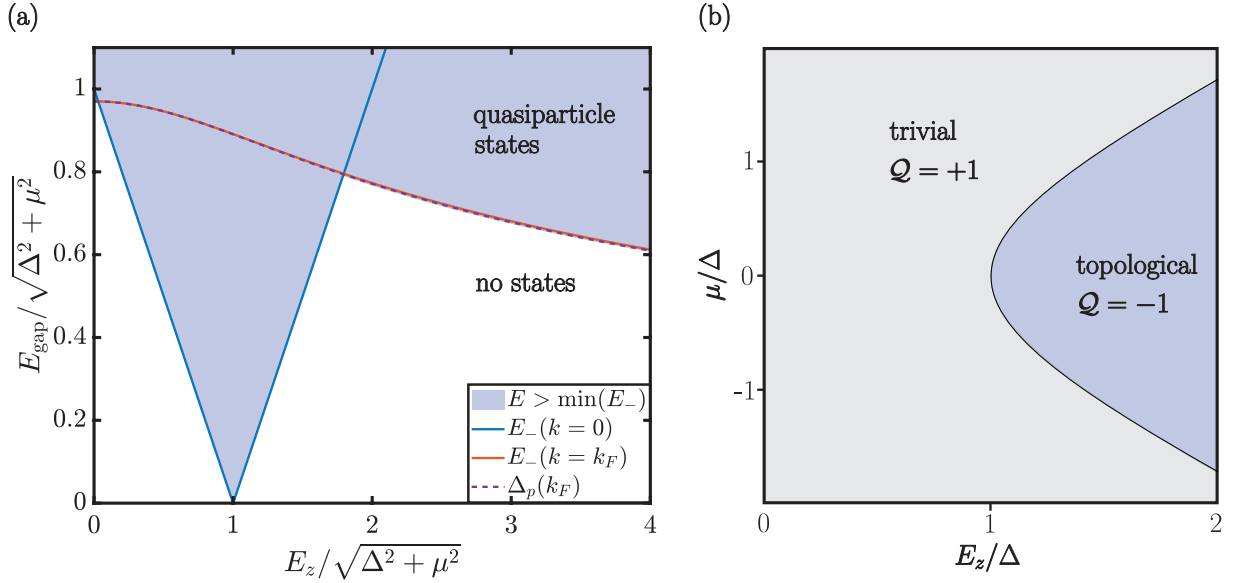


Figure 2.6: (a) Excitation gap at $k = 0$ and at Fermi-level k_F together with the p-wave gap $|\Delta_p(k_F)|$ for $\mu = 0.5E_{\text{so}}$, $\Delta = 2E_{\text{so}}$. The energies $E > \min(E_-)$ for which quasi-particle states exist are marked in light blue [cf. 265]. Here $\min(E_-)$ is obtained numerically from Eq. (2.83). (b) Phase diagram of the Rashba wire in the chemical potential μ , Zeeman field E_z plane [cf. 84].

Thus, the Hamiltonian is in the BDI symmetry class, whose \mathbb{Z} topological invariant can be determined as a winding number. As discussed in Sec. 2.1, however, it is sufficient for us to determine the Pfaffian invariant Eq. (2.41), which gives the parity of the winding number [346, 347]. With $\mathcal{P} = \tau_y \otimes \sigma_y \mathcal{K} = U_{\mathcal{P}} \mathcal{K}$, which yields $U_{\mathcal{P}} = \tau_y \otimes \sigma_y$, we find

$$\mathcal{Q} = \prod_{k=-k} \text{sgn}(\text{Pf}[\mathbf{i}\mathcal{H}(k)U_{\mathcal{P}}]) \quad (2.96)$$

$$= - \prod_{k=-k} \text{sgn} \text{Pf} \begin{pmatrix} 0 & \tilde{\Delta} & 2\tilde{k} & \tilde{E}_z + \tilde{k}^2 + \tilde{\mu} \\ -\tilde{\Delta} & 0 & \tilde{E}_z + \tilde{k}^2 - \tilde{\mu} & -2\tilde{k} \\ 2\tilde{k} & -\tilde{E}_z - \tilde{k}^2 + \tilde{\mu} & 0 & -\tilde{\Delta} \\ -\tilde{E}_z + \tilde{k}^2 - \tilde{\mu} & -2\tilde{k} & \tilde{\Delta} & 0 \end{pmatrix} \quad (2.97)$$

$$= - \prod_{k=-k} \text{sgn}(-\tilde{\Delta}^2 + 4\tilde{k}^2 + \tilde{E}_z^2 - (\tilde{k}^2 - \tilde{\mu})^2) \quad (2.98)$$

$$= -\text{sgn}(\tilde{E}_z^2 - (\tilde{\Delta}^2 + \tilde{\mu}^2)) = \begin{cases} +1 & \text{for } E_z^2 < \Delta^2 + \mu^2 \\ -1 & \text{for } E_z^2 > \Delta^2 + \mu^2 \end{cases}. \quad (2.99)$$

Thus, the system is in the topological phase for $|E_z| > \sqrt{\Delta^2 + \mu^2}$ and in the trivial phase for $|E_z| < \sqrt{\Delta^2 + \mu^2}$, and the gap closes at $|E_z| = \sqrt{\Delta^2 + \mu^2}$. The corresponding phase diagram is shown in Fig. 2.6(b).

Majorana condition

For the case of the p-wave superconductor, we considered a domain wall as a boundary between the topological and trivial phase in the previous section. There we found an analytical solution

for the zero energy state, which is localized at the phase boundary. Due to the absence of the spin degree of freedom, the Hamiltonian is only a 2×2 matrix. In addition, particle-hole symmetry $\mathcal{P}_p = \tau_x \mathcal{K}$ and pseudo-time reversal symmetry $\mathcal{T}_p = \mathcal{K}$ could be utilized to simplify the calculation for the zero-energy state $\psi_0 = (u_0, v_0)^T$. It turned out to be helpful that $\psi_0 = (u_0, v_0)^T = \mathcal{P}_p \psi_0 = (v_0^*, u_0^*)^T = \mathcal{T}_p \psi_0 = (u_0^*, v_0^*)^T$ could be fulfilled simultaneously by a real solution $u_0 = v_0 \in \mathbb{R}$.

In the case of the Majorana wire, the situation is more complicated by the fact that we have to consider the spin degrees of freedom. In addition, for a zero energy solution $\psi_0 = (u_\uparrow, u_\downarrow, v_\downarrow, -v_\uparrow)^T$, application of the symmetry operator yields $\mathcal{P}\psi_0 = (v_\uparrow^*, v_\downarrow^*, u_\downarrow^*, -u_\uparrow^*)^T$ and $\mathcal{T}\psi_0 = (u_\uparrow^*, -u_\downarrow^*, v_\downarrow^*, v_\uparrow^*)^T$. However, here $\psi_0 = \mathcal{T}\psi_0 = \mathcal{P}\psi_0 \Leftrightarrow \psi_0 = 0$, which is why a state cannot have eigenvalue +1 for both operators \mathcal{P} and \mathcal{T} at the same time. However, $\mathcal{P}\psi_0$ and $\mathcal{T}\psi_0$ are also zero energy states because of $\mathcal{P}\mathcal{H}(y)\mathcal{P}^{-1} = -\mathcal{H}(y)$ and $\mathcal{T}\mathcal{H}(y)\mathcal{T}^{-1} = \mathcal{H}(y)$. If a single zero energy state ψ_0 exists, we can choose it so that $\mathcal{T}\psi_0 = \psi_0$ and $\mathcal{P}\psi_0 = i\psi_0$. However, we will later consider long wires with two ends, where two such solutions occur with energies $\pm\varepsilon \approx 0$. If exactly two such quasi zero energy states $\psi_0^{(1)}$ and $\psi_0^{(2)}$ exist, then (since $\mathcal{T}\psi_0$ lies in the subspace to eigenvalue $E \approx 0$ spanned by $\psi_0^{(1)}$ and $\psi_0^{(2)}$) one can always choose the state such that $\mathcal{T}\psi_0^{(1)} = \psi_0^{(1)}$. Further, we can take $\psi_0^{(2)} = \mathcal{P}\psi_0^{(1)}$, since $\mathcal{P}\psi_0^{(1)}$ also lies in the eigenspace of $E \approx 0$ and $\mathcal{P}\psi_\varepsilon = \psi_{-\varepsilon}$. Hence, $\mathcal{T}\psi_0^{(1)}$ and $\mathcal{P}\psi_0^{(1)}$ form an orthonormal basis in the eigenspace corresponding to the eigenvalue $E \approx 0$.

It will be advantageous for transport calculations to separate the particle-hole redundant eigenvectors [367], which is automatically the case for states with energy $E \neq 0$, since $\mathcal{P}\psi_E = \psi_{-E}$ and $E \neq -E$ in this case. For the Majorana zero modes with $E = 0$ we can achieve this separation by selecting the two eigenmodes as described above. Then because of $\{\mathcal{T}, \mathcal{P}\} = 0$, the eigenstates are chosen such that

	\mathcal{T}	\mathcal{P}
$\psi_0^{(1)}$	$\psi_0^{(1)}$	$\psi_0^{(2)}$
$\psi_0^{(2)}$	$-\psi_0^{(2)}$	$\psi_0^{(1)}$

This choice also implies

$$\psi_0^{(1)} = \begin{pmatrix} u_\uparrow \\ u_\downarrow \\ v_\downarrow \\ -v_\uparrow \end{pmatrix} = \begin{pmatrix} u_\uparrow^* \\ -u_\downarrow^* \\ v_\downarrow^* \\ v_\uparrow^* \end{pmatrix} = \mathcal{T}\psi_0^{(1)} \quad (2.100)$$

and therefore $u_\uparrow, iu_\downarrow, v_\downarrow, iv_\uparrow \in \mathbb{R}$.

We next consider a one-dimensional Rashba wire of length L with two ends at $y = 0$ and $y = L$. The real space Hamiltonian in unitless parameters with confinement potential $\tilde{V}(y)$ is given by

$$\mathcal{H}/E_{\text{so}} = \tau_z \otimes \left[-\sigma_0 \left(\partial_y^2 + \mu + \tilde{V}(y) \right) - 2i\sigma_x \partial_y \right] - \tilde{E}_z \tau_0 \otimes \sigma_z + \tilde{\Delta} \tau_x \otimes \sigma_0. \quad (2.101)$$

Assuming that the wire is in the topological phase, i.e. $E_z > \sqrt{\Delta^2 + \mu}$, there are interfaces between the trivial phase (outside the wire) and topological phase (inside the wire) at both ends. We therefore expect two Majorana zero modes, each exponentially located at one end of the wire. For a symmetric potential $V(y)$ in the wire, the Hamiltonian Eq. (2.82) has the

additional symmetry Π that reflects $y \rightarrow L - y$ and applies a complex conjugation [304], i.e.

$$\Pi = (y \rightarrow L - y)\mathcal{K} \quad (2.102)$$

$$\mathcal{H}(y) = \Pi\mathcal{H}(y)\Pi^{-1}. \quad (2.103)$$

Then, in addition to $\mathcal{T}\psi_0^{(1)} = \psi_0^{(1)}$ (for exactly two zero modes $\psi_0^{(1)}, \psi_0^{(2)}$), we can also ensure that

$$\Pi\psi_0^{(1)} = \begin{pmatrix} u_{\uparrow}(L - y) \\ -u_{\downarrow}(L - y) \\ v_{\downarrow}(L - y) \\ +v_{\uparrow}(L - y) \end{pmatrix} = \lambda\psi_0^{(1)} \quad (2.104)$$

with $\lambda = \pm 1$. Because Π commutes with the particle-hole operator, $[\Pi, \mathcal{P}] = 0$, this also implies that

$$\Pi\psi_0^{(2)} = \Pi\mathcal{P}\psi_0^{(1)} = \lambda\psi_0^{(2)}. \quad (2.105)$$

Thus $u_{\uparrow}, v_{\downarrow}$ have equal parity under spatial inversion, while u_{\downarrow} and v_{\uparrow} have opposite parity. Eq. (2.105) already shows that the chosen eigenvectors in the zero energy subspace are a superposition of the localized Majorana zero modes. This is also evident because of $u_{\sigma} \neq v_{\sigma}^*$, since we have chosen the $\psi_0^{(i)}$ as eigenstates of \mathcal{T} and not \mathcal{P} . We can also construct eigenstates of \mathcal{P} using the following linear combinations of the zero energy states

$$\psi_{\pm} = e^{\mp i\pi/4} \frac{\psi_0^{(1)} \pm i\psi_0^{(2)}}{\sqrt{2}} = \frac{e^{\mp i\pi/4}}{\sqrt{2}} \begin{pmatrix} u_{\uparrow}(y) \mp i v_{\uparrow}(y) \\ u_{\downarrow}(y) \pm i v_{\downarrow}(y) \\ v_{\downarrow}(y) \mp i u_{\downarrow}(y) \\ -[v_{\uparrow}(y) \pm i u_{\uparrow}(y)] \end{pmatrix}. \quad (2.106)$$

For these zero energy states, application of the symmetry operators yields

$$\mathcal{P}\psi_{\pm}(y) = \psi_{\pm}(y) \quad (2.107)$$

$$\mathcal{T}\psi_{\pm}(y) = \pm i\psi_{\pm}(y) \quad (2.108)$$

$$\Pi\psi_{\pm}(y) = \lambda\psi_{\mp}(y) \quad (2.109)$$

with $\lambda = \pm 1$. The particle-hole invariance ensures that the Majorana condition between hole and particle wavefunctions in ψ_{\pm} is fulfilled. If ψ_{-} is localized at the left end of the wire, i.e. $\psi_{-} = \chi(y) \propto e^{-y/\xi}$, then $\psi_{+} = \lambda\Pi\psi_{-} \propto \chi(L - y)$ is the Majorana state localized at the right edge. Thus, if two localized zero modes ψ_{\pm} are present, Eq. (2.106) and the localization imply that $u_{\uparrow}(y) + i v_{\uparrow}(y) = 0$ for all y that are not at the left end, and $u_{\uparrow}(y) - i v_{\uparrow}(y) = 0$ for any y not located at the right edge – and analogously for the inverse spin. Therefore, as long as the Majorana modes do not overlap, which is true if their localization length is much shorter than the wire length, $\xi \ll L/2$, the Majorana condition for localized zero modes can be written as

$$|u_{\sigma}| = |v_{\sigma}|. \quad (2.110)$$

In the case of a short wire with overlap between the Majorana wave functions, the states move slightly away from zero energy to $E = \pm\varepsilon$ (see Sec. 2.2.3). Then the above equations apply only approximately, and it becomes exact in the limit $L \gg \xi$.

In the following, we choose the wavefunction $u_\sigma^{(E)}$ and $v_\sigma^{(E)}$ for all energies $E \geq 0$, so that $\mathcal{H}\psi_E(y) = E\psi_E(y)$ with $\psi_E(y) = (u_\uparrow^{(E)}, u_\downarrow^{(E)}, v_\downarrow^{(E)}, -v_\uparrow^{(E)})^T = \mathcal{T}\psi_E(y) = \Pi\psi_E(y)$. As a result, even for the zero energy case, they do not fulfill $u_\sigma = v_\sigma^*$ – often referred to the Majorana condition. Instead, we have shown that $|u_\sigma| = |v_\sigma|$ is the correct Majorana condition in this representation. We will see later that this representation is useful for transport calculations and if necessary one can still convert to the conventional Majorana representation by using Eq. (2.106), such that $\mathcal{P}\psi_\pm = \psi_\pm$. For a wire of finite length it turns out that in the presence of a Zeeman field, $E_z > 0$, all states with energy $E > 0$ are non-degenerate⁴. Furthermore, for all $E \geq 0$, we choose

$$u_\uparrow^{(E)}, \mathbf{i}u_\downarrow^{(E)}, \mathbf{i}v_\downarrow^{(E)}, v_\uparrow^{(E)} \in \mathbb{R}, \quad (2.111)$$

and for any symmetrical potential the parities of the wave functions under spatial inversion are related by [304]

$$\text{parity}(u_\uparrow^{(E)}) = \text{parity}(v_\downarrow^{(E)}) \neq \text{parity}(u_\downarrow^{(E)}) = \text{parity}(v_\uparrow^{(E)}). \quad (2.112)$$

Analytic solution for a Majorana zero mode

By making use of the symmetries, it is also possible to obtain an analytical solution for the Majorana states (cf. supplemental material of [368]; parts of the analytical solution presented here have been published in [131]). To obtain the zero-energy solution, we consider the BdG equation

$$\mathcal{H}\psi = 0. \quad (2.113)$$

From ψ_- , Eq. (2.106), it becomes apparent that the form of a Majorana solution is given by

$$\psi = (\chi_\uparrow, \chi_\downarrow, \mathbf{i}\chi_\downarrow, \mathbf{i}\chi_\uparrow)^T. \quad (2.114)$$

Using this vector in the BdG equation, one finds a system of four differential equations

$$\begin{aligned} -\partial_y^2 \chi_\uparrow - \tilde{\mu} \chi_\uparrow - \tilde{E}_z \chi_\uparrow - 2\mathbf{i}\partial_y \chi_\downarrow + \mathbf{i}\tilde{\Delta} \chi_\downarrow &= 0 \\ -\partial_y^2 \chi_\downarrow - \tilde{\mu} \chi_\downarrow + \tilde{E}_z \chi_\downarrow - 2\mathbf{i}\partial_y \chi_\uparrow + \mathbf{i}\tilde{\Delta} \chi_\uparrow &= 0 \\ \mathbf{i}\partial_y^2 \chi_\downarrow + \mathbf{i}\tilde{\mu} \chi_\downarrow - \mathbf{i}\tilde{E}_z \chi_\downarrow - 2\partial_y \chi_\uparrow + \tilde{\Delta} \chi_\uparrow &= 0 \\ \mathbf{i}\partial_y^2 \chi_\uparrow + \mathbf{i}\tilde{\mu} \chi_\uparrow + \mathbf{i}\tilde{E}_z \chi_\uparrow - 2\partial_y \chi_\downarrow + \tilde{\Delta} \chi_\downarrow &= 0. \end{aligned} \quad (2.115)$$

Furthermore, Eq. (2.108) allows us to define the real quantities $e^{i\pi/4}\chi_\uparrow = \hat{\chi}_\uparrow \in \mathbb{R}$, $\mathbf{i}e^{i\pi/4}\chi_\downarrow = \hat{\chi}_\downarrow \in \mathbb{R}$, such that the BdG equations are reduced to the two real equations

$$\begin{aligned} -\partial_y^2 \hat{\chi}_\uparrow - \tilde{\mu} \hat{\chi}_\uparrow - \tilde{E}_z \hat{\chi}_\uparrow - 2\partial_y \hat{\chi}_\downarrow + \tilde{\Delta} \hat{\chi}_\downarrow &= 0 \\ -\partial_y^2 \hat{\chi}_\downarrow - \tilde{\mu} \hat{\chi}_\downarrow + \tilde{E}_z \hat{\chi}_\downarrow + 2\partial_y \hat{\chi}_\uparrow - \tilde{\Delta} \hat{\chi}_\uparrow &= 0. \end{aligned} \quad (2.116)$$

For the zero-energy solution, which is located at the left edge, we make the ansatz [368]

$$\begin{pmatrix} \hat{\chi}_\uparrow \\ \hat{\chi}_\downarrow \end{pmatrix} = e^{-A\tilde{y}} \begin{pmatrix} \varrho_\uparrow \\ \varrho_\downarrow \end{pmatrix} \quad (2.117)$$

⁴ This may seem like a contradiction as in momentum space $E(-k) = E(k)$ for periodic boundary conditions. However, similarly to the difference between periodic and open boundary conditions for a particle in a box, in the open case, k and $-k$ describe the same state such that there is in fact no degeneracy.

with boundary conditions $\hat{\chi}(y=0) = 0$. For a wire with two ends, this is only an approximate solution, since the exponential function can only be approximately zero at the right end such that there are corrections due to the overlap of the Majorana wave functions from both ends. For a very long wire $L \gg \xi_{\text{eff}}$ this is a good approximation, which becomes exact in the limit $L \rightarrow \infty$. Using the ansatz in Eq. (2.116), we obtain

$$\begin{pmatrix} -A^2 - \tilde{\mu} - \tilde{E}_z & 2A + \tilde{\Delta} \\ -2A - \tilde{\Delta} & -A^2 - \tilde{\mu} + \tilde{E}_z \end{pmatrix} \begin{pmatrix} \varrho_{\uparrow} \\ \varrho_{\downarrow} \end{pmatrix} = 0 . \quad (2.118)$$

For a non-trivial solution, the determinant of the coefficient matrix must vanish, yielding the following quartic equation in A

$$0 = (A^2 + \tilde{\mu})^2 - \tilde{E}_z^2 + (2A + \tilde{\Delta})^2 \quad (2.119)$$

$$= A^4 + A^2(2\tilde{\mu} + 4) + A(4\tilde{\Delta}) + \tilde{\mu}^2 - \tilde{E}_z^2 + \tilde{\Delta}^2 , \quad (2.120)$$

where for a localized Majorana wave function only solutions with $\text{Re}(A) > 0$ are relevant. This quartic equation can be solved analytically, already being in the reduced form $A^4 + \alpha A^2 + \beta A + \gamma = 0$ with

$$\alpha = 2\tilde{\mu} + 4 \quad (2.121)$$

$$\beta = 4\tilde{\Delta} \quad (2.122)$$

$$\gamma = \tilde{\mu}^2 - \tilde{E}_z^2 + \tilde{\Delta}^2 . \quad (2.123)$$

This allows us to use the solution formula for reduced quartic equations [369] for which we define

$$P = -\frac{\alpha^2}{12} - \gamma \quad (2.124)$$

$$Q = -\frac{\alpha^3}{108} + \frac{\alpha\gamma}{3} - \frac{\beta^2}{8} \quad (2.125)$$

$$U = \left(-\frac{Q}{2} + \sqrt{\frac{Q^2}{4} + \frac{P^3}{27}} \right)^{1/3} \quad (2.126)$$

$$Y = -\frac{5}{6}\alpha + U - \frac{P}{3U} \quad (2.127)$$

$$W = \sqrt{\alpha + 2Y} \quad (2.128)$$

$$Z = \frac{\beta}{2W} . \quad (2.129)$$

Then the solutions are given by

$$\begin{aligned} A_1 &= \frac{1}{2} \left[W + \sqrt{W^2 - 4(\alpha + Y + Z)} \right] \\ A_2 &= \frac{1}{2} \left[W - \sqrt{W^2 - 4(\alpha + Y + Z)} \right] \\ A_3 &= \frac{1}{2} \left[-W + \sqrt{W^2 - 4(\alpha + Y - Z)} \right] \\ A_4 &= \frac{1}{2} \left[-W - \sqrt{W^2 - 4(\alpha + Y - Z)} \right] . \end{aligned} \quad (2.130)$$

Factoring Eq. (2.120) using the four roots, i.e. $0 = (A - A_1)(A - A_2)(A - A_3)(A - A_4)$, we directly find that

$$\begin{aligned} 0 &= A_1 + A_2 + A_3 + A_4 \\ \tilde{\mu}^2 - \tilde{E}_z^2 + \tilde{\Delta}^2 &= A_1 A_2 A_3 A_4, \end{aligned} \quad (2.131)$$

and in the topological region, $\gamma < 0$, in addition $W, Y, Z, \alpha \in \mathbb{R}$ and a Majorana solution exists [368] with $\text{Re}(A_1), \text{Re}(A_2), \text{Re}(A_3) > 0$ and $\text{Re}(A_4) < 0$. We therefore need the coefficient of the term corresponding to A_4 to be zero in order to obtain a localized and normalizable solution in the limit $L \rightarrow \infty$. Then, for each A_i the system of equations Eq. (2.118) has to be solved. The solution can be expressed as

$$\begin{pmatrix} \varrho_{\uparrow, i} \\ \varrho_{\downarrow, i} \end{pmatrix} = \mathcal{N}_i \begin{pmatrix} 2A_i + \tilde{\Delta} \\ A_i^2 + \tilde{\mu} + \tilde{E}_z \end{pmatrix}, \quad (2.132)$$

where \mathcal{N}_i is a normalization constant. From Eq. (2.130) one finds in the topological regime that $A_1 = A_2^*$ and $A_3 \in \mathbb{R}$, and we therefore define

$$A_1 = \tilde{\xi}_1^{-1} + i\tilde{k}_{\text{eff}} \quad (2.133)$$

$$A_2 = \tilde{\xi}_1^{-1} - i\tilde{k}_{\text{eff}} \quad (2.134)$$

$$A_3 = \tilde{\xi}_2^{-1}. \quad (2.135)$$

Using $\xi_1 = \tilde{\xi}_1 l_{\text{so}}$, $\xi_2 = \tilde{\xi}_2 l_{\text{so}}$, and $y = \tilde{y} l_{\text{so}}$, $k_{\text{eff}} = \tilde{k}_{\text{eff}}/l_{\text{so}} \approx k_F/l_{\text{so}}$, the solution in the topological range is given by [131]

$$\begin{aligned} \hat{\chi}_L(y) = \mathcal{N} &\left[e^{-y/\xi_2} \begin{pmatrix} 2\tilde{\xi}_2^{-1} + \tilde{\Delta} \\ \tilde{\xi}_2^{-2} + \tilde{\mu} + \tilde{E}_z \end{pmatrix} \right. \\ &+ e^{-y/\xi_1} \left\{ a e^{i k_{\text{eff}} y} \begin{pmatrix} 2(\tilde{\xi}_1^{-1} + i\tilde{k}_{\text{eff}}) + \tilde{\Delta} \\ (\tilde{\xi}_1^{-1} + i\tilde{k}_{\text{eff}})^2 + \tilde{\mu} + \tilde{E}_z \end{pmatrix} \right. \\ &\quad \left. \left. + b e^{-i k_{\text{eff}} y} \begin{pmatrix} 2(\tilde{\xi}_1^{-1} - i\tilde{k}_{\text{eff}}) + \tilde{\Delta} \\ (\tilde{\xi}_1^{-1} - i\tilde{k}_{\text{eff}})^2 + \tilde{\mu} + \tilde{E}_z \end{pmatrix} \right\} \right]. \end{aligned} \quad (2.136)$$

Furthermore, from the boundary condition $\hat{\chi}_L(y=0) = 0$ we obtain [131]

$$a = b^* = \frac{(i\tilde{\xi}_2 + \tilde{\xi}_1(-i + \tilde{k}_{\text{eff}}\tilde{\xi}_2))(-2 + 2(\tilde{E}_z + \tilde{\mu})\tilde{\xi}_1\tilde{\xi}_2 - \tilde{\Delta}(\tilde{\xi}_1 + \tilde{\xi}_2) + i\tilde{k}_{\text{eff}}\tilde{\xi}_1(2 + \tilde{\Delta}\tilde{\xi}_2))}{4\tilde{k}_{\text{eff}}(1 + \tilde{\xi}_1(\tilde{\Delta} - (\tilde{E}_z - \tilde{k}_{\text{eff}}^2 + \tilde{\mu})\tilde{\xi}_1))\tilde{\xi}_2^2}. \quad (2.137)$$

Using $\chi_{\uparrow} = e^{-i\pi/4}\hat{\chi}_{\uparrow}$ and $\chi_{\downarrow} = -ie^{-i\pi/4}\hat{\chi}_{\downarrow}$ we have thus analytically found the zero-energy solution at the left end, and the corresponding solution at the right end is given by

$$\chi_R(y) = \Pi \chi_L(y) = \chi_L^*(L - y). \quad (2.138)$$

As we can see from Eq. (2.136), there are two decay constants ξ_1 and ξ_2 , where ξ_1 is dominant for large Zeeman fields and ξ_2 dominates close to the phase transition. Using the relations for the A_i , Eq. (2.131), we can express A_3 in terms of A_1 and A_2 , and using ξ_1 , we find

$$\xi_2 = A_3^{-1} l_{\text{so}} = \left(\frac{-1}{\xi_1} + \sqrt{\frac{1}{\xi_1^2} - \frac{\tilde{\mu}^2 + \tilde{\Delta}^2 - \tilde{E}_z^2}{\xi_1^{-2} + k_F^2}} \right)^{-1}. \quad (2.139)$$

Unfortunately, the analytical solution is very lengthy, which makes it difficult to learn something about the Majorana wave functions from it. However, we have already considered the p-wave superconductor and its similarity to the Rashba wire. The next step is therefore to derive approximations from the analytical solution by using the knowledge of the p-wave superconductor.

Approximations

To estimate the effective gap using the induced p-wave gap Δ_p , it would be advantageous to approximate the lengthy analytical expression for the localization lengths of the Majorana modes ξ_1 and ξ_2 using the result in the p-wave limit. As we discussed in Sec. 2.2.1, the coherence length of a superconductor is given by $\xi = \hbar v_F / E_g$ [352], where E_g is the effective gap. In the case considered here, $E_-(k)/E_{\text{so}} = \tilde{k}^2 - \mu - \sqrt{\tilde{E}_z^2 + 4\tilde{k}^2}$ and $E_-(k_F) = 0$ so that we obtain the Fermi momentum and Fermi velocity as [131]

$$k_{\text{eff}} l_{\text{so}} \approx \tilde{k}_F = \left(2 + \tilde{\mu} + \sqrt{\tilde{E}_z^2 + 4 + 4\tilde{\mu}} \right)^{1/2}, \quad (2.140)$$

$$\frac{\hbar v_F}{l_{\text{so}} E_{\text{so}}} = \frac{\partial_{\tilde{k}} E_-}{l_{\text{so}} E_{\text{so}}} \Big|_{\tilde{k}_F} = 2\tilde{k}_F - \frac{4\tilde{k}_F}{\sqrt{\tilde{E}_z^2 + 4\tilde{k}_F^2}}. \quad (2.141)$$

Here we again used the reduced quantities $\tilde{k}_F = k_F l_{\text{so}}$, $\tilde{E}_z = E_z / E_{\text{so}}$, $\tilde{\mu} = \mu / E_{\text{so}}$. Thus, by using Eq. (2.93) we obtain an approximation for the coherence length [131]

$$\xi_1 \approx \xi = \frac{\hbar v_F}{\Delta_{p,\text{ind}}} = \frac{1}{\Delta} \left(\frac{\hbar}{\alpha_R m} \sqrt{E_z^2 + \hbar^2 \alpha_R^2 k_F^2} - \hbar \alpha_R \right), \quad (2.142)$$

$$\frac{\xi_1}{l_{\text{so}}} \approx \frac{1}{\Delta} \left(\sqrt{\tilde{E}_z^2 + 4\tilde{k}_F^2} - 2 \right). \quad (2.143)$$

However, for Zeeman fields close to the phase transition, the gap closing is dominated by an s-wave type behavior $E_g \propto |E_z - \sqrt{\Delta^2 + \mu^2}|$. Using the approximation $\xi_1 \approx \xi$, we can also approximate the s-wave correction length with Eq. (2.139) as [131]

$$\xi_2 \approx \xi_s \equiv \left(\frac{-1}{\xi} + \sqrt{\frac{1}{\xi^2} - \frac{\tilde{\mu}^2 + \tilde{\Delta}^2 - \tilde{E}_z^2}{\xi^{-2} + k_F^2}} \right)^{-1} \sim \frac{1}{E_z - \sqrt{\Delta^2 + \mu^2}}. \quad (2.144)$$

Here, the s-wave correlation length ξ_s diverges at the phase transition according to the energy of the closing gap, $E_g = |E_z - \sqrt{\Delta^2 + \mu^2}|$, consistent with $\xi_s \sim 1/E_g$.

We depict these approximations together with the full result of the analytical solution for the Majorana zero modes ξ_1 and ξ_2 in Fig. 2.7(b). Here, the dominant correlation length $\xi_{\text{eff}} = \max\{\xi, \xi_2\}$ first decreases at the beginning of the topological range and then increases with E_z starting at about $E_z \approx 3.2 E_{\text{so}}$. In this range the p-wave correlation length $\hbar v_F / \Delta_p$ is an excellent approximation of the analytical result. In the limit $E_z / E_{\text{so}} \rightarrow \infty$, the localization length is asymptotically proportional to E_z , such that the Rashba wire in the topological region $E_z \gg \sqrt{\Delta^2 + \mu^2}$ behaves approximately like a p-wave superconductor from the last section with $k_F \Delta \rightarrow \Delta_p$ and $\xi_k \rightarrow \hbar^2 k^2 / (2m^*) - \mu - \sqrt{E_z^2 + \hbar^2 \alpha_R^2 k^2}$. For $E_z \gtrsim \sqrt{\Delta^2 + \mu^2}$ at the beginning of the topological region, however, the effective correlation length of the Majorana mode is proportional to the inverse of the Zeeman field, $\xi_{\text{eff}} \sim E_z^{-1}$, and a description by the spin-polarized p-wave superconductor is not justified. However, since the transition to $\xi_{\text{eff}} \sim E_z$

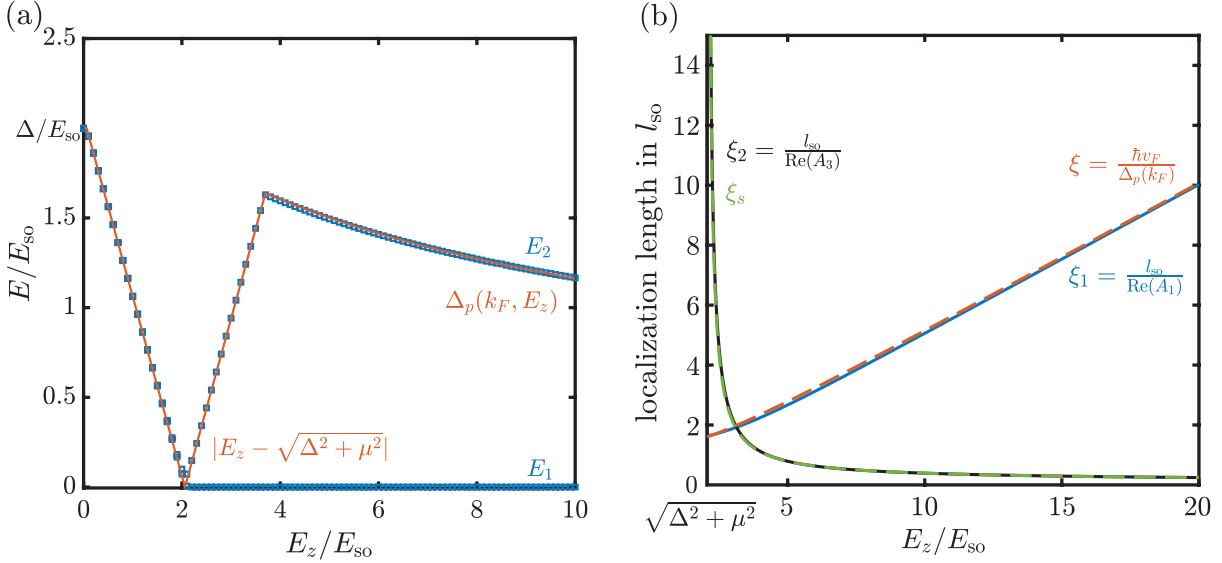


Figure 2.7: Validity of the approximations in the analytical solution for the Majorana zero mode: (a) The lowest two positive energy eigenvalues numerically computed for $\mu = 0.5 E_{\text{so}}$ and $\Delta = 2 E_{\text{so}}$ as a function of the Zeeman field E_z (blue solid line). The approximate excitation gap $\min\{\Delta_p, E_-(k=0)\}$ is depicted in red. The numerical calculations were performed for a very long wire $L = 130 l_{\text{so}}$. (b) Analytically determined localization length $\xi_1 = \text{Re}(A_1)^{-1} l_{\text{so}}$ and $\xi_2 = \text{Re}(A_3)^{-1} l_{\text{so}}$ of the Majorana wave functions as a function of the Zeeman field E_z for $\mu = 0.5 E_{\text{so}}$, $\Delta = 2 E_{\text{so}}$ and $a = 0.026 l_{\text{so}}$. In comparison, the approximations of the correlation lengths from the p-wave solution Eq. (2.143) and the s-wave correlation length Eq. (2.144) are shown (dashed lines), in excellent agreement with the analytical results.

occurs relatively shortly after entering the topological regime, one can often use the simple solution of the p-wave superconductor when interpreting numerical results for the Majorana state.

Comparison with numerical solution

For our later transport calculations we additionally need eigenstates with higher energies $E > 0$, which we have to determine numerically. For this purpose, we discretize the Hamiltonian Eq. (2.82) on a lattice of N sites with lattice constant $a = L/N$. We again work in reduced units $\tilde{a} = a/l_{\text{so}}$, $\tilde{\mu} = \mu/E_{\text{so}}$, $\tilde{E}_z = E_z/E_{\text{so}}$, $\tilde{\Delta} = \Delta/E_{\text{so}}$ and proceed analogously to Eq. (2.58), Eq. (2.59), and Eq. (2.60) of the previous section. Using the $N \times N$ matrices r and ℓ with $[r]_{ij} = \delta_{i,j-1}$ and $[\ell]_{ij} = \delta_{i,j+1}$, we can discretize the derivatives as $\partial_{\tilde{y}} = (\ell - r)/(2\tilde{a})$ and $\partial_{\tilde{y}}^2 = (r + \ell - 2\mathbb{1}_N)/\tilde{a}^2$. Thus, we find the discretized Hamiltonian

$$H = \Psi^\dagger \mathcal{H} \Psi \quad (2.145)$$

$$\begin{aligned} \mathcal{H}/E_{\text{so}} = & \frac{1}{\tilde{a}^2} [r + \ell - 2\mathbb{1}_N] \otimes \tau_z \otimes \sigma_0 - \tilde{\mu} \mathbb{1}_N \otimes \tau_z \otimes \sigma_0 - \frac{i}{\tilde{a}} [\ell - r] \otimes \tau_z \otimes \sigma_x \\ & - \tilde{E}_z \mathbb{1}_N \otimes \tau_0 \otimes \sigma_z + \tilde{\Delta} \mathbb{1}_n \otimes \tau_x \otimes \sigma_0, \end{aligned} \quad (2.146)$$

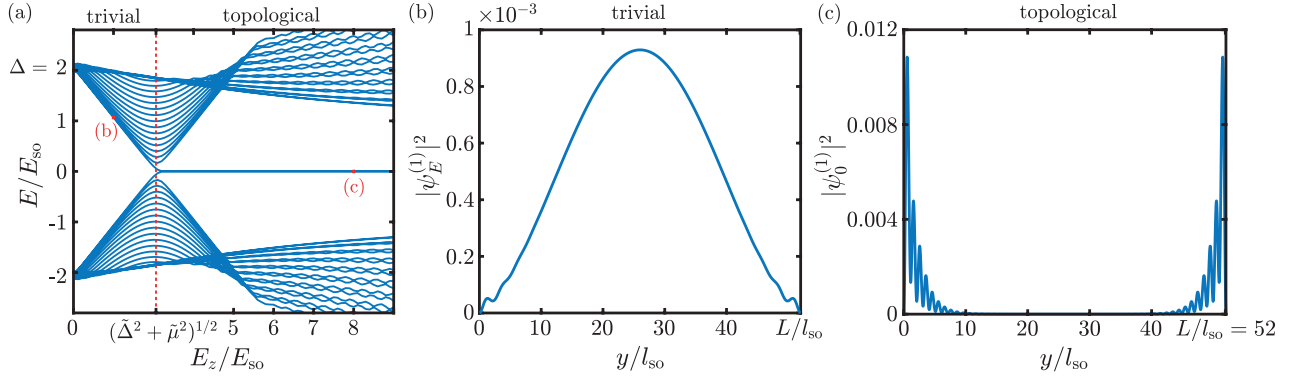


Figure 2.8: Eigenvalues and selected eigenstates of \mathcal{H}/E_{so} for $\mu = 0.5 E_{so}$, $\Delta = 2 E_{so}$ and $a = 0.026 l_{so}$. (a) 20 lowest energy eigenvalues as a function of the Zeeman field E_z . The red crosses mark the values for which the eigenvectors are shown in panel (b) and (c). At the red, vertical line, $E_z = \sqrt{\mu^2 + \Delta^2}$, the system transitions from the trivial to the topological phase by a closing of the bulk gap. After reopening of the gap for large fields, MZMs remain in the middle of the excitation gap. (b) Example of an eigenfunction $|\Psi_E|^2$ for $E_z = 1 E_{so}$ in the trivial phase corresponding to the lowest positive eigenvalue $E = 1.066 E_{so}$. The trivial wave function is extended through the wire and vanishes at the ends due to the open boundary conditions. (c) Example for MZMs $|\Psi_0^{(1)}|^2$ in the topological phase for $E_z = 8 E_{so}$ with energy $E = 1.12 \cdot 10^{-6} E_{so}$. The Majorana wave function is strongly localized at both wire ends.

where the basis spinor is given by

$$\Psi^\dagger = (\psi^\dagger(0), \psi^\dagger(a), \dots, \psi^\dagger(ja), \dots, \psi^\dagger(Na)) \quad (2.147)$$

$$\psi^\dagger(ja) = (c_{j\uparrow}^\dagger, c_{j\downarrow}^\dagger, c_{j\downarrow}, -c_{j\uparrow}) \quad (2.148)$$

To obtain the numerical solution we diagonalize the matrix \mathcal{H}/E_{so} using `Matlab`.

In Fig. 2.8(a) we show the 20 energy eigenvalues with the smallest absolute values as a function of the Zeeman energy E_z , numerically obtained for $\mu = 0.5 E_{so}$, $\Delta = 2 E_{so}$, and $a = 0.026 l_{so}$. We observe that after the transition to the topological phase, two states with nearly zero energy occur which lie in the middle of the excitation gap. In the trivial region these states are absent. In addition, the wave functions of the lowest level in the trivial and topological regime are shown in Fig. 2.8(b) and (c), respectively. It is apparent that the trivial wave function is extended over the wire, while the zero energy solutions are exponentially localized at the ends. As described, $\Psi_0^{(1)} = (u_{0\uparrow}, u_{0\downarrow}, v_{0\downarrow}, -v_{0\uparrow})^T$ is a superposition of the two Majorana wave functions $\Psi_+ = \chi_R$ and $\Psi_- = \chi_L$, which we obtain from the numerical results $\Psi_0^{(1)}$ using Eq. (2.106). We depict these Majorana wave functions in Fig. 2.9(a) and (b). For comparison, the solid red line shows the analytical solution for the Majorana wave functions according to Eq. (2.136) – in excellent agreement with the numerical results. Instead of the complete analytical solution, in Fig. 2.9(b) we plot only the exponentially decaying envelope $\propto e^{-(L-y)/\xi_{eff}}$ with effective correlation length $\xi_{eff} = \max\{\xi, \xi_2\}$. In order to also take into account the signs of the wave functions, we plot $u_\uparrow = \text{Re}(u_\uparrow)$ and $-iv_\downarrow = \text{Im}(v_\downarrow)$ in Fig. 2.9(c), which are superpositions of the two Majorana wave functions. We can see that u_\uparrow and v_\downarrow indeed have inverse parity under spatial inversion $y \rightarrow L-y$. This behavior under spatial inversion will become important later when we study electronic transport, in which the couplings between leads and the Majorana wire are determined by the values of the wave functions at the ends of the wire. In Fig. 2.7(a), we compare the estimator

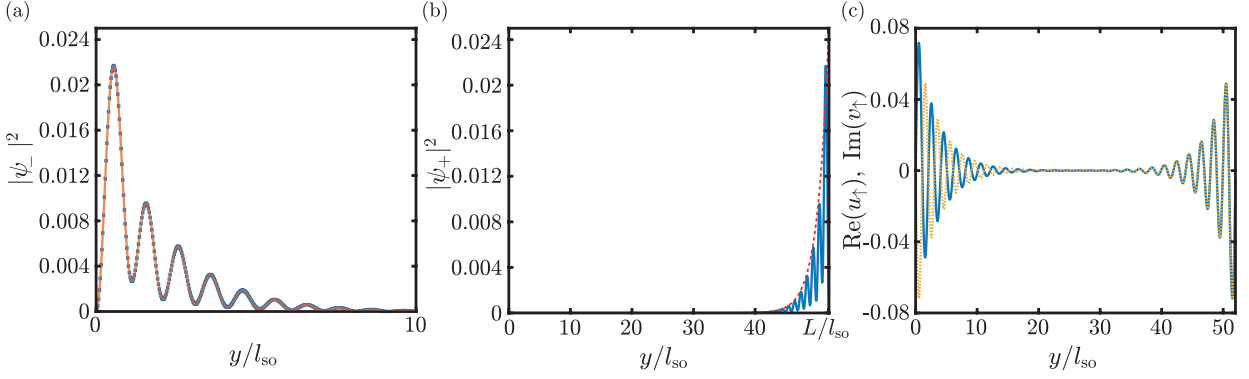


Figure 2.9: Majorana wave function in the topological phase for $E_z = 4 E_{\text{so}}$, $\mu = 0.5 E_{\text{so}}$, $\Delta = 2 E_{\text{so}}$, and $a = 0.026 l_{\text{so}}$. (a) Left localized analytical solution for the Majorana wave function Eq. (2.136) (red solid line) together with the numerical solution (blue squares) calculated from the solutions $\Psi_{\pm\varepsilon}$ for $\varepsilon = 1.12 \cdot 10^{-6} E_{\text{so}}$ using Eq. (2.106). (b) Right localized numerical solution (blue) together with $\mathcal{N}e^{-y/\xi_{\text{eff}}}$, where the localization length used for the envelope curve depicted in red is determined from the analytical solution as $\xi_{\text{eff}} = \max\{\xi, \xi_2\}$. We find excellent agreement between the numerical and analytical results. (c) Numerically calculated eigenfunctions $\text{Re}(u_{\uparrow})$ and $\text{Im}(v_{\uparrow})$ using the representation of the Majorana wave function according to Eq. (2.100), which demonstrates that the parity under spatial inversions is given by Eq. (2.112).

for the excitation gap from the p-wave superconductor $E_g = \min\{|E_z - \sqrt{\Delta^2 + \mu^2}|, \Delta_p(k_F)\}$ with the numerical energy eigenvalues E_i . For a long wire, we find excellent agreement between E_g and E_1 , and in the topological region the Majorana solution has zero energy, $E_0 \approx 0$. In the case of a shorter wire, which we will discuss in Sec. 2.2.3, oscillations of E_1 around E_g and of E_0 around zero occur.

Differences to trivial phase

To better understand what distinguishes the topological phase from the trivial phase of the Rashba wire, we consider linear combinations of the states with the lowest energies $\pm E_1$ [370] in both trivial and topological regime

$$\psi_{\pm}^{(E_1)} = e^{\mp i\pi/4} \frac{\psi_{E_1} \pm i\psi_{-E_1}}{\sqrt{2}}, \quad (2.149)$$

such that action of the symmetry operators yields $\mathcal{T}\psi_{\pm E_1} = \pm\psi_{E_1}$ and $\mathcal{P}\psi_{\pm E_1} = \psi_{\mp E_1}$. As can be seen from Eq. (2.106), these are the Majorana states in the case of $E_1 \approx 0$, however, if $E_1 > 0$ they are no eigenstates of the Hamiltonian. In Fig. 2.10 we depict $|\psi_{\pm}|^2$ together with the spin expectation value in the z -direction

$$\langle S_z \rangle_{\pm} / \hbar = \psi_{\pm}^{\dagger} (\tau_z \otimes \sigma_x) \psi_{\pm} \quad (2.150)$$

for the trivial and topological cases. As we can see in Fig. 2.10(a) and (c), in contrast to the trivial phase, the overlap between $\psi_{+}^{(E_1)}$ and $\psi_{-}^{(E_1)}$ disappears in the topological region. In fact, it can be shown that the smaller the overlap between $\psi_{+}^{(E_1)}$ and $\psi_{-}^{(E_1)}$, the closer E_1 is to zero [370]. We will further discuss the implications of this observation in two different ways: (i) In Sec. 2.2.3, we study Majorana zero modes in short wires where the Majorana wave functions

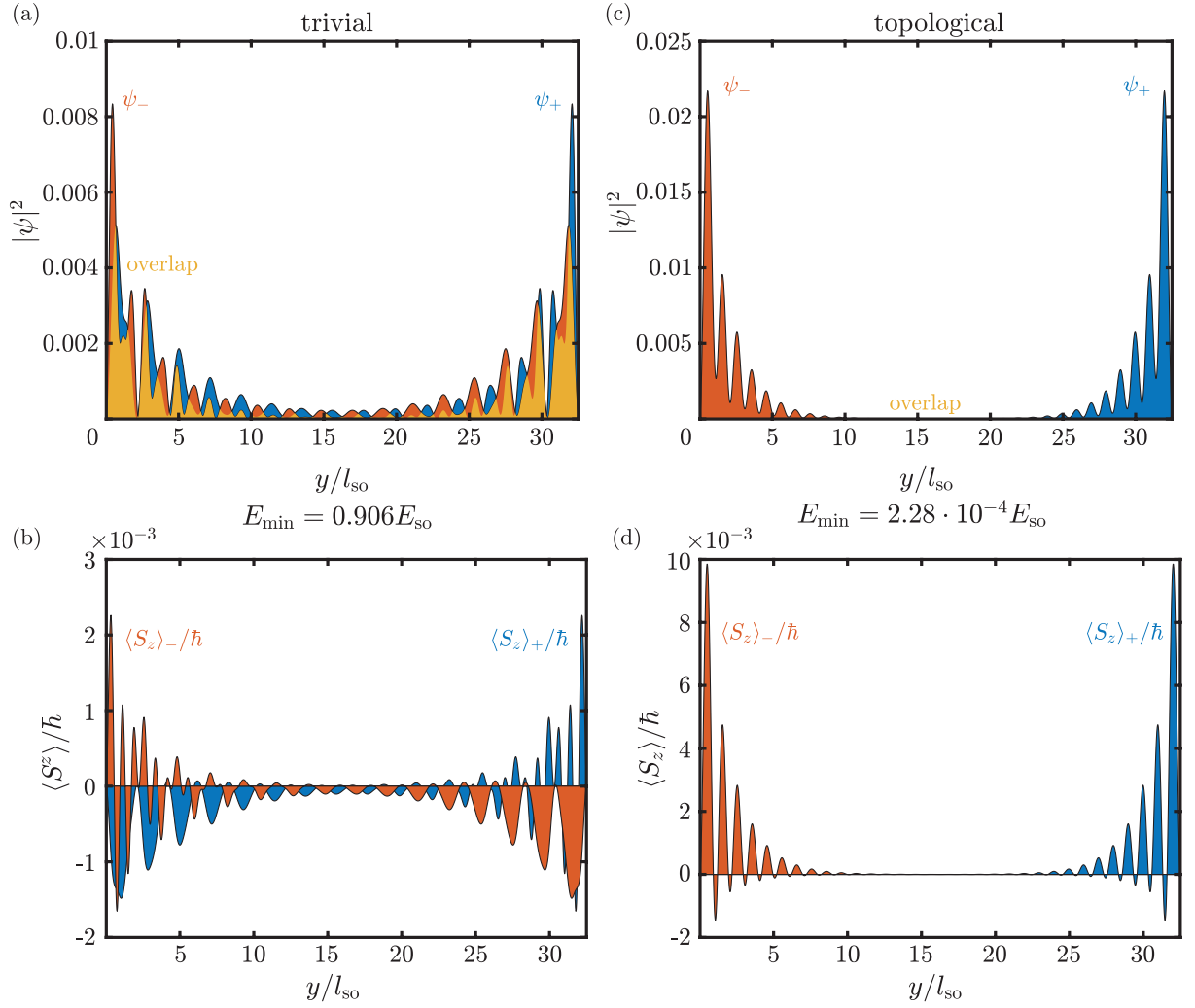


Figure 2.10: Wave function $|\psi_{\pm}^{(E)}|^2$, Eq. (2.149), and spin expectation value $\langle S_z \rangle_{\pm}$, Eq. (2.150), in the trivial phase for $E_z = 5 E_{\text{so}}$, $\mu = 8 E_{\text{so}}$ (a,b) and in the topological phase for $E_z = 7 E_{\text{so}}$, $\mu = 0.5 E_{\text{so}}$ (c,d). In all cases, the numerical solutions are computed for $\Delta = 2 E_{\text{so}}$, $a = 0.026 l_{\text{so}}$, and $L = 32.5 l_{\text{so}}$. (a) Wave function $|\Psi_+^{(E)}|^2$ for the lowest state with $E = 0.906 E_{\text{so}}$ in blue and $|\psi_-^{(E)}|^2$ in red. The overlap of both wave functions is shown in orange. (b) Spin expectation value $\langle S_z \rangle_+$ in blue and $\langle S_z \rangle_-$ in red for the solutions $\psi_{\pm}^{(E)}$ from (a). (c) Wave functions for the Majorana level with $E \approx 0$, where $|\psi_+^{(E)}|^2$ is shown in blue and $|\psi_-^{(E)}|^2$ in red. Here, the wave functions $\psi_{\pm}^{(E)}$ are not eigenstates of the Hamiltonian if $E > 0$. (d) Spin expectation value $\langle S_z \rangle_+$ in blue and $\langle S_z \rangle_-$ in red for the solutions $\psi_{\pm}^{(E)}$ from (c). Similar results are shown in Ref. [370].

from both ends overlap in the middle of the wire, which leads to a finite energy $E_1 \ll \Delta_p$. (ii) We artificially reduce the overlap in Sec. 2.2.4 by a smooth confinement potential in the trivial phase to generate Andreev-bound states. These are trivial excitations, which also lie in the excitation gap at energies $E_1 \ll \Delta_p$.

In addition, Fig. 2.10(b) and (d) depict the spin expectation value in z -direction for the $\psi_{\pm}^{(E_1)}$ states. In the topological region (Fig. 2.10(d)) it becomes apparent that mainly spin- \uparrow

components are present, since the Majorana solutions are approximately spin polarized. In the trivial region (Fig. 2.10(b)), however, both spin- \uparrow and \downarrow components are present at the Fermi level. It can also be observed that the overlap at the ends of the wire is mainly due to opposite spin components.

2.2.3 Finite energy Majorana modes

So far we have only considered the case of very long wires, where the length L is much larger than the localization length of the Majorana zero modes. We so far assumed that there is no overlap between the localized MZMs from the two wire ends. For this case we found an analytical zero-energy solution, Eq. (2.136), of the form

$$\chi_L(y) = (\chi_{L\uparrow}(y), \chi_{L\downarrow}(y), i\chi_{L\downarrow}(y), i\chi_{L\uparrow}(y))^T, \quad (2.151)$$

which is localized at the left end, and from the Π symmetry we obtained

$$\chi_R(y) = (\chi_{L\uparrow}^*(L-y), \chi_{L\downarrow}^*(L-y), -i\chi_{L\downarrow}^*(L-y), i\chi_{L\uparrow}^*(L-y))^T, \quad (2.152)$$

as the solution localized at the right end. Here we now consider wire lengths that are still larger than the localization length, but not so large that the overlap is negligible. In this case the two wave functions hybridize to eigenstates

$$\psi_{\pm E_0}(y) = \frac{e^{\pm i\pi/4}}{\sqrt{2}} (\chi_L(y) \mp i\chi_R(y)). \quad (2.153)$$

These eigenstates then have energies $\pm E_0$, given by [368]

$$E_0 = \frac{\int dy \psi_{E_0}^\dagger \mathcal{H} \psi_{E_0}}{\int dy \psi_{E_0}^\dagger \psi_{E_0}} \quad (2.154)$$

$$= \frac{1}{2} \frac{\int_0^L dy [\chi_L^\dagger(y) \mathcal{H} \chi_L(y) + \chi_R^\dagger(y) \mathcal{H} \chi_R(y) - i\chi_L^\dagger(y) \mathcal{H} \chi_R(y) + i\chi_R^\dagger(y) \mathcal{H} \chi_L(y)]}{\int dy \psi_{E_0}^\dagger \psi_{E_0}}. \quad (2.155)$$

Using the definitions from Eq. (2.106) of the last section, we see that $\mathcal{P}\chi_{L,R}(y) = \chi_{L,R}(y)$ and $\Psi_{E_0} = \mathcal{P}\Psi_{-E_0} = (u_\uparrow, u_\downarrow, v_\downarrow, -v_\uparrow)^T$. If we use the symmetries \mathcal{P} and Π of the wire Hamiltonian together, we obtain the unitary symmetry $\mathcal{P}\Pi = U_{\mathcal{P}}(y \rightarrow L-y)$ with $(\mathcal{P}\Pi)^\dagger \mathcal{H} \mathcal{P}\Pi = -\mathcal{H}$. Acting with this operator on the Majorana wave functions, we obtain $\mathcal{P}\Pi\chi_L = \chi_R$ and $\chi_L^\dagger(\mathcal{P}\Pi)^\dagger = \chi_R^\dagger$, such that

$$\int dy \chi_L^\dagger \mathcal{H} \chi_L = - \int dy \chi_L^\dagger (\mathcal{P}\Pi)^\dagger \mathcal{H} (\mathcal{P}\Pi) \chi_L = - \int dy \chi_R^\dagger \mathcal{H} \chi_R. \quad (2.156)$$

Hence, the diagonal terms add to zero, and the energy is determined by the overlaps according to

$$E_0 = -\frac{i}{2} \int dy [\chi_L^\dagger \mathcal{H} \chi_R - \chi_R^\dagger \mathcal{H} \chi_L]. \quad (2.157)$$

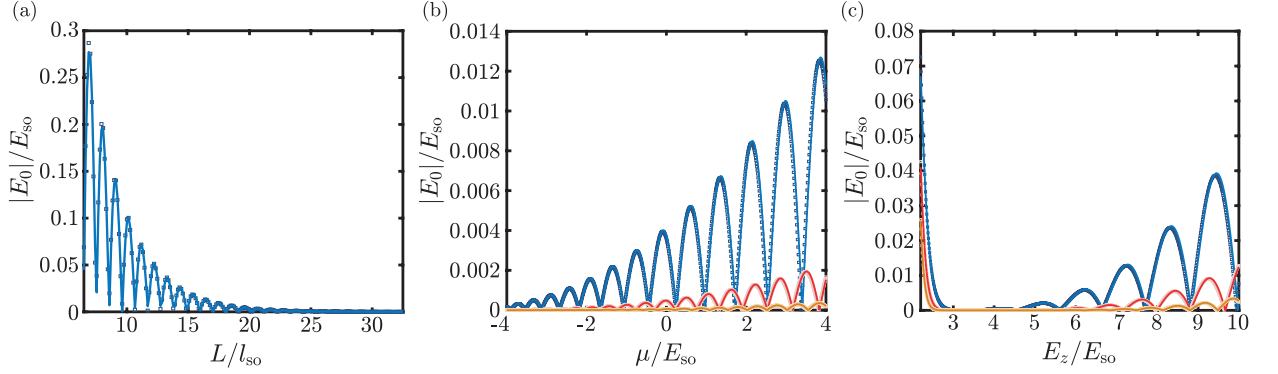


Figure 2.11: Energy of the Majorana eigenstate $|E_0|$ as a function of (a) wire length L for $E_z = 6 E_{so}$, $\mu = 0.5 E_{so}$, (b) chemical potential μ for $L \in \{19.5, 26.0, 32.5\} l_{so}$, $E_z = 6 E_{so}$, and (c) Zeeman energy E_z for $L \in \{19.5, 26.0, 32.5\} l_{so}$, $\mu = 0.5 E_{so}$. Solid lines are analytical results using Eq. (2.160) and the squares are determined numerically. Similar results are shown in Ref. [368] and for the Kitaev chain in Ref. [373].

To calculate the energy we use the analytical solution Eq. (2.136), where the Majorana wave functions truncated at the ends⁵ are given by

$$\chi_L^\dagger = \theta(L - y)e^{i\pi/4} (\hat{\chi}_{L\uparrow}(y), i\hat{\chi}_{L\downarrow}(y), \hat{\chi}_{L\downarrow}(y), -i\hat{\chi}_{L\uparrow}(y)) \quad (2.158)$$

$$\chi_R^\dagger = \theta(y)e^{-i\pi/4} (\hat{\chi}_{L\uparrow}(L - y), -i\hat{\chi}_{L\downarrow}(L - y), \hat{\chi}_{L\downarrow}(L - y), i\hat{\chi}_{L\uparrow}(L - y)) . \quad (2.159)$$

In this approximation, inside of the wire for $0 < y < L$ both χ_L and χ_R are zero energy solutions, i.e. $\mathcal{H}\chi_L = 0 = \mathcal{H}\chi_R$ (where $\theta(y) = 1$ and $\theta(L - y) = 1$), and due to the boundary conditions, the wave functions vanish at the end they are computed for, i.e. $\chi_L(0) = 0$ and $\chi_R(L) = 0$. Using that the derivative of the step function is a delta distribution, $\partial_y \theta(y) = \delta(y)$, there is only a non-vanishing contribution to E_0 when the term $\tau_z \sigma_0 \partial_y^2$ of the Hamiltonian acts on the step function (see App. A.4 for a detailed calculation). This yields [368, 371, 372]

$$E_0 = 2 \left(\hat{\chi}_{L\uparrow}'(0) \hat{\chi}_{L\uparrow}(L) - \hat{\chi}_{L\downarrow}'(0) \hat{\chi}_{L\downarrow}(L) \right) = Re^{-L/\xi} \cos(k_{\text{eff}}L - \delta) , \quad (2.160)$$

where $\hat{\chi}_{L\sigma}'$ denotes the first derivative. The last equation is valid if the s-wave localization length ξ_2 and the p-wave localization length ξ fulfill $e^{-L/\xi_2} \ll e^{-L/\xi}$, which is already the case shortly after entering the topological region (see Fig. 2.7(b)). The exact analytical expressions for R and δ are given in App. A.4; however, they are not important to understand the qualitative behavior. The crucial point is that the energy of the Majorana state is exponentially suppressed by the wire length and oscillates with $\cos(k_{\text{eff}}L) \approx \cos(k_FL)$. In Fig. 2.11 we present the analytical solution together with the numerical results. In addition to the wire length dependence in Fig. 2.11(a), the analytical solution also provides the chemical potential dependence (Fig. 2.11(b)) and the Zeeman field dependence (Fig. 2.11(c)) – all in excellent agreement with the numerical solution. As expected, because the finite energy is exponentially suppressed by the wire length, for $L \gg \xi$ there is effectively a zero mode.

⁵ As the wave functions were obtained analytically for a semi-infinite wire (or it was assumed that the wave function decays fast enough to be zero at the opposite end), we here approximate the wave functions by cutting them off at the opposite wire ends. The comparison with the numerics shows that this is an excellent approximation for sufficiently long wires.

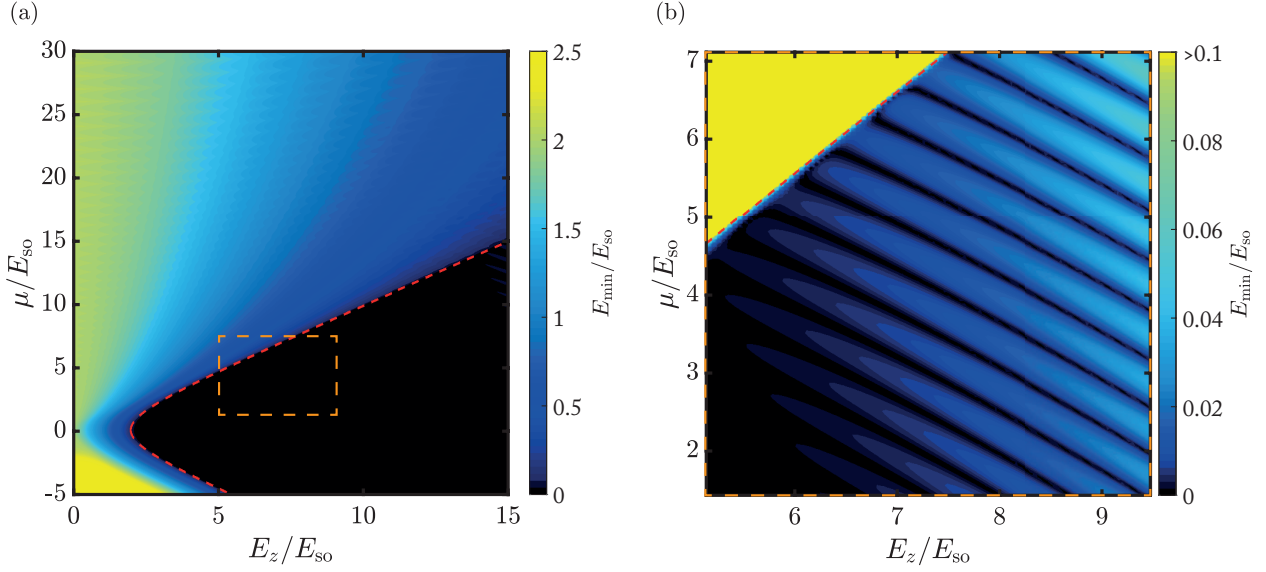


Figure 2.12: Energy of the lowest energy level as a function of the chemical potential μ and Zeeman energy E_z for a wire of length $L = 32.5 l_{so}$. The dashed red line marks the boundary between topological and trivial region. Panel (b) shows a zoom into the orange region marked in (a). We use a different colormap scale in (b) such that finite Majorana energy oscillations become visible.

The analytical wave function for the isolated Majorana states is independent of the wire length. Therefore, μ and E_z influence the energy E_0 indirectly via the correlation length $\xi(\mu, E_z)$. Here, ξ depends only weakly on μ , which is why E_0 increases only slightly with increasing μ (see Fig. 2.11(b)). As we saw in the last section (Fig. 2.7), ξ has a strong E_z dependence and is proportional to $\xi \propto E_z$ for large Zeeman fields, so that E_0 also increases with E_z .

As we have seen, zero energy solutions for separated Majorana states are independent of wire length, as the correlation length ξ does not depend on L . In the case of a finite overlap of the Majorana states from both ends, the wave functions hybridize in the topological region to eigenstates $\propto \chi_L \pm i\chi_R$ with finite energy $\propto e^{-L/\xi}$. While in the topological region, the spin \uparrow component is dominant, in the trivial region the contribution from the spin- \downarrow components have a finite overlap and thus give rise to an energy on the order of the induced gap.

In Fig. 2.12 we depict the energy of the lowest eigenstate as a function of both chemical potential μ and Zeeman energy E_z . In the topological region, we see the oscillations of finite energy values around zero due to the finite wire length (see Fig. 2.12(b)). In the trivial range the energy increases rapidly to the value of the induced gap $\Delta_{p\text{ind}}$, Eq. (2.93). Therefore, we can distinguish the topological from the trivial region only by means of the energy of the lowest level in this case. In the following section we introduce a smooth confinement potential, which reduces the overlap in the trivial region and thus enables trivial quasi-zero energy states to exist in the gap [112]. As a result, signatures based solely on the energy of the Majorana level cannot provide clear evidence for Majorana zero modes. Therefore, in later chapters we will consider signatures based on the localization property of Majorana zero modes and not their energy.

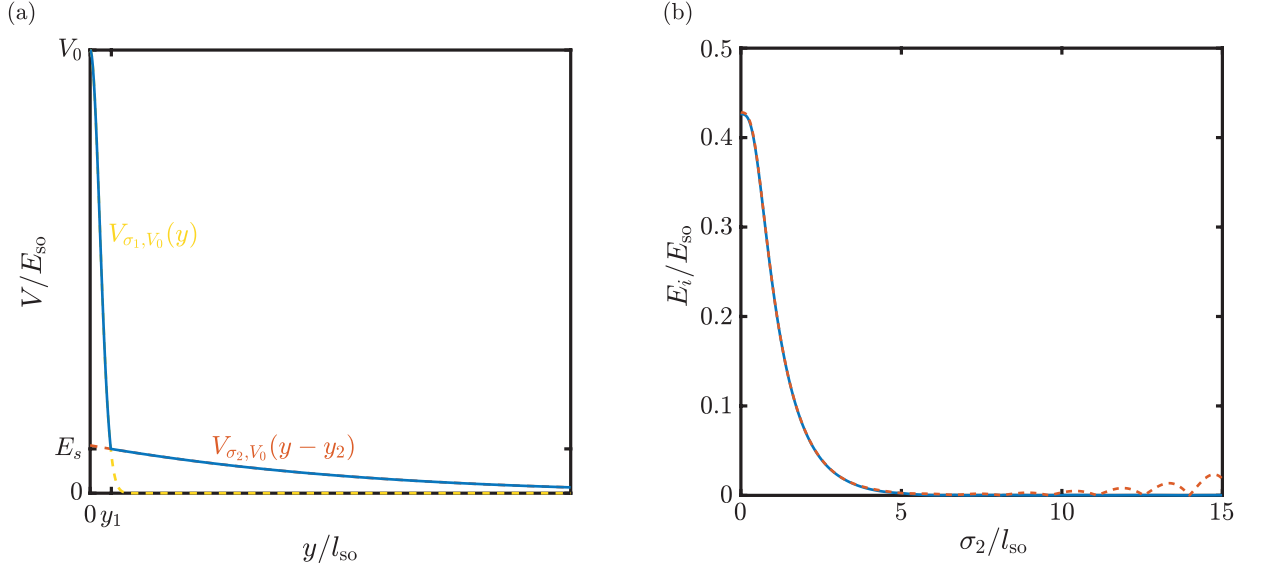


Figure 2.13: (a) Confinement potential at the left end of the wire (blue) according to Eq. (2.162). The solid line is the potential used in the Hamiltonian, which consists of the two Gaussian peaks represented by dashed lines. (b) The lowest two energy levels in the trivial region as a function of the width σ_2 of the wide peak. For the numerical calculations we use a confinement potential with $V_0 = 65 E_{\text{so}}$, $\sigma_1 = 0.1 l_{\text{so}}$, $E_s = 10 E_{\text{so}}$, Zeeman energy $E_z = 4.5 E_{\text{so}}$, chemical potential $\mu = 6 E_{\text{so}}$, induced superconductor gap $\Delta = 2 E_{\text{so}}$, lattice constant $a = 0.026 l_{\text{so}}$, and wire length $L = 39 l_{\text{so}}$.

2.2.4 Robust Andreev bound states

As we have seen in Sec. 2.2.2, the system is in the topological phase if – for a sufficiently large Zeeman field – only the lower band is occupied (see Fig. 2.4), so that the Majorana excitations are effectively spin polarized. On the other hand, in the trivial region both bands are occupied, and effectively two p-wave superconductors with opposite spin are present. Each of these p-wave superconductors contributes a Majorana mode at each end of the wire, differing in their Fermi momentum, $k_{F,-} < k_{F,+}$. If we employ hard wall boundary conditions for the ends of the wire, the step-like confinement potential has a Fourier expansion, which also contains components for large momenta $\sim |k_{F,+} - k_{F,-}|$, thus coupling each of the Majorana modes at the same end. This results in a large overlap (see Sec. 2.2.3) such that the lowest levels have energies close to the induced p-wave gap, $E_1 \approx E_2 \approx \Delta_{p,\text{ind}}$. If, on the other hand, the confinement potential is sufficiently smooth so that its Fourier expansion does not contain any contributions for such momenta, then the Majorana modes of the different effective p-wave superconductors cannot couple due to momentum conservation, and they remain at zero energy [111–128, 374–376]. A finite energy near zero is then only due to the overlap of the Majorana zero modes of opposite ends, i.e., of the same p-wave superconductor (analogous to Sec. 2.2.3). In addition, a potential influences the two spin directions to different extents due to their different momenta $k_{F,\pm}$ as a result of the Zeeman term, which means that the spin- \uparrow and spin- \downarrow Majorana wave functions at the same end are also spatially shifted against each other. This reduces their overlap and thus their energy even when they are coupled together. Thus, in the presence of a smooth boundary potential we expect the occurrence of pairs of quasi-zero energy states in the trivial phase [112], which are called Andreev zero modes or pseudo-Majorana states.

To take the confinement potential $V(y)$ into account, we consider the Hamiltonian

$$\mathcal{H} = \tau_z \otimes \left[-\sigma_0(\partial_{\tilde{y}}^2 + \tilde{\mu} - \tilde{V}(y)) - 2i\sigma_x \partial_{\tilde{y}} \right] - \tilde{E}_z \tau_0 \otimes \sigma_z + \tilde{\Delta} \tau_x \otimes \sigma_0 . \quad (2.161)$$

For a transport experiment, one connects the ends of the wire with leads so that electrons have to tunnel through the confinement potential to enter or leave the wire. Therefore, if we choose a high and wide potential, the wire is barely coupled to the leads. To still have a finite tunneling probability, we compose the confinement potential as a high narrow peak and a flat wide Gaussian peak, such that it is given by

$$\tilde{V}(y) = (V_L(y) + V_L(L - y))/E_{\text{so}} \quad (2.162)$$

with

$$V_L(y) = \begin{cases} V_{\sigma_1, V_0}(y) & \text{for } y < y_1 \\ V_{\sigma_2, V_0}(y + y_2 - y_1) & \text{for } y < y_1 \leq y \leq L \end{cases} , \quad (2.163)$$

where we use the abbreviations

$$V_{\sigma, V_0}(y) \equiv V_0 e^{-y^2/(2\sigma^2)} \quad (2.164)$$

$$y_j \equiv \sqrt{-2\sigma_j^2 \ln \frac{E_s}{V_0}} . \quad (2.165)$$

Here, the parameters of the potential are the standard deviation of the sharp Gaussian peak σ_1 , the standard deviation of the wide Gaussian peak σ_2 , and E_s is the height at which the narrow peak transitions into the wide one (see Fig. 2.13(a)).

We choose $V_0 = 65 E_{\text{so}}$ and $\sigma_1 = 0.1 l_{\text{so}}$ for the narrow high potential, and to investigate the influence of the broad peak on the energy eigenvalues of the Hamiltonian in the trivial phase, we numerically calculate the first two levels E_1, E_2 as a function of σ_2 for fixed $E_s = 10 E_{\text{so}}$ (see Fig. 2.13(b)). As expected, the first two Andreev bound states have almost vanishing energy for sufficiently wide potentials. The required width depends on how deep the system is in the trivial phase, as determined by μ and E_z . To better understand how robust these trivial pseudo-Majorana zero modes are under variation of chemical potential and Zeeman energy, we calculate the energy E_{min} of the first level as a function of E_z and μ numerically. To do this we consider two different potentials for comparison

- (a) a steep potential Eq. (2.162) with $\sigma_1 = \sigma_2 = 0.1 l_{\text{so}}$, $V_0 = 65 E_{\text{so}}$, $E_s = V_0$,
- (b) a smooth potential Eq. (2.162) with $\sigma_1 = 0.1 l_{\text{so}}$, $\sigma_2 = 6 l_{\text{so}}$, $V_0 = 65 E_{\text{so}}$, $E_s = 10 E_{\text{so}}$.

For the steep potential (a) (see Fig. 2.14(a)), we find energies similar to hard wall confinement (see Fig. 2.12), i.e., immediately after entering the trivial phase (indicated by the red dotted line), the energy increases to $\sim \Delta_{p,\text{ind}}$. Therefore, the topological phase can here be identified by the occurrence of the zero energy state. In contrast, if the potential is smooth (see Fig. 2.14(b)), pseudo-Majorana zero modes occur for extended parameter ranges in the trivial phase. How far these states extend into the trivial phase depends on the width of the confinement potential. In experiments, both the confinement and the effective chemical potential are controlled by electrostatic gating [130], such that a shallow confinement can often not be excluded, and the sole presence of signatures based on the zero-energy character is not sufficient to show that the system is in the topological phase.

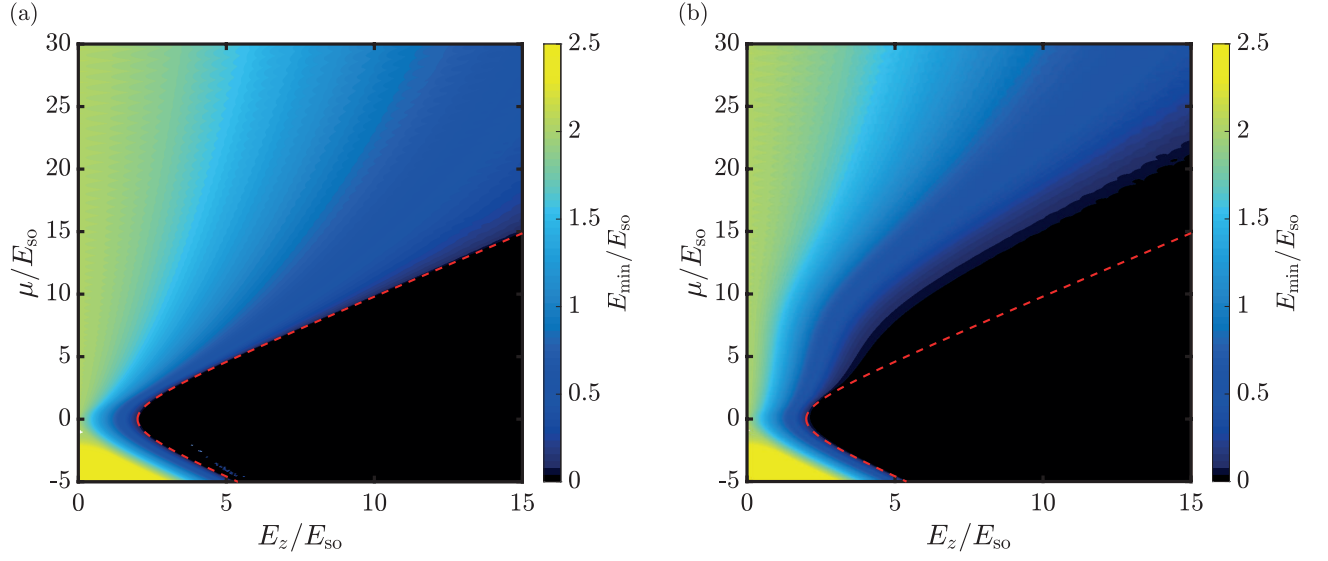


Figure 2.14: Energy of the lowest level as a function of chemical potential and Zeeman energy (cf. Ref. [124]) for $L = 39 l_{\text{so}}$ and confinement potentials according to Eq. (2.162) with (a) $\sigma_1 = \sigma_2 = 0.1 l_{\text{so}}$, $V_0 = 65 E_{\text{so}}$, $E_s = V_0$, and (b) $\sigma_1 = 0.1 l_{\text{so}}$, $\sigma_2 = 6 l_{\text{so}}$, $V_0 = 65 E_{\text{so}}$, $E_s = 10 E_{\text{so}}$. The dotted red lines mark the boundary between the topological and trivial phases.

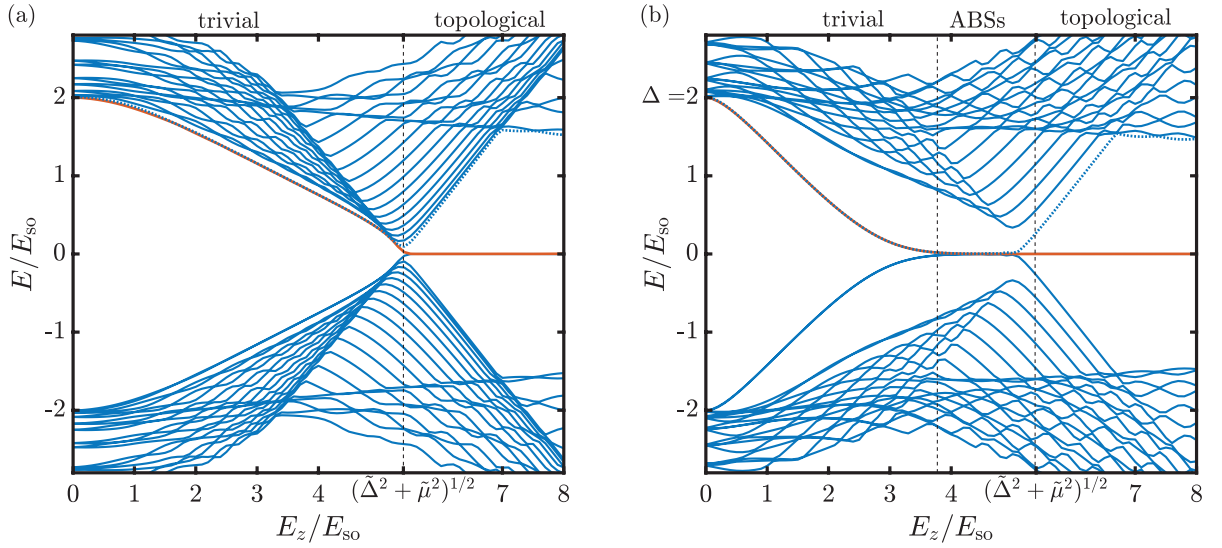


Figure 2.15: Lowest 20 energy eigenvalues of $\mathcal{H}/E_{\text{so}}$, Eq. (2.161), for $\mu = 5 E_{\text{so}}$, $\Delta = 2 E_{\text{so}}$, and $L = 32.5 l_{\text{so}}$ using (a) the steep and (b) the shallow confinement potential. The dotted, vertical lines are the boundaries between trivial case, pseudo Majorana bound states, and topological Majorana bound states. Here, E_2 is depicted with a dotted blue line and E_1 with a solid, red line.

In Fig. 2.15 we show the lowest few energy levels as a function of the Zeeman energy. In the case of the steep potential (a), we again find the same results as for the hard wall confinement (compare Fig. 2.15(a) with Fig. 2.12). Here the system is in the topological region for Zeeman energies $E_z > \sqrt{\mu^2 + \Delta^2}$ where the Majorana state is formed with zero energy, $E_1 \approx 0$, in the gap, while the second level is on the order of the induced gap, $E_2 \approx \Delta_{p,\text{ind}}$. On the other hand for the smooth confinement potential (b) there is, in addition to the topological domain for $E_z > \sqrt{\mu^2 + E_z^2}$ (which is similar to case (a)), another region for smaller Zeeman fields $E_z \lesssim \sqrt{\mu^2 + \Delta^2}$, in which the two pseudo Majorana zero modes with $E_1 \approx E_2 \approx 0$ are present. If the Zeeman energy is sufficiently large, a transition to the topological phase occurs, just like for hard wall confinement. We note that $V(x)$ appears in the Hamiltonian with the same Pauli matrices as the chemical potential, which together with the Zeeman field determines the boundary for the topological phase. Therefore, it is not surprising that the topological phase transition occurs slightly before $E_z = \sqrt{\mu^2 + E_z^2}$, as can be seen from the fact that E_2 increases before the dotted vertical line in Fig. 2.15(b).

In order to better understand the influence of the confinement potential on the overlap of the Majorana wave functions, we again depict the wave functions $|\psi_{\pm}^{(E_i)}|^2$, Eq. (2.149), and the spin expectation values $\langle S_z^{E_i} \rangle$, Eq. (2.150), for potentials (a) and (b) in Fig. 2.16.

We find in all cases that there are two contributions to the wave functions that oscillate with different frequencies. This is not surprising in the trivial region, since the Fermi level intersects both bands (see Fig. 2.4) at momenta $k_{F,\pm}$ with $k_{F,-} < k_{F,+}$. The slow oscillating component with $k_{F,+}$ is mainly spin- \downarrow polarized, while the fast oscillating $k_{F,-}$ component is predominantly spin- \uparrow (in agreement with Fig. 2.5). For the steep potential (a), both components have a large weight at the ends of the wire, where mainly opposite spin components have a large overlap (see Fig. 2.16(a)). Therefore, the Andreev bound states in case (a) have finite energies $E_1 = 0.479 E_{\text{so}}$ and $E_2 = 0.494 E_{\text{so}}$, which are on the order of the induced gap. In contrast, the smooth potential pushes the spin- \downarrow peaks away from the ends further into the wire (see Fig. 2.16(b)). This reduces the overlap and causes the contributions at the wide peak to vanish due to the rapidly oscillating tail of the spin- \uparrow component [cf. 370], i.e., no coupling between the components is possible, and the energies $E_1 \approx 1.41 \cdot 10^{-3} E_{\text{so}}$, $E_2 = 7.45 \cdot 10^{-3} E_{\text{so}}$ are several orders of magnitude smaller.

2.2.5 Two-dimensional Majorana wires

There are promising experimental results [91, 93, 130] for realizing MZMs in hybrid devices of two-dimensional electron gases (2DEGs) with spin orbit coupling in proximity to thin s-wave superconductors, e.g. an Al superconductor placed on top of a InAs 2DEG. In this section, we consider the case of such a two-dimensional strip of length L_y and width L_x , and study under which conditions we obtain localized Majorana modes at the short ends and how they relate to the one-dimensional description considered so far. For this we also reproduce some results found by Potter et al. [366, 377], who showed that for a strip thinner than the Majorana localization length $L_x < \xi_{\text{eff}}$, the approximation of a one-dimensional wire is justified. They argue that for a wide wire with $L_x \gg \xi_{\text{eff}}$, the Majorana states appear as edge states in the sample which exponentially decay into the bulk with a characteristic length of ξ_{eff} . If, however, L_x is decreased below this correlation length, the counter propagating edge states from the two sites overlap and mix and what remains are the localized Majorana modes at the two short ends separated by the distance $L_y \gg \xi_{\text{eff}}$ [377].

The two-dimensional case mainly differs from the previous one dimensional considerations by three effects: i) the appearance of transversal subbands, ii) an additional component of the Rashba spin-orbit coupling, and iii) the orbital effect due to the magnetic field. To separate

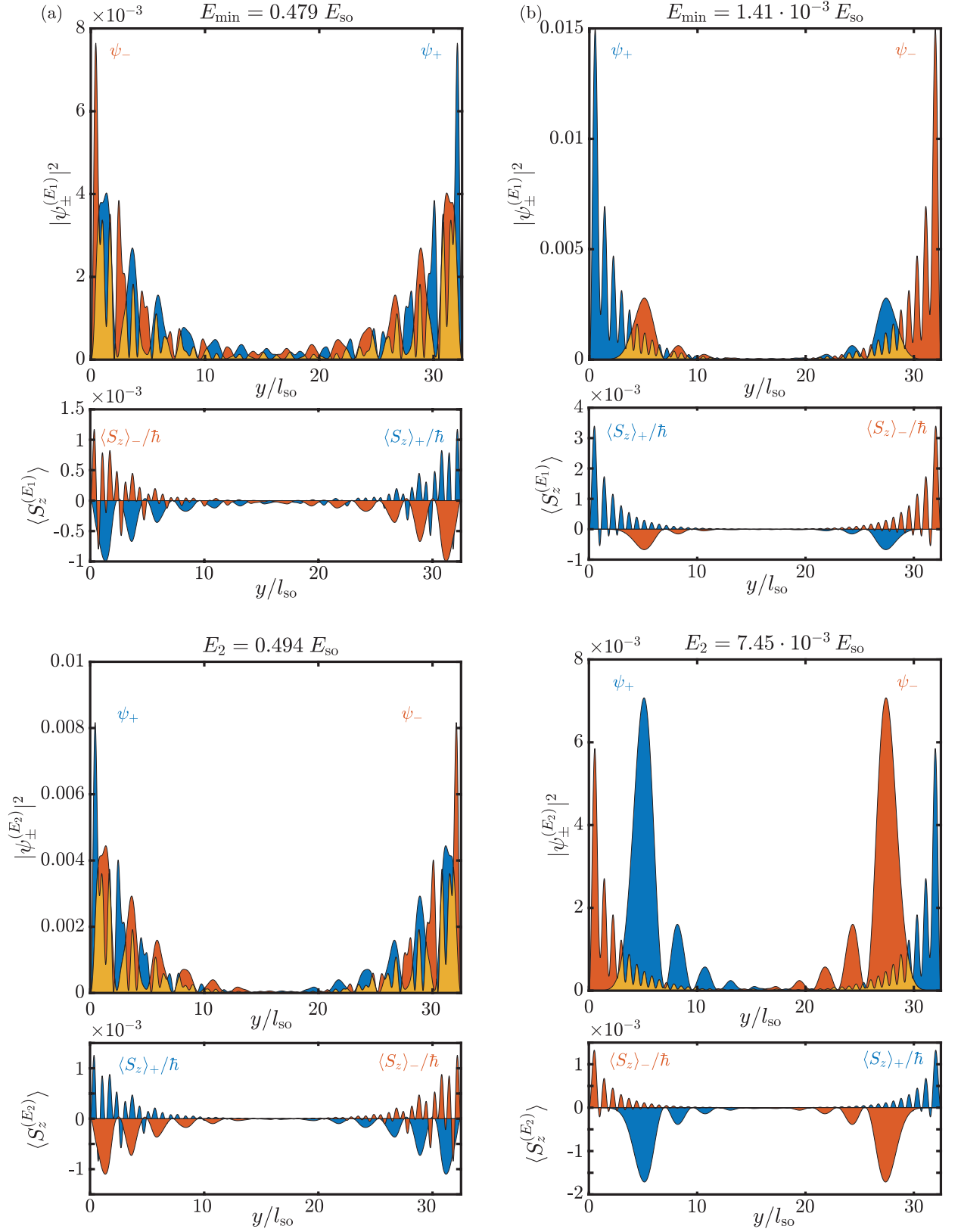


Figure 2.16: Wave functions $|\psi_{\pm}^{(E_i)}|^2$, Eq. (2.149), and spin expectation values $\langle S_z^{(E_i)} \rangle$ Eq. (2.150), in the trivial phase for the first two levels ($i = 1, 2$) using $E_z = 8 E_{\text{so}}$, $\mu = 10 E_{\text{so}}$, and (a) the steep confinement and (b) the shallow confinement. The overlap of the wave functions is depicted in orange. Similar results are shown in Ref. [370].

the influence of these factors, we turn them on one by one and consider the band structure and Majorana wave functions numerically.

Without the orbital effect of the magnetic field

We consider the Hamiltonian of a two-dimensional wire with induced s-wave superconductivity [378]

$$H_{2d} = \int_0^{L_x} dx \int_0^{L_y} dy \psi^\dagger(x, y) \mathcal{H}_{2d} \psi(x, y) \quad (2.166)$$

$$\mathcal{H}_{2d} = \tau_z \left[-\frac{\hbar^2}{2m^*} (\partial_y^2 + \partial_x^2) \sigma_0 - \mu \sigma_0 - i\hbar \alpha_R (\sigma_x \partial_y - \sigma_y \partial_x) \right] - E_z \tau_0 \sigma_z + \Delta \tau_x \sigma_0, \quad (2.167)$$

where the basis spinor is again given by $\psi^\dagger = (c_\uparrow^\dagger, c_\downarrow^\dagger, c_\downarrow, -c_\uparrow)$. We have used that the Rashba term in two dimensions is given by $-i(\boldsymbol{\sigma} \times \nabla) \cdot \mathbf{e}_z$, where for simplicity we assume that the spin-orbit coupling strength is the same in both directions. Here, the term $i\hbar \alpha_R \sigma_y \partial_x$ breaks the pseudo time reversal symmetry \mathcal{T} symmetry, while the anti-unitary reflection Π and particle hole symmetry \mathcal{P} remain intact. To determine numerical results, we also discretize the two-dimensional Hamiltonian \mathcal{H}_{2d} on a lattice with lattice constant $a = 0.026 l_{\text{so}}$ in both directions. The number of lattice sites for length L_x and width L_y is thus given as $4N_y N_x = 4L_x L_y / a^2$. As a basis we choose $\Psi^\dagger = (\psi_x^\dagger(y=a), \psi_x^\dagger(y=2a), \dots, \psi_x^\dagger(y=N_y a))$ with $\psi_x^\dagger(y) = (\psi^\dagger(x=a, y), \psi^\dagger(x=2a, y), \dots, \psi^\dagger(x=N_x a, y))$ and $\psi^\dagger(x, y) = (c_\uparrow^\dagger(x, y), c_\downarrow^\dagger(x, y), c_\downarrow(x, y), -c_\uparrow(x, y))$. Thus, operators can be expressed as Kronecker products $O_y \otimes O_x \otimes \tau \otimes \sigma$, where O_y acts in y -direction, O_x in x -direction, τ in the particle-hole basis, and σ in the spin basis. For example, the operator $c_\downarrow^\dagger(x, y)$ then has position $4(x/a - 1) + 4N_x(y/a - 1) + 2$ in the basis vector Ψ^\dagger . For the discretization, we therefore express

$$\partial_x \rightarrow \frac{1}{2a} \mathbb{1}_{N_y} \otimes (\ell_{N_x} - r_{N_x}) \quad (2.168)$$

$$\partial_y \rightarrow \frac{1}{2a} (\ell_{N_y} - r_{N_y}) \otimes \mathbb{1}_{N_x} \quad (2.169)$$

$$\partial_x^2 \rightarrow \frac{1}{a^2} \mathbb{1}_{N_y} \otimes (\ell_{N_x} + r_{N_x} - 2\mathbb{1}_{N_x}) \quad (2.170)$$

$$\partial_y^2 \rightarrow \frac{1}{a^2} (\ell_{N_y} + r_{N_y} - 2\mathbb{1}_{N_y}) \otimes \mathbb{1}_{N_x}, \quad (2.171)$$

analogous to the one-dimensional case. Here we define the $N \times N$ matrices $[\mathbb{1}_N]_{ij} = \delta_{ij}$, $[\ell_N]_{ij} = \delta_{i,j+1}$, and $[r_N]_{ij} = \delta_{i,j-1}$. In addition, we again use $E_{\text{so}} = m^* \alpha_R^2 / 2$ and $l_{\text{so}} = \hbar / (m^* \alpha_R)$ to separate parameters from their units, i.e. $\tilde{\mu} = \mu / E_{\text{so}}$, $\tilde{E}_z = E_z / E_{\text{so}}$, $\tilde{\Delta} = \Delta / E_{\text{so}}$, $\tilde{a} = a / l_{\text{so}}$.

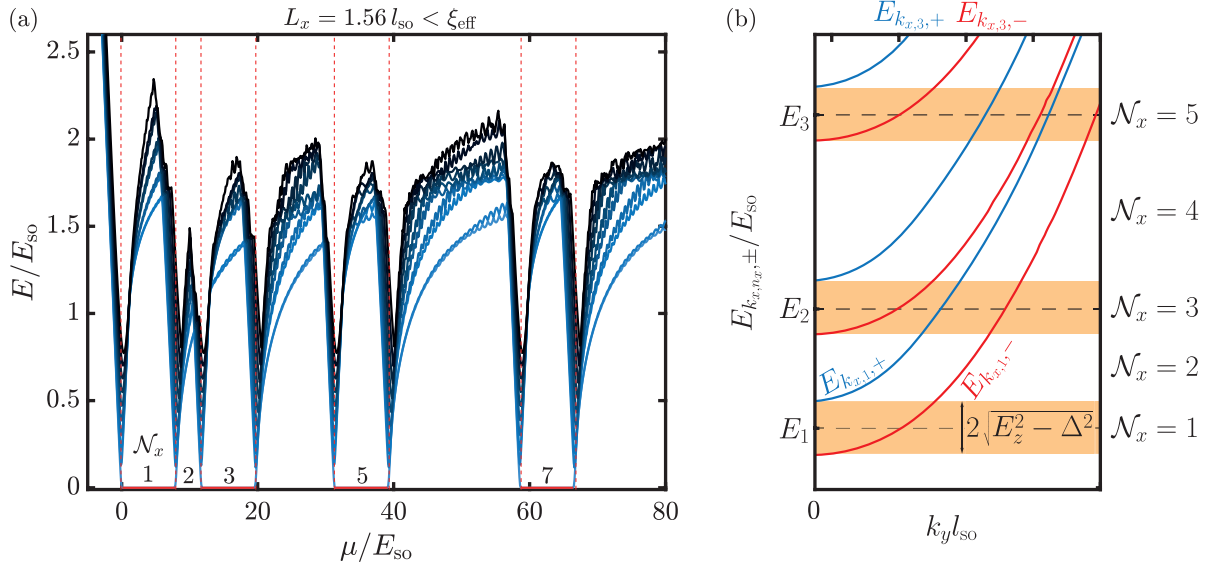


Figure 2.17: (a) The lowest ten energy eigenvalues of \mathcal{H}_{2d} without the $i\partial_x\tau_z\sigma_y$ term for $E_z = 7 E_{so}$, $\mu = 0$, $a = 0.026 l_{so}$, $\Delta = 2 E_{so}$, $L_y = 32.5 l_{so}$, $L_x = 1.56 l_{so}$. The red dashed lines mark the bottoms of the transverse bands at energies $E_{n,\pm}^{(x)} = (n\pi/(L_x/l_{so}))^2 \pm \sqrt{E_z^2 - \Delta^2}$, and \mathcal{N}_x is the number of occupied transverse bands [366]. (b) Dispersion $E_{\pm}(k_x, k_y)$ projected in the (E, k_y) plane, assuming $L_y \rightarrow \infty$ and $L_x < \xi_{eff}$, such that $k_{x,n_x} = \pi n_x/L$ is quantized and k_y is continuous. Here \mathcal{N}_x is the number of occupied bands when the chemical potential μ lies in the corresponding region. If μ lies in the orange highlighted regions, the number of occupied bands is odd and Majorana zero modes are present. The centers of these regions are given by $E_{n_x} = \hbar^2 k_{x,n_x}^2/(2m)$.

Thus, we obtain the discretized Hamiltonian as

$$H_{2d} = \Psi^\dagger \mathcal{H}_{2d} \Psi, \quad (2.172)$$

$$\begin{aligned} \frac{\mathcal{H}_{2d}}{E_{so}} = & -\frac{1}{\tilde{a}^2} [\mathbb{1}_{N_y} \otimes (r_{N_x} + \ell_{N_x} - 2\mathbb{1}_{N_x}) + (r_{N_y} + \ell_{N_y} - 2\mathbb{1}_{N_y}) \otimes \mathbb{1}_{N_x}] \otimes \tau_z \otimes \sigma_0 \\ & - \tilde{\mu} [\mathbb{1}_{N_y} \otimes \mathbb{1}_{N_x}] \otimes \tau_z \otimes \sigma_0 \\ & + \frac{i}{\tilde{a}} [\mathbb{1}_{N_y} \otimes (\ell_{N_x} - r_{N_x})] \otimes \tau_z \otimes \sigma_y \\ & - \frac{i}{\tilde{a}} [(\ell_{N_y} - r_{N_y}) \otimes \mathbb{1}_{N_x}] \otimes \tau_z \otimes \sigma_x \\ & - \tilde{E}_z [\mathbb{1}_{N_y} \otimes \mathbb{1}_{N_x}] \otimes \tau_0 \otimes \sigma_z \\ & + \tilde{\Delta} [\mathbb{1}_{N_y} \otimes \mathbb{1}_{N_x}] \otimes \tau_x \otimes \sigma_0. \end{aligned} \quad (2.173)$$

First, suppose the spin orbit coupling term $i\hbar\alpha_R\sigma_y\partial_x$ was absent. In this case we would have a Rashba wire in the y -direction and a free electron gas in an infinitely high potential well in the x -direction. The momentum in the x -direction is thus quantized to $k_{x,n_x} = \pi n_x/L_x$ with integer n_x . This results in transverse bands with n_x nodes, energies $E_{n_x} = (\pi l_{so}/(L_x + a))^2 E_{so} n_x^2$, and topological regions with Majorana zero modes for

$$E_{n_x} - \sqrt{E_z^2 - \Delta^2} < \mu < \sqrt{E_z^2 - \Delta^2} + E_{n_x}. \quad (2.174)$$

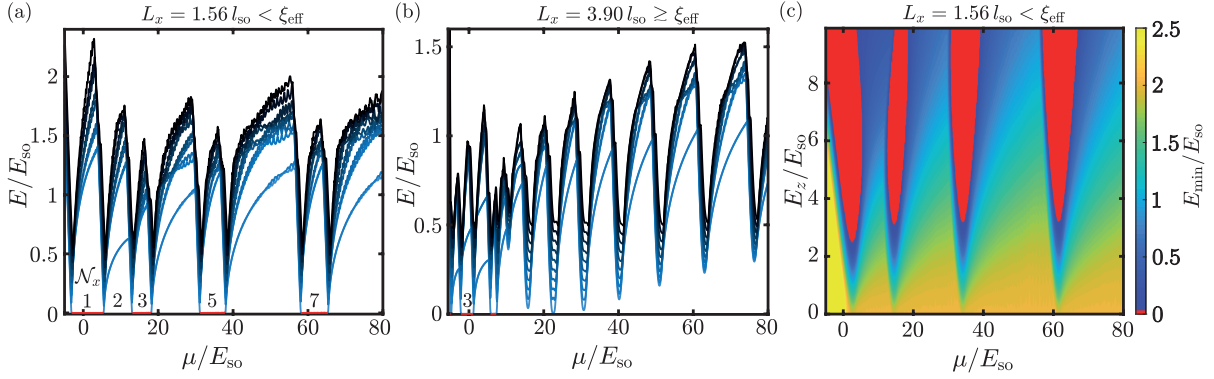


Figure 2.18: The lowest ten energy levels of the full Hamiltonian \mathcal{H}_{2d} for $E_z = 7 E_{so}$, $\Delta = 2 E_{so}$, $L_y = 32.5 l_{so}$, for (a) a thin wire $L_x = 1.56 l_{so} < \xi_{\text{eff}}$ [cf. 377] and (b) a wide wire with $L_x = 3.90 l_{so} \gtrsim \xi_{\text{eff}}$ as a function of the chemical potential μ , as well as for (c) the thin wire $L_x = 1.56 l_{so} < \xi_{\text{eff}}$ as a function of both the chemical potential μ and the Zeeman energy E_z . Similar results are shown in Ref. [377].

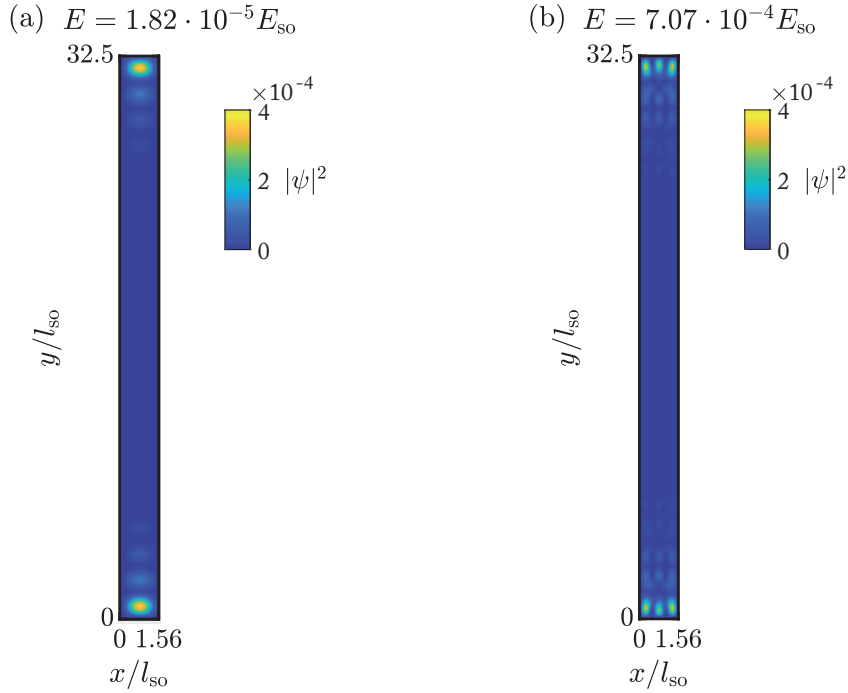


Figure 2.19: Wave function $|\psi|^2$ of the lowest energy level for $E_z = 7 E_{so}$, $\Delta = 2 E_{so}$, $L_x = 1.56 l_{so} < \xi_{\text{eff}}$, and $L_y = 32.5 l_{so} \gg \xi_{\text{eff}}$. In panel (a) for $\mathcal{N}_x = 1$ occupied subbands at $\mu = 0$, and panel (b) for $\mathcal{N}_x = 5$ occupied subbands at $\mu = 3^2(\pi l_{so}/(L_x + a))^2 E_{so}$.

The original topological region is thus shifted by an offset of $E_1/E_{\text{so}} = (\pi l_{\text{so}}/(L_x + a))^2$ from the origin and additional Majorana zero modes occur around $\mu = E_{n_x}$ for $n_x \in \mathbb{N}$. This is shown in Fig. 2.17(a), where we depict the ten lowest energy levels as a function of the chemical potential. These energy levels are obtained numerically for a Hamiltonian Eq. (2.173) without the $i\sigma_y\partial_x$ spin-orbit coupling term. As parameters we choose $a = 0.026 l_{\text{so}}$, $E_z = 4.5 E_{\text{so}}$, $\Delta = 2 E_{\text{so}}$, $L_y = 32.5 l_{\text{so}}$, $L_x = 1.56 l_{\text{so}} < \xi_{\text{eff}}$, and indicate a quasi-zero energy Majorana state by a solid red line. The red dotted vertical lines mark the positions $E_n \pm \sqrt{E_z^2 - \Delta^2}$ where the gap closes, and \mathcal{N}_x is the number of transverse bands crossed by the Fermi level for $\Delta = 0$ [377]. This is sketched in Fig. 2.17(b) where we show the dispersion with several transverse subbands (k_x) for a wire infinitely extended in y -direction as a function of the longitudinal momentum k_y . We note that for the parameters we have chosen, the lower bands for each k_x also have only a single minimum (cf. Fig. 2.5(d)). As expected, a Majorana zero energy state occurs only if the number of crossed bands \mathcal{N}_x is odd, and for an even number \mathcal{N}_x there are two low energy states with nearly degenerate energy that is significantly larger than zero. If we now consider larger widths L_x , the transverse subbands E_{n_x} move closer together. If a band $E_{-,k_x,n+1}$ then falls below $E_{+,k_x,n}$ at $k_y = 0$, the areas with odd \mathcal{N}_x become smaller.

Now that we have discussed the principles in the simple case without spin-orbit coupling in y -direction, we consider the full Hamiltonian. We will find that the numerical results can still be qualitatively explained with the simple picture with a single spin-orbit coupling term $-i\sigma_x\partial_y$ and refer to Refs. [366, 377] for a more detailed discussion. In Fig. 2.18(a) and (b) we again show the ten lowest energy eigenvalues of the full Hamiltonian as a function of the chemical potential μ . Here we choose $E_z = 7 E_{\text{so}}$, $\Delta = 2 E_{\text{so}}$, $a = 0.026 l_{\text{so}}$, and $L_y = 32.5 l_{\text{so}} \gg \xi_{\text{eff}}$. In Fig. 2.18(a) the width of the wire is given by $L_x = 1.56 l_{\text{so}} < \xi_{\text{eff}}$. We find the same pattern of repeating topological regions near $E_{n_x}/E_{\text{so}} = (\pi l_{\text{so}}/(L_x + a))^2 n_x^2$. However, the bands are slightly shifted due to the extra spin-orbit coupling term. Again, for the regions with Majorana zero modes an odd number \mathcal{N}_x of transverse subbands is occupied. In addition, we show the energy of the lowest state as a function of μ and E_z in Fig. 2.18(c). In the yellow regions this energy is of the order of the induced gap, $E_{\text{min}} \approx \Delta$, and we indicate regions with Majorana zero modes with $E_{\text{min}} \approx 0$ in red. For small Zeeman energies $E_z \lesssim 2 E_{\text{so}}$, the wire is always trivial. When increasing the Zeeman field, the topological regions are initially parabolic and resemble copies of the one-dimensional phase diagram for the different subbands. However, for large Zeeman fields, $E_z \gtrsim 6 E_{\text{so}}$, successive subbands cross each other such that the topological regions (red) are distorted, and their size can even decrease. In Fig. 2.18(b) we consider a wide wire with $L_x = 3.9 l_{\text{so}} \gtrsim \xi_{\text{eff}}$. As described above, in this case the topological regions are reduced due to subband crossing. In addition, the simplified model of the potential well in the x -direction breaks down for $L_x > \xi_{\text{eff}}$. As a consequence, the topological zero energy region for higher chemical potentials can even disappear entirely.

We further consider the Majorana wave functions: In the quasi one-dimensional case $L_x < \xi_{\text{eff}}$, the wave functions are localized at the short edges, as shown in Fig. 2.19(a) for $L_x = 1.56 E_{\text{so}}$, $L_y = 32.5 l_{\text{so}}$, $E_z = 7 E_{\text{so}}$, $\mu = 0$, $\Delta = 2 E_{\text{so}}$. This is the zero energy state with $\mathcal{N}_x = 1$ close to $E_1/E_{\text{so}} = (\pi l_{\text{so}}/(L_x + a))^2 - 1/2$. We observe a single node in the transverse direction with high weight, exponentially localized with localization length $\xi_{\text{eff}} > L_x$ at each end of the wire. In Fig. 2.19(b), we show the topological phase with $\mathcal{N}_x = 5$ for $\mu \approx (\pi l_{\text{so}}/(L_x + a))^2 3^2 E_{\text{so}}$, where the Majorana wave function has $\lfloor \mathcal{N}_x/2 \rfloor + 1 = 3$ nodes in the transverse direction and is exponentially localized at the short edges.

Fig. 2.20 depicts the first occurring zero energy state (μ is tuned to the topological phase) for $E_z = 7 E_{\text{so}}$, $\Delta = 2 E_{\text{so}}$, $L_y = 32.5 l_{\text{so}}$, and different widths L_x of the wire. For this set of

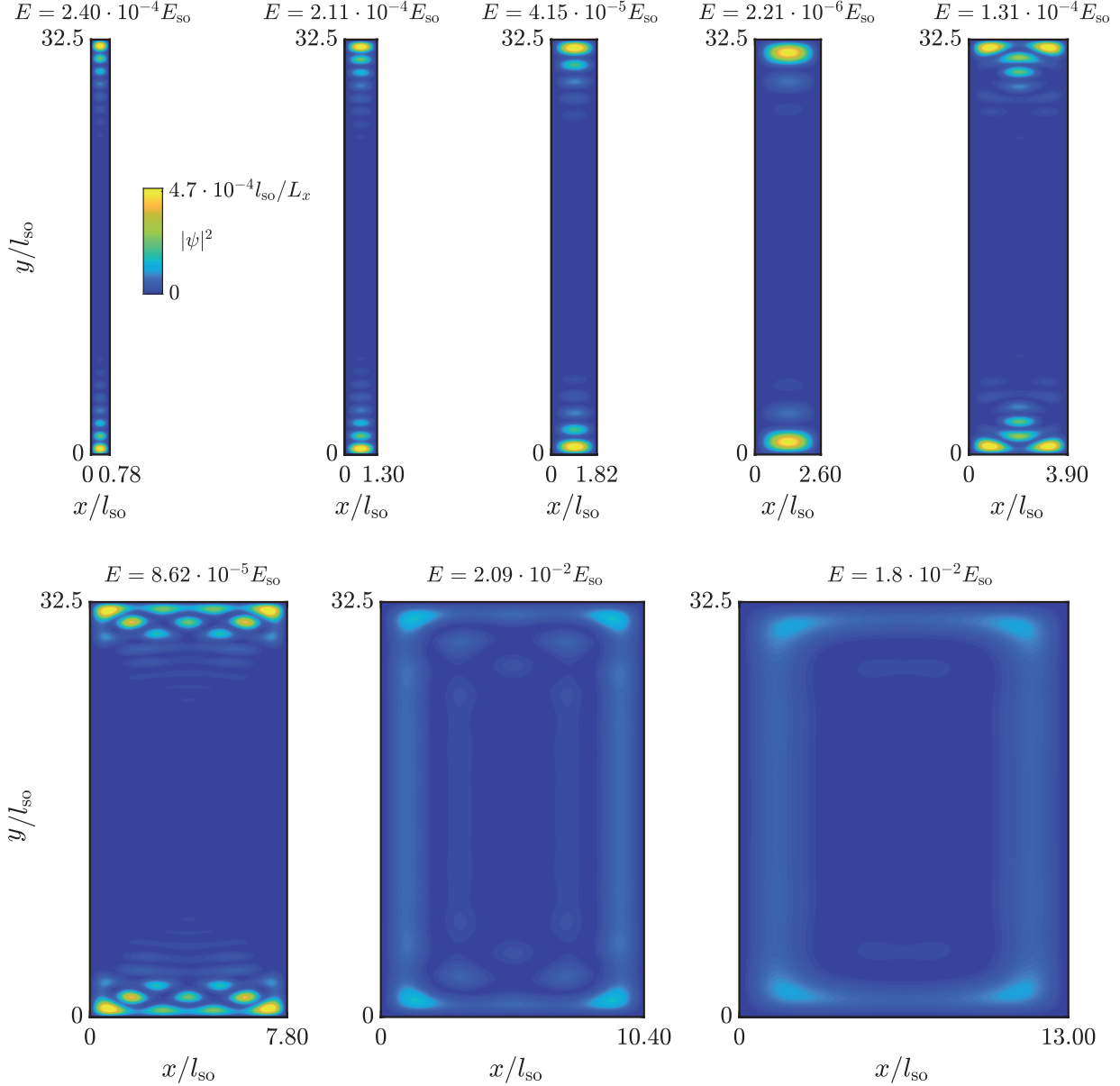


Figure 2.20: Wave function $|\psi|^2$ of the zero energy states for different widths L_x and $E_z = 7 E_{\text{so}}$, $\Delta = 2 E_{\text{so}}$, $L_y = 32.5 l_{\text{so}}$, and $\mu = \pi^2 l_{\text{so}}^2 / (L_x + a)^2 E_{\text{so}}$, so that $\mathcal{N}_x = 1$ subband is occupied. We choose L_x as $\{0.78, 1.30, 1.82, 2.60, 3.90, 7.80, 19.50\} l_{\text{so}}$, and the localization length of the MZMs is roughly $\xi_{\text{eff}} \approx 3.52 l_{\text{so}}$ according to the analytic solution for the one-dimensional wire. Similar results are shown in Ref. [377].

parameters, the analytical solution for the one-dimensional wire Eq. (2.133) yields a localization length of $\xi_{\text{eff}} \approx 3.52 l_{\text{so}}$, and it can be clearly seen that for $L_x < \xi_{\text{eff}}$ there are two Majoranas at the short ends of the wire, while for $L_x \gg \xi_{\text{eff}}$ the zero-energy wave functions are localized along the edges of the true two-dimensional region. For the former quasi-one-dimensional case, we also compare the one-dimensional analytical solution $\psi_{1d}(y)$ with the two-dimensional solution $\mathcal{N}\psi_{2d}(x = L_x/2, y)$, where the factor \mathcal{N} accounts for the fact that the two-dimensional solution is normalized on a larger region. Fig. 2.21 shows the magnitude of these wave functions for

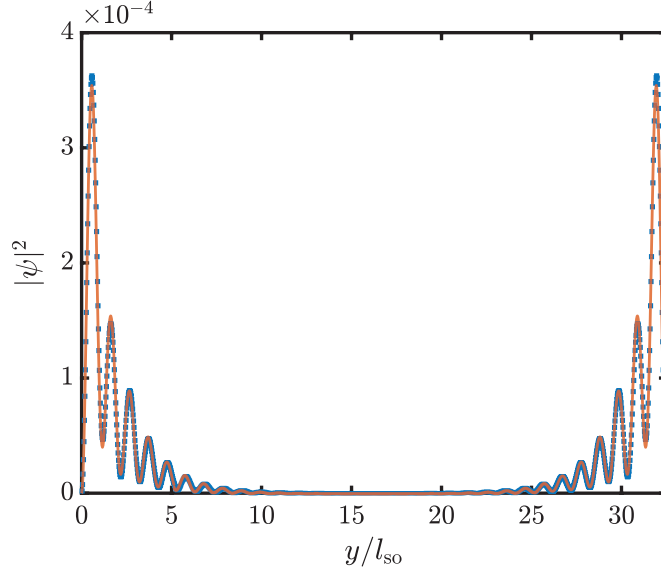


Figure 2.21: Comparison of the analytical solution of the one-dimensional wave function $\mathcal{N}|\psi(y)|^2$ with the cross-section of the two-dimensional wave function $|\psi_{2d}(x = L_x/2, y)|^2$. Here \mathcal{N} is a rescaling factor, since $|\psi_{2d}|^2$ is normalized on the domain $(0, L_x) \times (0, L_y)$. For both solutions $L_y = 32.5 l_{so}$, $E_z = 7 E_{so}$, and $\Delta = 2 E_{so}$. For the numerical two-dimensional solution, we consider a thin wire with $L_x = 1.56 l_{so} < \xi_{eff}$. For the two-dimensional case, a chemical potential of $\mu_{2d} = [(\frac{\pi l_{so}}{L_x + a})^2 - \frac{1}{2}] E_{so}$ corresponds to the case $\mu_{1d} = 0$ of the one-dimensional, analytical solution because of the shift due to the transverse confinement.

$E_z = 7 E_{so}$, $\Delta = 2 E_{so}$, $L_y = 32.5 l_{so}$, and a chemical potential of $\mu_{1d} = 0$ for the one-dimensional solution, which corresponds to $\mu_{2d} = (\pi l_{so}/(L_x + a))^2 - 1/2$ in the two-dimensional case with $L_x = 1.56 l_{so}$. We find excellent correspondence between the one-dimensional solution and the cross-section of the two-dimensional solution by a suitable choice of the constant pre-factor. Hence, the one-dimensional solution provides an excellent estimate of the first topological regime of the more realistic two-dimensional wire.

The orbital effect of the magnetic field

Up to this point, we have considered only the influence of the magnetic field on the electron spin via the Zeeman effect, which contributes the term $H_Z = -\frac{1}{2}g\mu_B \mathbf{B} \cdot \boldsymbol{\sigma} \tau_0$ to the Hamiltonian [379]. Here μ_B is the Bohr magneton, g the Landé factor (we use $g = -14.9$ for InAs [380, 381]), and we consider a magnetic field $\mathbf{B} = -B_z \mathbf{e}_z$ orthogonal to the plane of the two-dimensional hybrid system such that $H_Z = \frac{1}{2}g\mu_B B_z \sigma_z \tau_0 \equiv -E_z \tau_0 \sigma_z$. Furthermore, the magnetic field also affects the trajectory (classically described by the Lorentz force), and the phase of the wave function of an electron with charge $-e$, which we account for using the minimal substitution $\mathbf{p} \rightarrow -i\hbar\nabla + e\mathbf{A}\tau_z$ in the Hamiltonian [379], where \mathbf{A} is the vector potential with $\mathbf{B} = \nabla \times \mathbf{A}$ and the Pauli matrix τ_z allows describing particles and holes in a single expression [382]. In the presence of the magnetic field, we thus obtain the Hamiltonian

$$H_{2d} = \int_0^{L_x} dx \int_0^{L_y} dy \psi^\dagger(x, y) \left\{ \tau_z \left[\frac{1}{2m^*} (-i\hbar\nabla + e\mathbf{A}\tau_z)^2 \sigma_0 - \mu \sigma_0 \alpha_R (\boldsymbol{\sigma} \times (-i\hbar\nabla + e\mathbf{A}\tau_z)) \cdot \mathbf{e}_z \right] - E_z \tau_0 \sigma_z + \Delta \tau_x \sigma_0 \right\} \psi(x, y), \quad (2.175)$$

where the basis is given by $\psi^\dagger(x, y) = (\psi_\uparrow^\dagger(x, y), \psi_\downarrow^\dagger(x, y), \psi_\downarrow(x, y), -\psi_\uparrow(x, y))$. Due to the orbital effect where $\mathbf{A}(x, y)$ depends on the position in the wire, direct discretization according to the rules of the previous section no longer leads to a physically correct Hamiltonian. This has the two main reasons that i) \mathbf{A} is also mixed with the derivatives ∂_i and thus with the hopping elements, but these are in general not symmetric if one defines \mathbf{A} naively on the lattice points, resulting in a non-Hermitian Hamiltonian, and ii) in addition, the wave function has to pick up a phase factor $e^{i\pi\phi/\phi_0}$ with $\phi_0 = h/2e$ if an electron encircles a surface penetrated by a flux ϕ [383], which cannot be guaranteed by naive discretization. To circumvent these problems, we resort to a trick and first define new fermion field operators [383, 384]

$$\tilde{\psi}(\mathbf{r}) = e^{\frac{ie}{\hbar}\tau_z \int_{\mathbf{R}}^{\mathbf{r}} \mathbf{A}(\tilde{\mathbf{r}}) \cdot d\tilde{\mathbf{r}}} \psi(\mathbf{r}) , \quad (2.176)$$

where \mathbf{R} is some reference point on the lattice. Here, the phase factor is chosen such that the gradient cancels the vector potential⁶

$$e^{\frac{ie}{\hbar}\tau_z \int_{\mathbf{R}}^{\mathbf{r}} \mathbf{A}(\tilde{\mathbf{r}}) \cdot d\tilde{\mathbf{r}}} (-i\hbar\nabla + e\mathbf{A}\tau_z) e^{-\frac{ie}{\hbar}\tau_z \int_{\mathbf{R}}^{\mathbf{r}} \mathbf{A}(\tilde{\mathbf{r}}) \cdot d\tilde{\mathbf{r}}} = -i\hbar\nabla . \quad (2.177)$$

When using Eq. (2.176) in the Hamiltonian Eq. (2.175) only the superconducting pairing term is special, as it contains two creation and two annihilation operators, respectively. We therefore define

$$\Delta(\mathbf{r})\tau_x\sigma_0 \equiv \Delta e^{-\frac{ie}{\hbar}\tau_z \int_{\mathbf{R}}^{\mathbf{r}} \mathbf{A}(\tilde{\mathbf{r}}) \cdot d\tilde{\mathbf{r}}} e^{-\frac{ie}{\hbar}\tau_z \int_{\mathbf{R}}^{\mathbf{r}} \mathbf{A}(\tilde{\mathbf{r}}) \cdot d\tilde{\mathbf{r}}} \tau_x\sigma_0 \quad (2.178)$$

which is fortunately a pure on-site term and thus does not cause any problems in the discretization. Using these relations, the Hamiltonian becomes

$$H_{2d} = \int_0^{L_x} dx \int_0^{L_y} dy \tilde{\psi}^\dagger(x, y) \left\{ \tau_z \left[\frac{-\hbar^2}{2m^*} \nabla^2 \sigma_0 - \mu \sigma_0 - i\hbar\alpha_R (\boldsymbol{\sigma} \times \nabla) \cdot \mathbf{e}_z \right] - E_z \tau_0 \sigma_z + \Delta(x, y) \tau_x \sigma_0 \right\} \tilde{\psi}(x, y) , \quad (2.179)$$

which is the same form we found in the case without orbital effect, and is discretized on the square lattice with lattice constant a by using

$$\partial_i^2 \tilde{\psi}(\mathbf{r}) \rightarrow \frac{1}{a^2} [\tilde{\psi}(\mathbf{r} + a\mathbf{e}_i) + \tilde{\psi}(\mathbf{r} - a\mathbf{e}_i) - 2\tilde{\psi}(\mathbf{r})] \quad (2.180)$$

$$\partial_i \tilde{\psi}(\mathbf{r}) \rightarrow \frac{1}{2a} [\tilde{\psi}(\mathbf{r} + a\mathbf{e}_i) - \tilde{\psi}(\mathbf{r} - a\mathbf{e}_i)] . \quad (2.181)$$

Only after performing these approximations in the Hamiltonian, we change back to the original operators using

$$\tilde{\psi}_\sigma^\dagger(\mathbf{r} + a\mathbf{e}_i) \tilde{\psi}_{\sigma'}(\mathbf{r}) = e^{-\frac{ie}{\hbar}\tau_z \int_{\mathbf{r}}^{\mathbf{r}+a\mathbf{e}_i} \mathbf{A}(\tilde{\mathbf{r}}) \cdot d\tilde{\mathbf{r}}} \psi_\sigma^\dagger(\mathbf{r} + a\mathbf{e}_i) \psi_{\sigma'}(\mathbf{r}) \quad (2.182)$$

$$\tilde{\psi}_\sigma^\dagger(\mathbf{r}) \tilde{\psi}_{\sigma'}(\mathbf{r}) = \psi_\sigma^\dagger(\mathbf{r}) \psi_{\sigma'}(\mathbf{r}) \quad (2.183)$$

$$\tilde{\psi}^\dagger(\mathbf{r}) \Delta(\mathbf{r}) \tau_x \sigma_0 \tilde{\psi}(\mathbf{r}) = \psi^\dagger(\mathbf{r}) \Delta \tau_x \sigma_0 \psi(\mathbf{r}) . \quad (2.184)$$

⁶This result is only valid if the integral is path independent. We can still use it for discretizing the Hamiltonian, since we only consider nearest neighbor terms for which the path between the two sites is unique on the lattice.

Thus, the discretized Hamiltonian contains the same on-site terms as for the case without the orbital effect, and also the hoppings are only adjusted by a spatially dependent Peierls phase [384, 385], i.e., a hopping $t_{\mathbf{r}_1, \mathbf{r}_2}$ between lattice sites \mathbf{r}_1 and \mathbf{r}_2 acquires a phase

$$\varphi(\mathbf{r}_1, \mathbf{r}_2) = -\frac{e}{\hbar}\tau_z \int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{A}(\tilde{\mathbf{r}}) \cdot d\tilde{\mathbf{r}} \quad (2.185)$$

$$t_{\mathbf{r}_1, \mathbf{r}_2} \rightarrow t_{\mathbf{r}_1, \mathbf{r}_2} e^{i\varphi(\mathbf{r}_1, \mathbf{r}_2)} = t_{\mathbf{r}_1, \mathbf{r}_2} [\tau_0 \cos(\varphi(\mathbf{r}_1, \mathbf{r}_2)) - i\tau_z \sin(\varphi(\mathbf{r}_1, \mathbf{r}_2))] . \quad (2.186)$$

Here, the relation $\int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{A}(\tilde{\mathbf{r}}) \cdot d\tilde{\mathbf{r}} = -\int_{\mathbf{r}_2}^{\mathbf{r}_1} \mathbf{A}(\tilde{\mathbf{r}}) \cdot d\tilde{\mathbf{r}}$ ensures that hoppings are symmetric such that the Hamiltonian is Hermitian, and when hopping along a closed loop \mathcal{C} that is the boundary of a surface $F(\mathcal{C})$ penetrated by a flux ϕ , Stokes theorem $\int_{\mathcal{C}} \mathbf{A}(\mathbf{r}) \cdot d\mathbf{r} = \int_{F(\mathcal{C})} [\nabla \times \mathbf{A}(\mathbf{r})] \cdot d^2\mathbf{r} = \phi$ makes sure that the correct phase is picked up.

Having discussed how to incorporate a vector potential into the discretized Hamiltonian, the next step is to derive a suitable expression of the vector potential for the magnetic field orthogonal to the wire plane. As we have already seen, the gauge freedom $\mathbf{A} \rightarrow \mathbf{A} + \nabla\chi$ does not change the magnetic field $\mathbf{B} = \nabla \times \mathbf{A} = \nabla \times (\mathbf{A} + \nabla\chi)$, but transforms the operators via a phase factor $\psi \rightarrow \psi e^{-\frac{ie}{\hbar}\chi}$. Therefore, the Hamiltonian is invariant under the gauge transformation, and the pairing term is transformed according to $\Delta e^{i\theta(\mathbf{r})} \rightarrow \Delta e^{i(\theta(\mathbf{r}) - \frac{2e}{\hbar}\chi)}$, such that the vector potential \mathbf{A} and the phase of the superconducting order parameter θ are not independent. To obtain a physically correct choice for \mathbf{A} and θ corresponding to the magnetic field $\mathbf{B} = -B_z \mathbf{e}_z$, the following conditions have to be satisfied [386]:

- i) The vector potential needs to describe the magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$.
- ii) The energy $\propto \int d^2r \mathbf{j}^2$ of the supercurrent

$$\mathbf{j} = \frac{\hbar \varrho_S}{m^*} \left(\nabla\theta + \frac{e}{\hbar} \mathbf{A} \right) \quad (2.187)$$

needs to be minimized [265, 386, 387], where ϱ_S is the density of Cooper pairs.

- iii) The supercurrent needs to be conserved $\nabla \cdot \mathbf{j} = 0$.

For an infinitely long Majorana wire in the x - y plane $\{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq L_x\}$ this can be achieved by choosing [381, 386]

$$\mathbf{A}(x, y) = -B_z x \mathbf{e}_y \quad (2.188)$$

$$\theta(x, y) = 0, \quad (2.189)$$

which clearly satisfies i). With this choice, the vector potential on the symmetry axis vanishes, i.e. $\mathbf{A}(x=0, y) = 0$, and therefore $\int dx x^2$ is minimal over the wire, and in addition \mathbf{A} is independent of the coordinate y extending to infinity in the wire. Thus, this vector potential also meets the requirement ii), and iii) is trivially satisfied in the infinite system as \mathbf{e}_y is parallel to the infinite wire direction and $\nabla \cdot (x \mathbf{e}_y) = 0$. In the case of a one-dimensional wire along the y axis, \mathbf{j} is zero everywhere in the wire by using Eq. (2.188) and Eq. (2.189) which explains why we can choose a real superconducting order parameter and why the orbital effect is not visible in the one-dimensional case.

As there cannot be any MZMs in the infinite wire due to the lack of short ends, we consider a finite wire with length $L_y \gg L_x$ next, i.e. $\{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq L_x \text{ and } 0 \leq y \leq L_y\}$. Unfortunately, as the supercurrent is zero outside the wire, the choice Eq. (2.188) and Eq. (2.189)

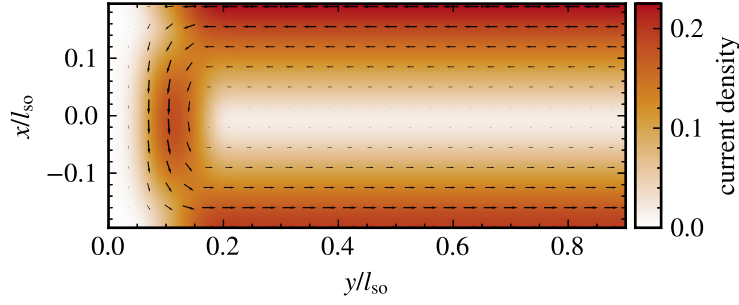


Figure 2.22: Supercurrent \mathbf{j} , Eq. (2.187), and current density $|\mathbf{j}|^2$ at the left end of a two-dimensional wire in an orthogonal magnetic field using the gauge Eq. (2.190) for the vector potential \mathbf{A} , and the superconducting phase $\theta(x, y) = 0$. (adapted from [303]).

does not conserve the current in the finite wire. In principle one would have to perform resource intensive self-consistent numerical computations to determine θ and \mathbf{A} such that conditions i)-iii) are met [381, 386], which would currently not be computationally feasible with the resources available to us. Instead, we use an approximation for the vector potential that reproduces the infinite wire case in the bulk and makes it smoothly vanish over a distance λ at the ends [303]

$$\mathbf{A} = -a(y)B_z x \mathbf{e}_y + \frac{a'(y)}{2} B_z (x^2 - (L_x/2)^2) \mathbf{e}_x, \quad (2.190)$$

$$a(y) = \begin{cases} f_{0,\lambda}(y) & 0 \leq y < \lambda \\ 1 & \lambda \leq y \leq L_y - \lambda \\ 1 - f_{L_y-\lambda, L_y}(y) & L_y - \lambda < y \leq L_y \end{cases}, \quad (2.191)$$

$$f_{y_1, y_2}(y) = \frac{h(y - y_1)}{h(y - y_1) + h(y_2 - y)} \quad (2.192)$$

$$h(y) = \begin{cases} \exp(-\lambda/y) & y > 0 \\ 0 & y \leq 0 \end{cases}, \quad (2.193)$$

where we choose $\lambda = L_x/2 \ll \xi_{\text{eff}} \ll L_y$. Together with $\theta(x, y) = 0$ this ensures that the current smoothly vanishes at the short wire ends such that $\nabla \cdot \mathbf{j}$ is satisfied everywhere in the wire. The resulting smooth current \mathbf{j} is depicted in Fig. 2.22.

As the last part of this section, we use the derived rules for the two-dimensional system to numerically compute the energy levels and compare locations of topological regions for the case with and without the orbital effect of the magnetic field. In Fig. 2.23 we show the dispersions for length $L_x = 13 l_{\text{so}}$, and Landé factor $g = -14.9$. The topological regions are clearly visible by the vanishing of the lowest energy level (a and b) accompanied by closing of the gap at the phase transitions, apparent from the energy of the second level in panel (c). For small E_z , the position of the topological phases remains unchanged even when taking into account the orbital effect, but for larger Zeeman fields the topological regions are bent towards larger chemical potentials [382]. Starting at a point in the topological phase and increasing the Zeeman field E_z can therefore bring the system back into the trivial phase, which cannot be understood from the one dimensional model. In the presence of the orbital effect, it also becomes apparent that zero-energy crossings occur more frequently outside the topological phases, which means that it is no longer possible to localize the phases on the basis of the energy of the lowest level alone. The closing of the gap in the bulk at the topological phase transitions can be used for this

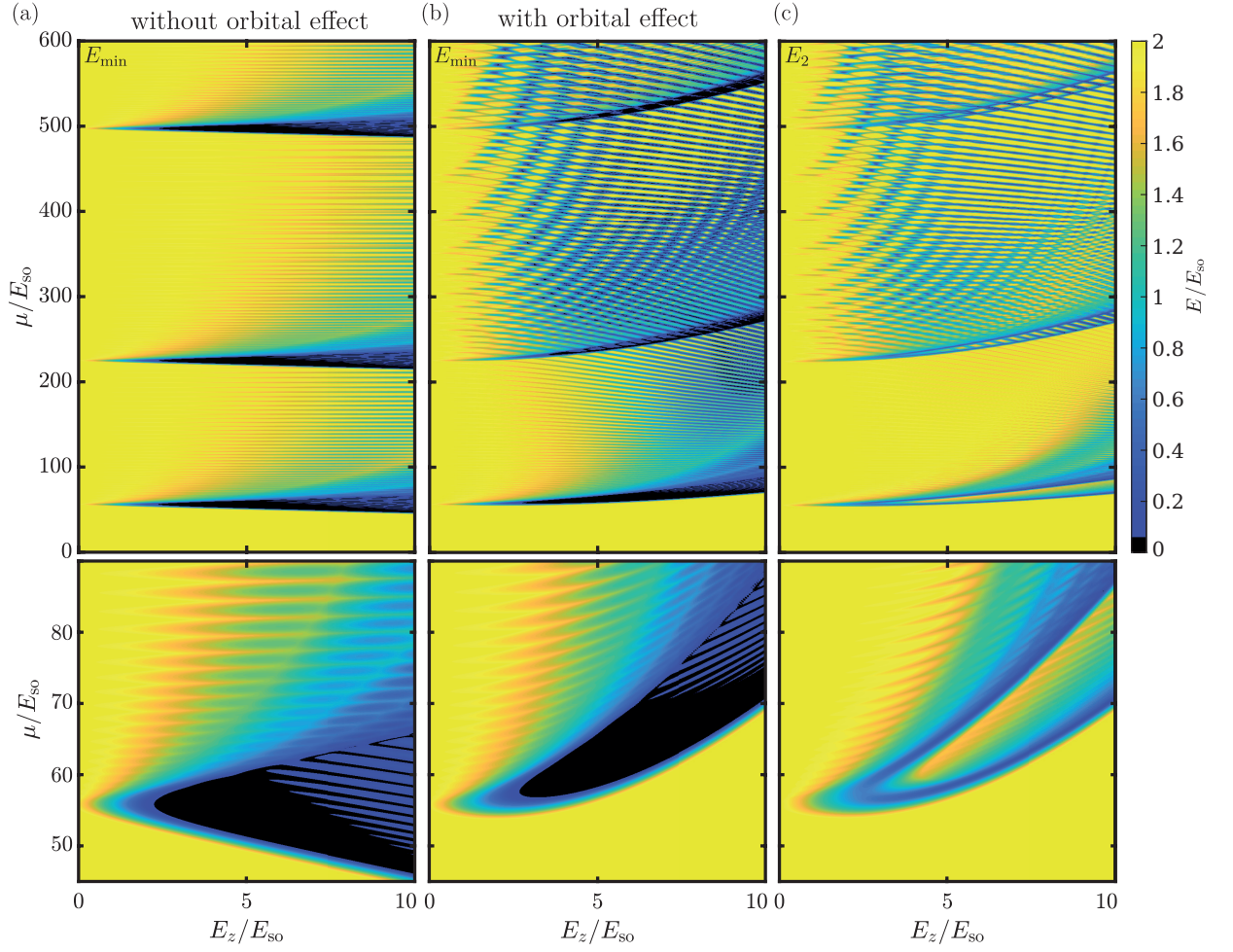


Figure 2.23: Energy of the lowest energy eigenvalues of the Hamiltonian as a function of the chemical potential μ and the Zeeman energy E_z for a thin wire with width $L_y = 0.39 l_{so}$, length $L_x = 13 l_{so}$, superconducting gap $\Delta = 2 E_{so}$, and a Landé factor of $g = -14.9$. (a) Energy of the first level without taking into account the orbital effect. The other panels depict results for the (b) first and (c) second level taking into account the orbital effect. The bottom panels show a zoom into the region of the first topological phase. Similar results are shown in [382].

purpose, which is still clearly visible from the energy of the second level in panel (c).

Fig. 2.24 further depicts MZM wave functions for the cases of (a) $\mathcal{N} = 1$, (b) $\mathcal{N} = 3$, and (c) $\mathcal{N} = 5$ occupied subbands at the Fermi level, which behave very similarly to the case without orbital effect discussed in the previous section. They can be identified to originate from the n -th topological phase based on the n nodes of the wave function.

In summary, we conclude that the orbital effect complicates the phase diagram, but the rough location of the topological domains is preserved [382] and for a thin wire with $L_x < \xi_{eff}$, the first topological regime can be approximated by the one-dimensional model for Zeeman energies close to the phase transition [366].

So far, we have demonstrated that the presence of a zero energy state is not a proof of the topological phase. However, we have further seen that the wave functions of the Andreev zero modes are also located at the ends of the wire. Thus, not only is it impossible to detect the presence of Majorana zero modes by energy-based signatures such as zero-bias peaks, but local

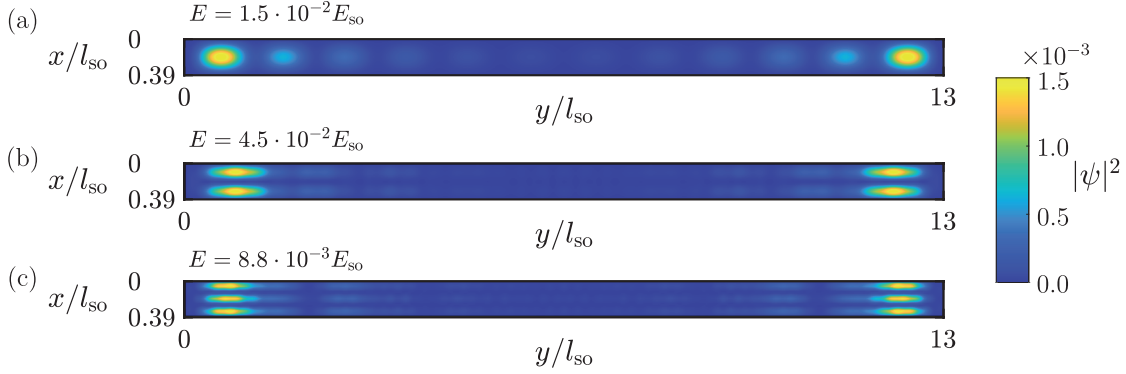


Figure 2.24: Majorana wave functions $|\psi(x, y)|^2$ of the lowest level form different topological phases (as shown in Fig. 2.23b) for $E_z = 7 E_{\text{so}}$, $L_x = 13 l_{\text{so}}$, $L_y = 0.39 l_{\text{so}}$, $g = -14.9$, and $\Delta = 2 E_{\text{so}}$. (a) In the first topological phase with $\mathcal{N} = 1$ subbands at the Fermi level for $\mu = 67 E_{\text{so}}$, (b) in the second topological phase, $\mathcal{N} = 3$, for $\mu = 250 E_{\text{so}}$, and (c) for the third one, $\mathcal{N} = 5$, with $\mu = 530 E_{\text{so}}$.

measurements cannot distinguish Majorana zero modes from pseudo-Majorana zero modes. This is a major problem for the experimental verification of Majorana zero modes [111–128, 376]. To be able to detect them successfully, we later consider a setup to measure coherent transport through the Majorana modes, to take advantage of the fact that the Majorana level is clearly separated from the next higher level by the induced gap, while the pseudo-Majorana zero modes appear in pairs, and their contributions can potentially destructively interfere with each other [132]. Furthermore, to exclude other processes such as Andreev reflections and tunneling of Cooper pairs, it will be advantageous to consider a non-grounded wire in the Coulomb blockade regime. Such a coherent transport experiment through a Rashba wire in Coulomb blockade was first theoretically described by Fu in 2008 [291]. However, this description is limited to the Majorana states only, and does not consider higher levels or Andreev zero modes.

After we have explored the properties of the Rashba hybrid wire in detail in this chapter, our goal is to develop a theoretical description of a coherent transport experiment in Coulomb blockade. We start in the next chapter by introducing a scattering matrix formalism that allows us to consider transmission phases and amplitudes for transport through e.g. a Rashba wire.

2.3 Scattering matrix formalism for coherent transport in Coulomb-blockade

In this section we describe the scattering matrix formalism used to compute coherent transport through Majorana wires. In order to study a signature based on the non-locality of MZMs that also allows distinguishing MZMs from Andreev zero modes and from other zero modes induced by disorder [131, 132], we consider Majorana wires embedded into one arm of an electron interferometer. By measuring interference with electrons through the reference arm one can extract the coherent transmission amplitude [130], which allows us in Sec. 2.4 to explain results of a recent experiment [130] and to propose a unique signature for MZMs by studying wires of different lengths [131]. Furthermore, in Sec. 2.5, we consider optimizing voltages of an array of gates, in proximity to a Majorana wire, using the CMA-ES machine learning algorithm, with the transmission amplitude as a metric for maximization [303].

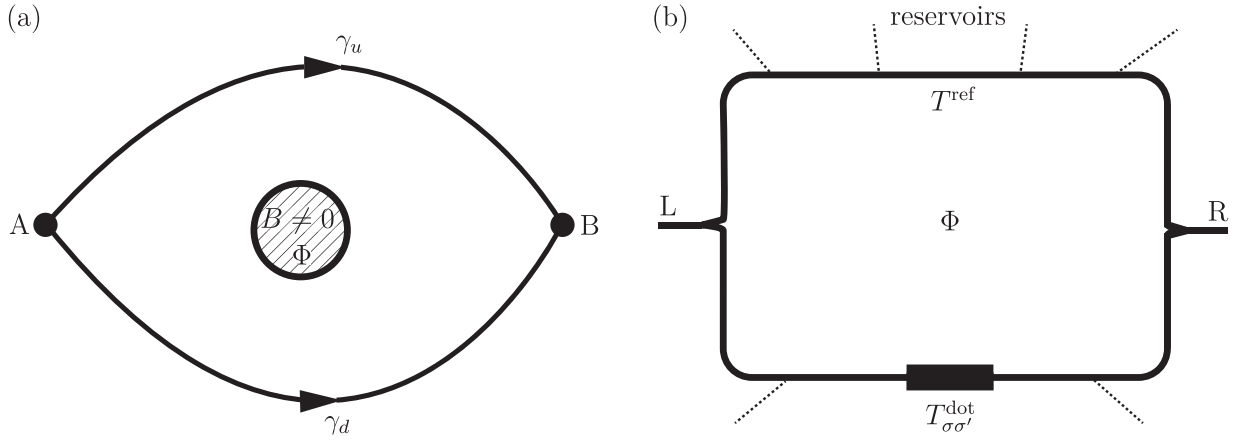


Figure 2.25: Using the Aharonov-Bohm effect for building an electron interferometer. (a) We consider points A and B connected by paths γ_u and γ_d . Inside the loop formed by both paths there is a flux Φ due to a cylinder with a magnetic field \mathbf{B} inside that cannot escape the cylinder. Even though there is no field on the paths, electrons traveling from A to B along different paths acquire a phase difference $\Delta\varphi = -e\Phi/\hbar$ [17, 388]. (b) Electron interferometer with two leads L and R, empty upper reference arm, lower arm with Majorana wire, and additional coupling to reservoirs to avoid the phase rigidity effect [390].

We begin here with a description of the Aharonov-Bohm interferometer [17, 388] and how the transmission amplitude can be experimentally determined in such a setup in Sec. 2.3.1. We then consider one of the most important ingredients for our considerations in Sec. 2.3.2, the Coulomb blockade, which makes it energetically unfavorable for additional electrons to enter the wire, allowing transport through the Majorana wire to be understood as co-tunneling through effective one-particle levels and enabling *teleportation* through MZMs [129]. In the following subsections, we provide details on the finite temperature scattering matrix formalism (Sec. 2.3.3), for which we use models with both effective and microscopic couplings (Sec. 2.3.4). In the last section, Sec. 2.3.5, we describe how to obtain the parameters for the scattering matrix formula using the open source python package KWANT [389], which simplifies studying systems with complex geometries as discretized tight-binding models, but does not support Coulomb blockade.

2.3.1 Aharonov-Bohm interferometer setup

As we have already discussed in Sec. 2.2.5, apart from the Zeeman effect, a magnetic field \mathbf{B} enters the Schrödinger equation by the minimal substitution $\mathbf{p} \rightarrow \mathbf{p} + e\mathbf{A}$, where e is the electron charge, \mathbf{p} is the kinetic momentum operator, and \mathbf{A} is the vector potential with $\mathbf{B} = \nabla \times \mathbf{A}$. We have seen that this results in a path dependent geometric phase [383]

$$\varphi(\gamma) = \frac{-e}{\hbar} \int_{\gamma} d\mathbf{r} \cdot \mathbf{A} , \quad (2.194)$$

which is collected by the wave function as the electron moves along the path γ .

We now consider two points A and B connected by two different paths γ_u and γ_d , between which an impenetrable cylinder containing a magnetic field \mathbf{B} is placed (Fig. 2.25a) [17]. As the field cannot escape the cylinder, there is no field on the paths, i.e. $0 = \mathbf{B}(\gamma_u) = \mathbf{B}(\gamma_d)$, however the closed path $\mathcal{C} = \gamma_u - \gamma_d$ encloses a flux $\Phi = \int_{\text{cylinder}} d^2\mathbf{r} \cdot \mathbf{B}$ due to the magnetic

field inside the cylinder. This results in a phase difference collected on the two paths between A and B given by

$$\Delta\varphi = \varphi(\gamma_u) - \varphi(\gamma_d) = \frac{-e}{\hbar} \int_{\mathcal{C}} d\mathbf{r} \cdot \mathbf{A} \quad (2.195)$$

$$= \frac{-e}{\hbar} \int_{F(\mathcal{C})} d^2\mathbf{r} \cdot (\nabla \times \mathbf{A}) \quad (2.196)$$

$$= -\frac{e}{\hbar} \Phi, \quad (2.197)$$

where we used Stokes theorem to express the integral over the path \mathcal{C} as an integral over the surface $F(\mathcal{C})$ with boundary \mathcal{C} . The occurrence of the phase difference – even though there is no magnetic field present on the paths of the electrons – is called the Aharonov-Bohm effect [17, 388], which can be exploited to build electron interferometers.

We consider in the following the setup of an Aharonov-Bohm interferometer shown in Fig. 2.25b, where the upper reference arm contains only a normal wire, and a short Majorana wire is embedded into the lower arm. This allows the transmission amplitude and transmission phase of coherent transport through the Majorana wire to be determined by measuring the oscillations of the current through the interferometer between A and B as a function of the flux Φ , as we will show below. Here, the Majorana wire is effectively treated as a quantum dot, where we determine the couplings of the dot to the leads and the energy levels of the dot from the eigenstates and energy eigenvalues of the full Majorana wire Hamiltonian (see also Sec. 2.3.3).

As depicted in Fig. 2.25b, it is also crucial to open the interferometer by coupling it to additional reservoirs to circumvent the phase-regularity effect [390]: In a closed interferometer preserving the particle number, the Onsager relation $G(\Phi) = G(-\Phi)$ [391, 392] holds for the flux dependence of the conductance, which does not allow for a nontrivial transmission phase β , since a term like $\cos(\phi + \beta)$ can only satisfy the Onsager relation for $\beta = 0, \pi$. Opening the interferometer has another advantage: while in a closed interferometer it is likely that an electron entering at lead L will go around the loop several times before leaving the interferometer at lead R, in an open interferometer it is unlikely that coherent electrons from L that can interfere at R have not taken direct routes via the reference arm or the arm with the dot, since at each opening there is a probability of leaving the interferometer [393–395]. As a result, the transmission amplitude T^{RL} can be approximated by the double slit limit [393]

$$T^{\text{RL}} = T_1 e^{i\Delta\varphi} + T_2, \quad (2.198)$$

where T_1 (T_2) is the transmission amplitude through the upper (lower) arm, and where we can attach the total phase difference $\Delta\varphi$ from the Aharonov-Bohm effect to one of the amplitudes.

For our setup, the situation is slightly more complicated because spin is not a good quantum number due to spin-orbit coupling in the Majorana wire, and thus the transmission matrix through the dot

$$\mathbf{T}^{\text{dot}} = \begin{pmatrix} T_{\uparrow\uparrow} & T_{\uparrow\downarrow} \\ T_{\downarrow\uparrow} & T_{\downarrow\downarrow} \end{pmatrix}, \quad (2.199)$$

also contains non-diagonal entries. We assume that the reference arm is diagonal in the spins, so that

$$\mathbf{T}^{\text{ref}} = \begin{pmatrix} T^{\text{ref}} & 0 \\ 0 & T^{\text{ref}} \end{pmatrix}, \quad (2.200)$$

where $T^{\text{ref}} \in \mathbb{R}$. Thus, the transmission amplitude through the interferometer is given by

$$T_{\sigma\sigma'}^{LR} = T_{\sigma\sigma'} + e^{i\Delta\varphi} \delta_{\sigma\sigma'} T^{\text{ref}} , \quad (2.201)$$

and using the Landauer formula [396], the current through the interferometer is obtained as

$$I = \frac{e^2}{h} \sum_{\sigma\sigma'} |T_{\sigma\sigma'}^{LR}|^2 = \frac{e^2}{h} \left\{ \sum_{\sigma\sigma'} |T_{\sigma\sigma'}|^2 + 2|T^{\text{ref}}|^2 + 2 \sum_{\sigma} \text{Re} \left[e^{i\Delta\varphi} T_{\sigma\sigma} T^{\text{ref}} \right] \right\} . \quad (2.202)$$

The current consists of the direct terms through the dot $\sum_{\sigma\sigma'} |T_{\sigma\sigma'}|^2$ and the reference arm $2|T^{\text{ref}}|^2$, and the interference current

$$I_{\text{intf}} \equiv \frac{2e^2}{h} \sum_{\sigma} \text{Re} \left[e^{i\Delta\varphi} T_{\sigma\sigma} T^{\text{ref}} \right] \quad (2.203)$$

$$= \frac{e^2}{h} |T^{\text{ref}}| \left[e^{i\Delta\varphi} (T_{\uparrow\uparrow} + T_{\downarrow\downarrow}) + e^{-i\Delta\varphi} (T_{\uparrow\uparrow} + T_{\downarrow\downarrow})^* \right] , \quad (2.204)$$

which oscillates as a function of the Aharonov-Bohm flux ϕ . Using $T_{\uparrow\uparrow} + T_{\downarrow\downarrow} = |T_{\uparrow\uparrow} + T_{\downarrow\downarrow}| e^{i\beta}$, where β denotes the phase of the complex number $T_{\uparrow\uparrow} + T_{\downarrow\downarrow}$, we find

$$I_{\text{intf}} = \frac{2e^2}{h} |T^{\text{ref}}| |T_{\uparrow\uparrow} + T_{\downarrow\downarrow}| \cos(\Delta\varphi + \beta) . \quad (2.205)$$

Here we denote $|T_{\uparrow\uparrow} + T_{\downarrow\downarrow}|$ as the amplitude of coherent transmission, or transmission amplitude for short, and β as the transmission phase. In an experiment, the transmission amplitude can be obtained from the amplitude of the current oscillations ΔI [130, 132] as

$$\Delta I \equiv \max_{\Delta\varphi} I - \min_{\Delta\varphi} I = \frac{4e^2}{h} |T^{\text{ref}}| |T_{\uparrow\uparrow} + T_{\downarrow\downarrow}| \quad (2.206)$$

$$|T_{\uparrow\uparrow} + T_{\downarrow\downarrow}| = \frac{h}{4e^2} \frac{\Delta I}{|T^{\text{ref}}|} . \quad (2.207)$$

By measuring an oscillation period of the current fluctuations $I_{\text{int}}(\Delta\varphi)$ and then observing the phase shift when varying other parameters of the dot, changes in β can also be obtained [304, 390].

Later, we consider finite temperatures and thermally average the current over occupations of the dot. Performing the average in Eq. (2.202), we find [303]

$$\langle I_{\text{intf}} \rangle = \frac{e^2}{h} |T^{\text{ref}}| \left[e^{i\Delta\varphi} \langle T_{\uparrow\uparrow} + T_{\downarrow\downarrow} \rangle + e^{-i\Delta\varphi} \langle (T_{\uparrow\uparrow} + T_{\downarrow\downarrow})^* \rangle \right] \quad (2.208)$$

$$\equiv \frac{2e^2}{h} |T^{\text{ref}}| |\langle T_{\uparrow\uparrow} + T_{\downarrow\downarrow} \rangle| \cos(\Delta\varphi + \gamma) , \quad (2.209)$$

where we defined $\langle T_{\uparrow\uparrow} + T_{\downarrow\downarrow} \rangle = |\langle T_{\uparrow\uparrow} + T_{\downarrow\downarrow} \rangle| e^{i\gamma}$. We thus denote $A = |\langle T_{\uparrow\uparrow} + T_{\downarrow\downarrow} \rangle|$ as the (finite temperature) transmission amplitude and γ as the transmission phase. This thermal average can be understood as dephasing due to averaging the current over a finite measurement time. It is crucial that the magnitude is only taken after averaging the amplitude $\langle T_{\uparrow\uparrow} + T_{\downarrow\downarrow} \rangle$ in the definition of A .

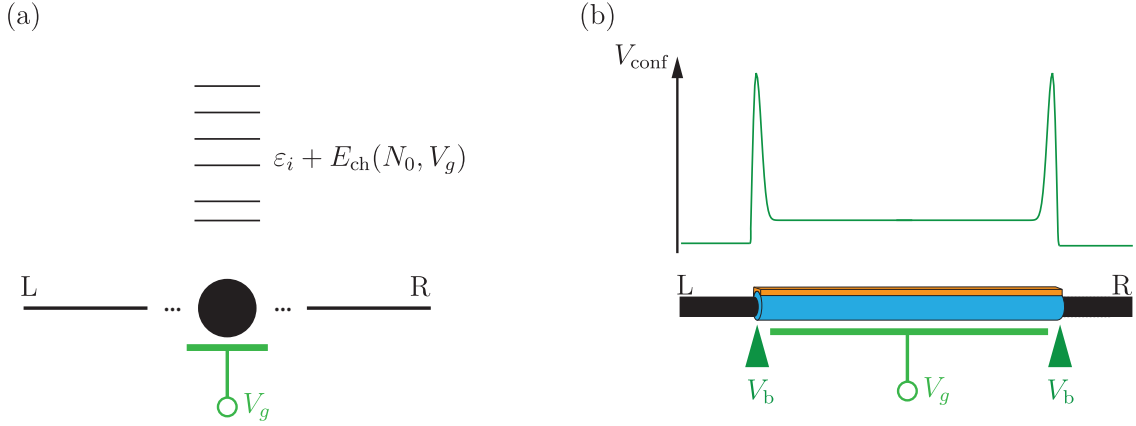


Figure 2.26: Setups for Coulomb-blockade in quantum dots. (a) A quantum dot containing N_0 electrons with energy levels ϵ_i and charging energy E_{ch} in proximity to a gate with voltage V_g and weakly coupled to two leads L and R. (b) Majorana wire hybrid wire made of a semiconductor with strong spin orbit coupling (blue) and an s-wave superconductor (orange) in proximity to a gate V_g . The wire is separated from the two leads L, R by pinch-off gates V_b which create a steep confinement potential V_{conf} .

2.3.2 Coulomb-blockade

In this section, we study Coulomb blockade which is caused by the repulsion of electrons when they are confined to a small region, e.g. a quantum dot, that couples only weakly to the environment [397, 398]. If a gate is placed in proximity to the dot, the energy of the electrons can be lowered so that, depending on the gate voltage, a fixed number of particles N_0 in the dot is energetically optimal. To then bring another electron to the island, or remove one, requires a charging energy. If one couples the region very weakly to two leads and applies a voltage between the leads, no electrons can tunnel sequentially as long as the voltage difference is less than the charging energy, which is why it is called Coulomb blockade. We use that due to the charging energy, tunneling of multiple particles at once can be effectively excluded and the remaining transport can be described by so-called cotunneling, which takes place via effective single-particle levels. We will see that this transport is enhanced by MZMs especially by their (non-local) localization properties.

However, before considering the more complicated situation of a Majorana wire with superconductivity, we study a quantum dot with energy levels ϵ_i with N_0 electrons (Fig. 2.26a) first. We assume that the level broadening in the dot is much smaller than the level spacing, i.e., that the dot only couples weakly to the outside world. Then we approximate the repulsion of each of the N electrons with the $N - 1$ other electrons by an energy contribution $E_c N(N - 1)/2$ [399–402], where the charging energy per pair E_c is inversely proportional to the volume of the dot. Therefore, in the case of a one-dimensional dot of length L , it is proportional to the inverse of the wire length, $E_c \propto 1/L$, which we will use later for Majorana wires. In addition, a gate in proximity to the dot so that the energy for each electron is lowered by the gate voltage V_g , giving rise to a contribution $-eV_g N$. This results in the charging energy

$$E_{\text{ch}}(N, V_g) = \frac{E_c}{2} N(N - 1) - eV_g N . \quad (2.210)$$

The particle number of the ground state in the dot that minimizes E_{ch} thus depends on the gate voltage. In Fig. 2.27a we plot E_{ch} as a function of V_g for various N , where we first consider

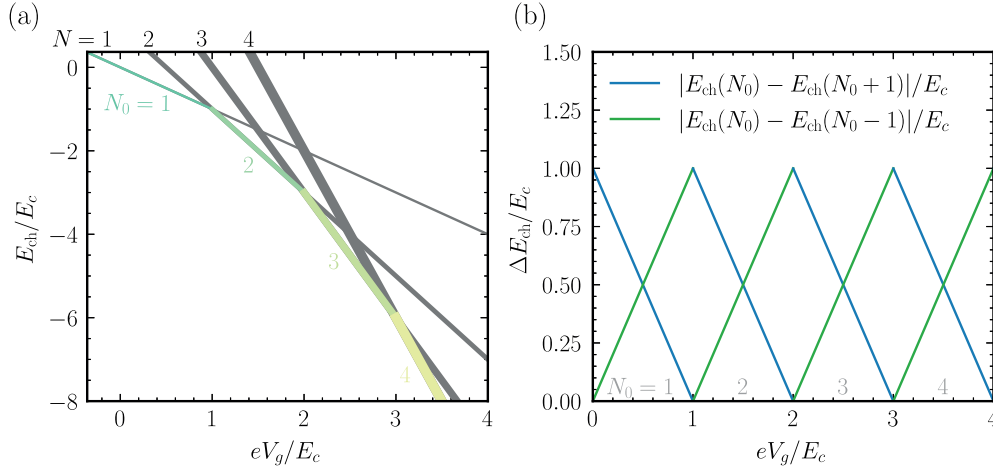


Figure 2.27: Dependence of the charging energy on the gate voltage V_g and particle number N in the dot. (a) Charging energy Eq. (2.210) as a function of V_g for several particle numbers N . A thicker line indicates a larger number of particles. The ground state particle number N_0 minimizes the charging energy (colored lines) and increases by one at the degeneracy points where $E_{\text{ch}}(N_0) = E_{\text{ch}}(N_0 + 1)$ given by $eV_g = E_c N_0$. (b) Distance of the charging energy ΔE_{ch} between optimal particle number N_0 and $N_0 \pm 1$ (blue for an additional particle and green for one electron less).

the metallic limit in which only the charging energy and not the single particle levels are taken into account. Initially, for small voltages $eV_g < E_c$, $N_0 = 1$ gives the smallest charging energy, then increasing V_g , N_0 increases by one after each intersection of $E_{\text{ch}}(N_0, V_g)$ with the energy for the subsequent particle number $E_{\text{ch}}(N_0 + 1, V_g)$ given by

$$E_{\text{ch}}(N, V_g) = E_{\text{ch}}(N + 1, V_g) \Rightarrow eV_g = E_c N . \quad (2.211)$$

In case of the quantum dot, we also need to consider the energy levels ε_i , so that the total energy for N_0 electrons occupying the first N_0 levels is given by [397, 399]

$$E(N_0, V_g) = \sum_{i=1}^{N_0} \varepsilon_i + \frac{E_c}{2} N_0(N_0 - 1) - eV_g N_0 . \quad (2.212)$$

Here, the degeneracy points of energies for N and $N + 1$ particles, $E(N, V_g) = E(N + 1, V_g)$, are given by $eV_g = \varepsilon_{N_0+1} + E_c N$. This means that for the transition $N_0 \rightarrow N_0 + 1$, we additionally need to account for occupying level ε_{N_0+1} , such that the N_0 that minimizes the energy can be expressed by the self-consistency expression as a function of V_g [305]

$$N_0 \approx \left\lceil \frac{eV_g - \varepsilon_{N_0+1}}{E_c} \right\rceil , \quad (2.213)$$

where $\lceil \cdot \rceil$ denotes the ceiling function.

In Fig. 2.27b, we show the energy difference between the charging energy for N_0 and for $N_0 \pm 1$ electrons, yielding the typical Coulomb diamonds often found in the literature [398, 401, 402]. For our later considerations, we are particularly interested in gate voltages in the middle between the degeneracy points $eV_g \sim E_c(N_0 + 1/2)$, where the charging energy difference to both

transitions becomes the largest, and thus a well-defined particle number exists for sufficiently large E_c . In this case, the charging energy forbids sequential tunneling, and the dominant transport mechanism from lead L to R via the dot is elastic cotunneling [397, 398, 403, 404], which can be described by effective one-particle levels $\varepsilon_{\text{eff},i}$. For an associated electron-like process, the electron arriving from lead L occupies an empty level $\varepsilon_{\text{eff},i>N_0}$ before an electron leaves the dot to R, so that there are $N_0 + 1$ electrons in the dot for a very short time $t \sim \hbar/E_c$ only. The seemingly missing energy E_c for this short time can be explained by fluctuations allowed by the Heisenberg uncertainty principle. The analogous hole-like process involves an occupied level $\varepsilon_{\text{eff},i \leq N_0}$. To obtain the effective energy levels for cotunneling as a combination of dot levels ε_m and charging energy $E_{\text{ch}}(N_0, V_g)$, we consider the dot Hamiltonian

$$H_{\text{dot}} = \sum_m \left[\varepsilon_m + \frac{E_c}{2} \sum_{l \neq m} n_l - eV_g \right] n_m, \quad (2.214)$$

where $n_l = d_l^\dagger d_l$ is the particle number operator for level ε_l . We perform a Hartree approximation, i.e., we define fluctuations around the mean $\delta n_l = n_l - \langle n_l \rangle$ and neglect fluctuations to second order, $\mathcal{O}(\delta n^2)$, which gives the Hamiltonian

$$H_{\text{dot}} = \sum_m \left[\varepsilon_m + E_c \sum_{l \neq m} \langle n_l \rangle - eV_g \right] n_m + \mathcal{O}(\delta n^2) + \text{const.} \quad (2.215)$$

This yields the effective energy levels as [405]

$$\varepsilon_{\text{eff},m} = \begin{cases} \varepsilon_m + E_c N_0 - eV_g & , m > N_0 \\ \varepsilon_m + E_c (N_0 - 1) - eV_g & , m \leq N_0 \end{cases}, \quad (2.216)$$

These levels can be used to describe both electron-like and hole-like processes [398]. If the gate voltage V_g is chosen such that $\varepsilon_{\text{eff},N_0+1}$ vanishes, an electron can jump from the Fermi level in the lead to the dot without additional energy, and sequential tunneling is possible, resulting in a conductance resonance after which the optimal particle number increases by one if eV_g is further increased. If V_g is tuned sufficiently far away from the resonances, where in Coulomb blockade cotunneling is the dominant transport mechanism, a much smaller conductance is observed. We will address the calculation of this conductance in the next sections, but before that we discuss what influence superconductivity in the dot has on the effective levels and couplings.

For this purpose, we consider a short Majorana wire (Fig. 2.26b), which is separated from the leads by pinch-off gates that generate a steep confinement potential, so that we can describe the wire as a quantum dot. Diagonalizing the Hamiltonian of the Majorana wire Eq. (2.146), we obtain energy eigenvalues $\pm \mathcal{E}_i$ and eigenstates $(\mathbf{u}_i, \mathbf{v}_i)$ with particle-like wave functions \mathbf{u}_i and hole-like wave functions \mathbf{v}_i . Here, particle-hole symmetry ensures that for every $\mathcal{E}_i > 0$ with $(\mathbf{u}_i, \mathbf{v}_i)$, there also exists an eigenstate with negative energy $-\mathcal{E}_i$ and $(\mathbf{v}_i^*, \mathbf{u}_i^*)$. The wave functions are important because their value at the lead-wire interfaces y_α , $\alpha = \text{L, R}$, determine the couplings $\lambda_{\text{eff},i,\sigma,\alpha}^e$ of the lead α and the dot level i [304] (detail later in Sec. 2.3.3). For an unoccupied Bogoliubov level \mathcal{E}_i , the electron-like effective levels and couplings are then given as

$$\varepsilon_{\text{eff},m}^e = \mathcal{E}_i + E_c N_0 - eV_g \quad (2.217)$$

$$\lambda_{\text{eff},m,\sigma,\alpha}^e = u_{i,\sigma}(y_\alpha), \quad (2.218)$$

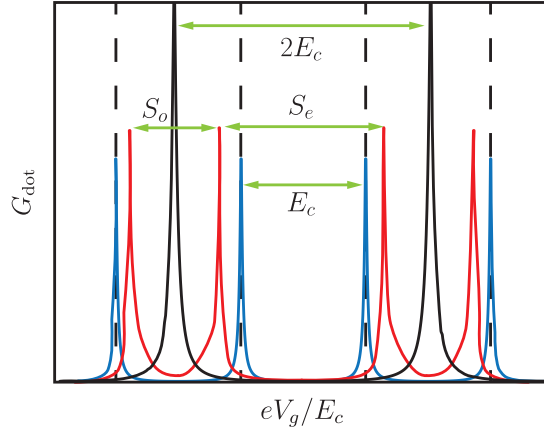


Figure 2.28: Spacing of conductance resonances depending on the lowest energy level \mathcal{E}_0 in the superconductor. Here, $S_o \approx E_c - 2\mathcal{E}_0$ is the spacing between conductance resonances for odd particle numbers and $S_e \approx E_c + 2\mathcal{E}_0$ for an even number of electrons. As the particle number N_0 increases by one after each resonance, the spacing alternates between S_o and S_e (red). In the case of MZMs where $\mathcal{E}_0 \approx 0$, this spacing difference vanishes (blue), which is a signature for any zero energy mode in the middle of the superconducting gap. In the case where the charging energy is smaller than \mathcal{E}_0 (black), adding an electron to obtain odd N_0 costs more energy than adding two electrons (a Cooper pair) to the condensate and keeping an even particle number, such that the spacing becomes $2E_c$. Vanishing of the even-odd splitting in Majorana wires is demonstrated in Ref. [95], where they show a similar figure.

and for a hole-like process

$$\varepsilon_{\text{eff},m}^h = -\mathcal{E}_i + E_c(N_0 - 1) - eV_g \quad (2.219)$$

$$\lambda_{\text{eff},m,\sigma,\alpha}^h = v_{i,\sigma}(y_\alpha) . \quad (2.220)$$

However, Cooper pairs in the superconductor consist of pairs of electrons, which leaves an unpaired electron in the case of an odd particle number N_0 , which has to occupy the lowest Bogoliubov level \mathcal{E}_1 for the ground state at zero temperature [129, 304]. Therefore, in the case of an odd particle number, the additional energy \mathcal{E}_1 is necessary, which also leads to a signature for MZMs: The distance between conductance resonances for an even particle number N_0 it is given by $S_e \propto E_c + 2\mathcal{E}_1$, whereas for odd particle number $S_o \propto E_c - 2\mathcal{E}_1$ [95, 406, 407]. On the one hand, in a trivial superconductor with proximity gap Δ_{prox} , the spacings clearly alternate with length differences $4\Delta_{\text{prox}}$ (blue curve Fig. 2.28). The vanishing of these even-odd spacings differences, on the other hand, shows that the energy of the first level is zero, $\mathcal{E}_1 = 0$, which is therefore a signature for MZMs [95] (red line in Fig. 2.28), but also for any other zero-energy excitation in the middle of the superconductor gap. For the case where the charging energy is smaller than the gap, $E_c < \Delta_{\text{prox}}$, the ground states with odd particle numbers are completely skipped, since it is energetically more favorable to directly add two electrons as a Cooper pair into the dot which costs an energy of $2E_c$, than to add a single electron that occupies the first level and in addition requires the energy E_c (black line in Fig. 2.28).

For the case where a Bogoliubov level i is occupied, such as for the ground state of odd N_0 , we find in Sec. 2.3.3 that the associated effective energies are modified by $\mathcal{E}_i \rightarrow -\mathcal{E}_i$ and for the couplings $u_{i,\sigma} \rightarrow v_{i,\sigma}^*$, which is consistent with the previous descriptions. However, there is a problem with the BdG description of the superconductor: the Hamiltonian of the Majorana wire

Eq. (2.146) does not describe a fixed electron number N_0 , but is a grand-canonical description with chemical potential μ . To obtain a good approximation for a wire with N_0 electrons [408], we choose the chemical potential μ self-consistently, so that

$$\langle \hat{N} \rangle = \sum_i |\mathbf{v}_i(\mu)|^2 = N_w, \quad (2.221)$$

i.e., the average particle number in the wire is given by N_w . We distinguish N_0 , the number of electrons in the dot consisting of wire and superconductor, from the average number of particles N_w in the wire. Here, we model only the Majorana wire with superconductivity induced by the proximity effect [291, 409, 410], since its energy level and eigenstates are relevant to the scattering matrix description because the induced gap is much smaller than the s-wave gap in the superconductor. However, most of the charge of an electron tunneling into the dot lives in the superconductor [304], where we assume that $\langle \hat{N} \rangle$ in the wire effectively changes only by $1/20$ when an electron tunnels into the dot [304]. As a result, the self-consistently chosen chemical potential μ also only weakly changes when $N_0 \rightarrow N_0 + 1$, $N_w \rightarrow N_w + 1/20$.

Because MZMs give rise to a highly non-local Bogoliubov wave function exponentially localized at both ends of the wire, the chance of electrons entering the dot at the wire ends is greatly enhanced [129–132, 304]. This gives a relatively large chance of a cotunneling process to take place, where one electron enters the left MZM and one electron exits the dot from the right MZM, such that a charge e is transmitted from lead L to lead R, giving rise to an increased conductance. Due to the localization, the weight of the wave function in the middle of the wire is zero, yet cotunneling allows for transport through the dot which is why this is sometimes called *electron teleportation* in the literature [129].

2.3.3 Scattering matrix approach at finite temperatures

In this section, we discuss the scattering matrix formalism used to compute the amplitude of coherent transmission $|\langle T_{\uparrow\uparrow} + T_{\downarrow\downarrow} \rangle|$ through a Majorana wire in Coulomb-blockade, where $|T_{\sigma\sigma}|^2$ is the probability of an electron with spin σ to tunnel from the left to the right lead. As discussed in Sec. 2.3.2, this transmission amplitude is experimentally accessible [130, 411] from measuring current oscillations through an Aharonov-Bohm interferometer [17]. We published a more condensed version of the following description of the modified scattering matrix formalism that allows us to include Coulomb-blockade for large charging energies in Ref. [131].

In the following, we assume that the Majorana wire is tuned to the middle between conductance resonances by an adjacent gate with voltage $V_{g,\text{mid}}$ [131, 304]. In this case, the number of electrons N_0 in the ground state of the dot is well-defined, and we again assume that this results in an average particle number $N_w = N_0/20$ in the wire, which we ensure in our calculations by self-consistently determining the chemical potential in the wire [304]. If the charging energy E_c for adding an electron to the dot is much larger than the level broadening due to coupling to the two leads, elastic cotunneling is the dominant transport mechanism [398], and adding or removing more than one electron from the dot is energetically very unfavorable. It is therefore a good approximation to truncate the Hilbert space to only including states with N_0 and $N_0 \pm 1$ electrons in the dot and to consider cotunneling through effective single particle energy levels.

The scattering matrix formalism also makes it possible to consider finite temperatures by allowing the dot to be in a thermally excited state when the tunneling process takes place. For describing the elastic coherent transport, we assume that the dot has the same final and initial state with occupation numbers $\{n_i\}$ of the Bogoliubov levels \mathcal{E}_i with excitation energy $E(\{n_i\}) = \sum_i n_i \mathcal{E}_i$. The scattering matrix S , whose entries contain the complex transmission amplitudes

$\langle T_{\uparrow\uparrow} \rangle = S_{13}$, and $\langle T_{\downarrow\downarrow} \rangle = S_{24}$, is then determined by a modified Mahaux-Weidenmüller formula [131, 304, 394]

$$S = 1 - 2\pi i \left\langle W \frac{1}{\varepsilon - H_{\text{eff}} + i\pi W^\dagger W} W^\dagger \right\rangle \quad (2.222)$$

with the thermal average

$$\langle \mathcal{O} \rangle = \frac{1}{Z} \sum_{\{n_i\}} e^{-\beta E(\{n_i\})} \mathcal{O}(\{n_i\}) . \quad (2.223)$$

Here, the Boltzmann weights depend on the excitation energy and the inverse temperature $\beta = (k_B T)^{-1}$, and the thermal average is normalized by $Z = \sum_{\{n_i\}} \exp[-\beta E(\{n_i\})]$. We perform the thermal average over the entries of the scattering matrix and hence $T_{\sigma\sigma}$, since this correctly describes the oscillation amplitude of the total current through the interferometer when thermally averaging the current as we found in Eq. (2.209). Here, ε is the energy of incoming electrons from the leads, H_{eff} is the diagonal matrix of effective energy levels, and W is the coupling matrix containing the effective hole-like and electron-like dot-lead couplings for spin σ and lead α [131], i.e.

$$H_{\text{eff}} = \begin{pmatrix} \text{diag} [\varepsilon_{\text{eff},j}^{(h)}(N_0, \{n_i\})]_{j=1..j_{\text{max}}} & 0 \\ 0 & \text{diag} [\varepsilon_{\text{eff},j}^{(e)}(N_0, \{n_i\})]_{j=1..j_{\text{max}}} \end{pmatrix} , \quad (2.224)$$

$$\frac{(W)_{\alpha\sigma}}{\sqrt{\varrho_F}} = \left(\lambda_{\text{eff},\alpha 1\sigma}^{(h)}(N_0, \{n_i\}), \dots, \lambda_{\text{eff},\alpha j_{\text{max}}\sigma}^{(h)}(N_0, \{n_i\}), \lambda_{\text{eff},\alpha 1\sigma}^{(e)}(N_0, \{n_i\}), \dots, \lambda_{\text{eff},\alpha j_{\text{max}}\sigma}^{(e)}(N_0, \{n_i\}) \right) . \quad (2.225)$$

Here, ϱ_F is the density of states at the Fermi level of the leads, and we include a number of j_{max} levels in the calculation of S , where contributions of levels get smaller for higher levels with larger energies. In addition, we will find later that contributions of higher levels cancel each other.

In the following, we discuss the details of the scattering matrix formalism and derive expressions for effective levels and couplings.

Truncation of the Hilbert space

The strong Coulomb-blockade fixes the particle number N_0 in the ground state of the dot and repels additional electrons. Therefore, simultaneous tunneling processes of multiple electrons or holes, and thus also tunneling of Cooper pairs and Andreev processes, are suppressed [131, 304]. To describe an elastic cotunneling process, we start from an initial dot state $|N_0, \{n_i\}\rangle$ with N_0 particles, which reside in the condensate or occupy Bogoliubov levels as described by the occupation numbers $\{n_i\}$. An electron-like tunneling process then happens via an intermediate dot state $|N_0 + 1, \{\tilde{n}_i\}\rangle$, where an electron hops into the dot and then an electron quickly leaves the dot towards the right lead to produce the final dot state $|N_0, \{n_i\}\rangle$. For a hole-like process, the intermediate state has the form $|N_0 - 1, \{\tilde{n}_i\}\rangle$, where the dot has $N_0 - 1$ electrons.

Hence, strong Coulomb-blockade allows us to truncate the Hilbert space to only these states when describing scattering processes.

Hamiltonian

We consider the Hamiltonian [131, 304, 305]

$$H = H_{\text{dot}} + H_{\text{tun}} + H_{\text{lead}} , \quad (2.226)$$

where H_{dot} is the dot Hamiltonian, H_{lead} the Hamiltonian for the leads, and H_{tun} describes tunneling between dot and leads. In the following, we describe the individual contributions to H in detail. We start with the dot Hamiltonian

$$H_{\text{dot}} = H_w + H_{\text{ch}} , \quad (2.227)$$

where H_w is the Majorana wire Hamiltonian and H_{ch} the charging contribution due to Coulomb-blockade. In the Nambu basis $\psi^\dagger(y) = (\psi_\uparrow^\dagger(y), \psi_\downarrow^\dagger(y), \psi_\downarrow(y), -\psi_\uparrow(y))$, we model the one-dimensional proximitized nanowire as

$$H_w = \frac{1}{2} \int dy \psi^\dagger(y) \mathcal{H}_w \psi(y) \quad (2.228)$$

$$= \frac{1}{2} \int dy \psi^\dagger(y) \left\{ \tau_z \left[\sigma_0 \frac{-\hbar^2}{2m^*} \partial_y^2 - \sigma_0 \mu - i\hbar \alpha_R \sigma_x \partial_y \right] - E_z \tau_0 \sigma_z + \Delta \tau_x \sigma_0 \right\} \psi(y) , \quad (2.229)$$

where Δ is the proximity s-wave gap after integrating out the superconductor (here approximated by a real constant), E_z is the Zeeman energy due to the orthogonal magnetic field, μ is the chemical potential, and $-i\hbar \alpha_R \sigma_x \partial_y$ describes the strong Rashba spin orbit coupling. We discretize this Hamiltonian on a grid of N sites with spacing $a = 0.026 l_{\text{so}}$ and then diagonalize the resulting matrix

$$\mathcal{H}_w \begin{pmatrix} \Psi_i(1) \\ \vdots \\ \Psi_i(N) \end{pmatrix} = \mathcal{E}_i \begin{pmatrix} \Psi_i(1) \\ \vdots \\ \Psi_i(N) \end{pmatrix} , \quad (2.230)$$

where \mathcal{E}_i are the energy eigenvalues, and wave functions are given by $\Psi_i(y) = (\mathbf{u}_i(y), \mathbf{v}_i(y))^T$. This diagonalization gives rise to new Bogoliubov operators [305]

$$\beta_n = \sum_\sigma \int dy \left[u_{n\sigma}^*(y) \psi_\sigma(y) e^{i\phi/2} + v_{n\sigma}^*(y) \psi_\sigma^\dagger(y) e^{-i\phi/2} \right] , \quad (2.231)$$

where the particle-hole symmetry of H_w ensures that for each β_n with $\mathcal{E}_n > 0$, $(\mathbf{u}_n, \mathbf{v}_n)$ there exists a state $-\mathcal{E}_n$, $(\mathbf{v}_n^*, \mathbf{u}_n^*)$ corresponding to β_n^\dagger . We use particle-hole symmetry to express the original fermion creation and annihilation operators as

$$\psi_\sigma(y) = \sum_n u_{n\sigma}(y) \beta_n e^{-i\phi/2} = \sum_{\substack{n \\ \mathcal{E}_n > 0}} \left[u_{n\sigma}(y) \beta_n + v_{n\sigma}^*(y) \beta_n^\dagger \right] e^{-i\phi/2} \quad (2.232)$$

$$\psi_\sigma^\dagger(y) = \sum_n v_{n\sigma}(y) \beta_n e^{i\phi/2} = \sum_{\substack{n \\ \mathcal{E}_n > 0}} \left[v_{n\sigma}(y) \beta_n + u_{n\sigma}^*(y) \beta_n^\dagger \right] e^{i\phi/2} . \quad (2.233)$$

Then using $\{\beta_n, \beta_{n'}^\dagger\} = \delta_{nn'}$, the diagonal wire Hamiltonian is given by

$$H_w = \frac{1}{2} \sum_n \mathcal{E}_n \beta_n^\dagger \beta_n = \sum_{\substack{n \\ \mathcal{E}_n > 0}} \mathcal{E}_n \beta_n^\dagger \beta_n + \text{const} . \quad (2.234)$$

Eq. (2.232) and Eq. (2.233) directly follow from using the diagonalization Eq. (2.231), except for the non-trivial terms $e^{\pm i\phi/2}$ involving the condensate phase operator ϕ , conjugate to the particle number operator [412, p. 52 ff.]. This helps to overcome the problem that, on the one hand, we would like to describe single particle transport through the dot with defined particle number in Coulomb-blockade, but on the other hand, the convenient BdG formalism does not describe fixed particle numbers [333, 412]. This is resolved by formally separating particle number N_0 and quasi-particle occupations $\{n_i\}$, considering states $|N_0, \{n_i\}\rangle$ and using the following rules

i) We choose the chemical potential of the wire to fix the average particle number in the dot to N_0 [408].

ii) The operator β_j (β_j^\dagger) empties (occupies) the Bogoliubov level j but leaves the particle number unchanged

$$\beta_j^\dagger |N_0, \{n_j = 0, \dots\}\rangle = |N_0, \{n_j = 1, \dots\}\rangle \quad (2.235)$$

$$\beta_j^\dagger |N_0, \{n_j = 1, \dots\}\rangle = 0 . \quad (2.236)$$

iii) The operators $e^{\pm i\phi/2}$ increases/lowers the particle number by one

$$e^{\pm i\phi/2} |N_0, \{n_i\}\rangle = |N_0 \pm 1, \{n_i\}\rangle . \quad (2.237)$$

The second contribution to the dot Hamiltonian is the charging energy expressed with the particle number operator $d_j^\dagger d_j$ of dot level j as [131, 305]

$$H_{\text{ch}} = \sum_j \left[-eV_g + \frac{E_c}{2} \sum_{i \neq j} d_i^\dagger d_i \right] d_j^\dagger d_j , \quad (2.238)$$

with voltage V_g of a gate in proximity to the dot, and charging energy E_c due to the repulsion of two electrons. As discussed in Sec. 2.3.2, we treat this Hamiltonian in a mean field approximation which yields a charging energy for N_0 electrons in the dot given by [398, 402]

$$E_{\text{ch}}(N_0, V_g) = \frac{E_c}{2} N_0(N_0 - 1) - eV_g N_0 . \quad (2.239)$$

For simplicity, we assume that the leads are normal-conducting and without spin-orbit coupling such that particle and hole sector are independent, and spin is a good quantum number. Thus, the lead Hamiltonian is given by

$$H_{\text{lead}} = \int dy \mathbf{c}^\dagger(y) \left\{ \tau_z \sigma_0 \left[-\frac{\hbar^2}{2m^*} \partial_y^2 - V_{\text{lead}} \right] - E_z \tau_0 \sigma_z \right\} \mathbf{c}(y) , \quad (2.240)$$

where $V_{\text{lead}} = 100 E_{\text{so}}$ is chosen such that both spin species are present near the Fermi level of the lead, and $c_\sigma^\dagger(y)$ creates an electron with spin σ at position y in the lead.

The weak coupling between dot and leads ($\alpha = R, L$) is described by a tunneling Hamiltonian H_{tun} . We first discuss an effective coupling [304]

$$H_{\text{tun}} = \sum_{j\sigma\alpha} t_{j\alpha\sigma} c_\sigma^\dagger(y_\alpha) d_j + \text{h.c.} , \quad (2.241)$$

where $c_\sigma^\dagger(y_\alpha)$ creates an electron at the dot-lead interface y_α in lead α , d_j (d_j^\dagger) is the destruction (creation) operator of dot level j . Here, the effective dot-lead couplings $t_{j\alpha\sigma}$ are approximated

as the overlap between an incoming wave $\Psi_{\alpha\sigma}(y)$ from lead α that decays exponentially in the tunneling barrier and the dot eigenfunction $\varphi_j(y)$ in the normal-conducting state $\Delta = 0$ [131], i.e.

$$t_{j\alpha\sigma} = t_0 \int dy \Psi_{\alpha\sigma}(y) \varphi_j(y) . \quad (2.242)$$

In this effective description, t_0 is a real number that is chosen such that the coupling is sufficiently weak to justify Coulomb-blockade, and we assume that $\Psi_{\alpha\sigma}(y) \propto e^{-y/\lambda}$ with $\lambda = 0.26 l_{\text{so}}$ is just a generic incoming wave that decays together with φ_i within a tunnel barrier between lead and dot (not necessarily modelled in this simple picture). We will later also consider simulations with two different microscopic models, where both dot and leads are treated more carefully by (a) using a Hamiltonian that contains leads, dot, and tunneling barrier allowing us to obtain the lead wave functions and explicitly computing the wave function overlap in the barrier (Sec. 2.3.4), and in the second model by (b) integrating out translationally invariant leads and using the resulting self energies in the scattering approach (Sec. 2.3.5). We find that the effective model captures the most important ingredients to understand transport in the topological phase.

For obtaining the effective scattering energy levels and coupling, it will be advantageous to express the tunneling Hamiltonian in terms of the BdG operators, Eq. (2.231). For this, we use eigenfunctions φ_j of \mathcal{H}_w for the normal-conducting case, $\Delta = 0$, to express the annihilation operator in dot level j as [304]

$$d_j = \sum_{\sigma} \int dy \varphi_{j\sigma}^*(y) \psi_{\sigma}(y) . \quad (2.243)$$

Using the expressions of the fermion operators in terms of BdG operators Eqs. (2.233) and (2.232), one finds the tunneling Hamiltonian [131, 304, 305]

$$H_{\text{tun}} = \sum_{j\alpha\sigma} c_{\sigma}^{\dagger}(y_{\alpha}) e^{-i\phi/2} \left[\lambda_{j\alpha\sigma}^u(\mu) \beta_j(\mu) + \lambda_{j\alpha\sigma}^v(\mu) \beta_j^{\dagger}(\mu) \right] + \text{h.c.} \quad (2.244)$$

with couplings

$$\lambda_{j\alpha\sigma}^u(\mu) = \int dy \sum_{m\sigma'} t_{m\alpha\sigma} \varphi_{m\sigma'}^*(y) u_{n\sigma'}(y; \mu) \quad (2.245)$$

$$\lambda_{j\alpha\sigma}^v(\mu) = \int dy \sum_{m\sigma'} t_{m\alpha\sigma} \varphi_{m\sigma'}^*(y) v_{n\sigma'}^*(y; \mu) . \quad (2.246)$$

Using the definition of $t_{m\alpha\sigma}$, Eq. (2.242), and the orthonormality of the basis φ_i these couplings can be expressed as

$$\begin{aligned} \lambda_{j\alpha\sigma}^u(\mu) &= t_0 \int dy \Psi_{\alpha\sigma}(y) u_{n\sigma}(y; \mu) \approx t_0 u_{n\sigma}(y_{\alpha}; \mu) \\ \lambda_{j\alpha\sigma}^v(\mu) &= t_0 \int dy \Psi_{\alpha\sigma}(y) v_{n\sigma}^*(y; \mu) \approx t_0 v_{n\sigma}^*(y_{\alpha}; \mu) . \end{aligned} \quad (2.247)$$

Here, y_{α} is the position of the dot-lead interface for lead α , e.g. for a wire of length L with $y \in [0, L]$ this is $y_L = 0$, $y_R = L$. Hence, for the effective model, Eq. (2.247) shows that the couplings are proportional to the values of the hole $v_{n\sigma}^*(y; \mu)$ and electron $u_{n\sigma}(y; \mu)$ wave functions at the ends of the wire.

Effective energy levels and coupling matrix elements

We next describe how the effective couplings W and effective energy levels H_{eff} can be obtained from matrix elements of the Hamiltonian Eq. (2.226). Before the tunneling process takes place, we assume the system to be in an initial state with one particle of energy ε and spin σ in lead α , the Fermi sea $|0\rangle_{\bar{\alpha}}$ in the other lead $\bar{\alpha}$, and the dot with N_0 electrons with occupation numbers of the Bogoliubov levels given by $\{n_i\}$, i.e.

$$|\psi_i(N_0)\rangle = |(\varepsilon, \sigma)\rangle_{\alpha} |0\rangle_{\bar{\alpha}} |N_0, \{n_i\}\rangle. \quad (2.248)$$

The energy of this state relative to the energy of the incoming particle is then given by

$$E_i(N_0) - \varepsilon = \langle \psi_i(N_0) | H | \psi_i(N_0) \rangle - \varepsilon = \sum_j n_j \mathcal{E}_j + E_{\text{ch}}(N_0, V_g). \quad (2.249)$$

After the elastic tunneling process, a particle is transferred to the other lead $\bar{\alpha}$ and the dot is again in the initial state such that the energy is conserved

$$|\psi_f(N_0)\rangle = |0\rangle_{\alpha} |(\varepsilon, \sigma')\rangle_{\bar{\alpha}} |N_0, \{n_i\}\rangle \quad (2.250)$$

$$E_f(N_0) = E_i(N_0). \quad (2.251)$$

This transfer from lead α to $\bar{\alpha}$ by elastic cotunneling can be described via an intermediate, temporary state [305, 398]

$$|\psi_{\text{tmp}}(N_0 \pm 1)\rangle = |0\rangle_{\alpha} |0\rangle_{\bar{\alpha}} |N_0 \pm 1, \{\tilde{n}_i\}\rangle \quad (2.252)$$

$$E_{\text{tmp}}(N_0 \pm 1) = \sum_j \tilde{n}_j \mathcal{E}_j + E_{\text{ch}}(N_0 \pm 1, V_g), \quad (2.253)$$

where the electron or hole resides on the dot for a very short time allowed by the Heisenberg uncertainty. As most of the charge in the dot lives in the superconductor, not the wire, we can use the approximation $\mu(N_0) \approx \mu(N_0 \pm 1)$ in the wire Hamiltonian, such that we only self-consistently determine the chemical potential of the wire to reflect the particle number in the intermediate state, thus, reducing the number of self-consistent computations that require the full eigenstates of the wire Hamiltonian in each step.

From the form of the tunneling Hamiltonian H_{tun} , Eq. (2.246), it becomes apparent that an electron can enter the dot either by occupying an empty ($n_j = 0$) level $\beta_j^\dagger e^{i\phi/2}$ or by emptying an occupied ($n_j = 1$) level $\beta_j e^{i\phi/2}$ and forming a Cooper pair. In addition, in the BdG formalism an electron (e) with energy ε tunneling from $\alpha = L$ to $\alpha = R$ corresponds to a hole (h) with $-\varepsilon$ from R to L , and in the normal-conducting leads, electron and hole blocks are completely independent. Therefore, the effective electron and hole energy levels are defined via [131, 305]

$$\varepsilon_{\text{eff}}^{(e)} = E_{\text{tmp}}(N_0 + 1) - E_i(N_0) \quad (2.254)$$

$$\varepsilon_{\text{eff}}^{(h)} = E_i(N_0) - E_{\text{tmp}}(N_0 - 1), \quad (2.255)$$

and the corresponding couplings are given by matrix elements

$$\lambda_{\text{eff}}^{(e)} = \langle \psi_i(N_0) | H_{\text{tun}} | \psi_{\text{tmp}}(N_0 + 1) \rangle \quad (2.256)$$

$$\lambda_{\text{eff}}^{(h)} = \langle \psi_{\text{tmp}}(N_0 - 1) | H_{\text{tun}} | \psi_i(N_0) \rangle = \langle \psi_i(N_0) | H_{\text{tun}} | \psi_{\text{tmp}}(N_0 - 1) \rangle^*. \quad (2.257)$$

To also keep track of spin σ , lead α , and BdG level j (with $\mathcal{E} \geq 0$), we list the possible tunneling processes within the truncated Hilbert space in the following:

- i) An electron-like process ($N_0 \rightarrow N_0 + 1$) through an empty level j , where $n_j = 0$, results in intermediate occupations $\{\tilde{n}_i\} = \{\tilde{n}_{i \neq j} = n_i, \tilde{n}_j = 1\}$, such that

$$\begin{aligned}\varepsilon_{\text{eff},j}^{(e)}(\mu) &= \sum_i \tilde{n}_i \mathcal{E}_i(\mu) + E_{\text{ch}}(N_0 + 1, V_g) - \left[\sum_i n_i \mathcal{E}_i(\mu) + E_{\text{ch}}(N_0, V_g) \right] \\ &= \mathcal{E}_j(\mu) + E_c N_0 - eV_g \quad \text{for } n_j = 0\end{aligned}\quad (2.258)$$

$$\begin{aligned}\lambda_{\text{eff},j,\sigma,\alpha}^{(e)}(\mu) &= \langle (e, \varepsilon, \sigma) |_{\alpha} \langle 0 |_{\bar{\alpha}} \langle N_0, \{n_i\} | \left\{ \sum_{l\sigma'\alpha'} c_{\sigma'}^{\dagger}(y_{\alpha'}) e^{-i\phi/2} \left[\lambda_{l\alpha'\sigma'}^u(\mu) \beta_l(\mu) \right. \right. \\ &\quad \left. \left. + \lambda_{l\alpha'\sigma'}^v(\mu) \beta_l^{\dagger}(\mu) \right] + \text{h.c.} \right\} | N_0 + 1, \{\tilde{n}_i\} \rangle | 0 \rangle_{\bar{\alpha}} | 0 \rangle_{\alpha} \\ &= \lambda_{j\alpha\sigma}^u \quad \text{for } n_j = 0.\end{aligned}\quad (2.259)$$

Here, (e, ε, σ) indicates a particle of spin σ and energy ε in the electron-sector e of the lead Hamiltonian.

- ii) An electron-like process ($N_0 \rightarrow N_0 + 1$) through an occupied level j , where $n_j = 1$, results in intermediate occupations $\{\tilde{n}_i\} = \{\tilde{n}_{i \neq j} = n_i, \tilde{n}_j = 0\}$, such that

$$\varepsilon_{\text{eff},j}^{(e)}(\mu) = -\mathcal{E}_j(\mu) + E_c N_0 - eV_g \quad \text{for } n_j = 1 \quad (2.260)$$

$$\lambda_{\text{eff},j,\sigma,\alpha}^{(e)}(\mu) = \lambda_{j\alpha\sigma}^v \quad \text{for } n_j = 1. \quad (2.261)$$

- iii) A hole-like process ($N_0 \rightarrow N_0 - 1$) through an empty level j , where $n_j = 0$, results in intermediate occupations $\{\tilde{n}_i\} = \{\tilde{n}_{i \neq j} = n_i, \tilde{n}_j = 1\}$, such that

$$\begin{aligned}\varepsilon_{\text{eff},j}^{(h)}(\mu) &= \sum_i \delta_{n_i,1} \mathcal{E}_i(\mu) + E_{\text{ch}}(N_0, V_g) - \left[\sum_i \delta_{\tilde{n}_i,1} \mathcal{E}_i(\mu) + E_{\text{ch}}(N_0 - 1, V_g) \right] \\ &= -\mathcal{E}_j(\mu) + E_c(N_0 - 1) - eV_g \quad \text{for } n_j = 0\end{aligned}\quad (2.262)$$

$$\begin{aligned}\lambda_{\text{eff},j,\sigma,\alpha}^{(h)}(\mu) &= \langle (h, \varepsilon, \sigma) |_{\alpha} \langle 0 |_{\bar{\alpha}} \langle N_0, \{n_i\} | \left\{ \sum_{l\sigma'\alpha'} c_{\sigma'}(y_{\alpha'}) e^{i\phi/2} \left[\lambda_{l\alpha'\sigma'}^{u*}(\mu) \beta_l^{\dagger}(\mu) \right. \right. \\ &\quad \left. \left. + \lambda_{l\alpha'\sigma'}^{v*}(\mu) \beta_l(\mu) \right] + \text{h.c.} \right\} | N_0 - 1, \{\tilde{n}_i\} \rangle | 0 \rangle_{\bar{\alpha}} | 0 \rangle_{\alpha}^* \\ &= \lambda_{j\alpha\sigma}^v \quad \text{for } n_j = 0.\end{aligned}\quad (2.263)$$

Here, $c_{\sigma}(y_{\alpha})$ creates a hole in the hole-sector of lead α .

- iv) A hole-like process ($N_0 \rightarrow N_0 - 1$) through an occupied level j , where $n_j = 1$, results in intermediate occupations $\{\tilde{n}_i\} = \{\tilde{n}_{i \neq j} = n_i, \tilde{n}_j = 0\}$, such that

$$\varepsilon_{\text{eff},j}^{(e)}(\mu) = \mathcal{E}_j(\mu) + E_c(N_0 - 1) - eV_g \quad \text{for } n_j = 1 \quad (2.264)$$

$$\lambda_{\text{eff},j,\sigma,\alpha}^{(e)}(\mu) = \lambda_{j\alpha\sigma}^u \quad \text{for } n_j = 1. \quad (2.265)$$

Using these effective energy levels and couplings together with the appropriate chemical potentials μ in Eq. (2.222), we can compute the scattering matrix S and, hence, the amplitude of coherent transmission $|\langle T_{\uparrow\uparrow} + T_{\downarrow\downarrow} \rangle|$.

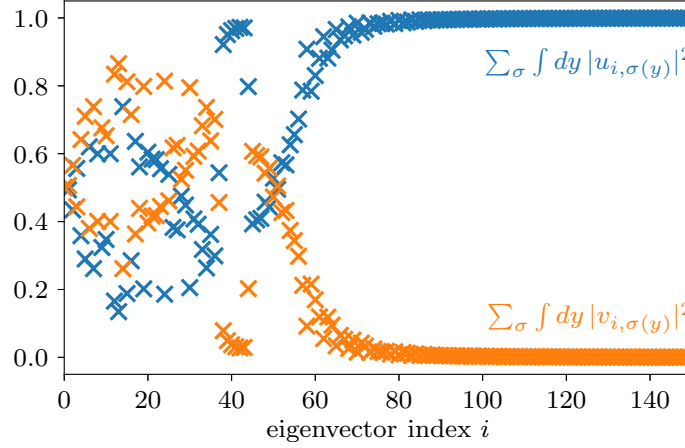


Figure 2.29: Electron and hole character of the energy ordered eigenstates i of a Majorana wire in the topological regime. We individually integrate the probability for the electron sector $\int dy |\mathbf{u}(y)|^2$ (blue crosses) and hole sector $\int dy |\mathbf{v}(y)|^2$ (orange crosses) for the first 150 eigenstates of a Majorana wire Hamiltonian, Eq. (2.161), for a wire of length $L = 32.5 l_{\text{so}}$. We choose $E_z = 6 E_{\text{so}}$, $\mu = 0.5 E_{\text{so}}$, $\Delta = 2 E_{\text{so}}$, and a steep confinement potential Eq. (2.162) with $\sigma_1 = \sigma_2 = 0.1 l_{\text{so}}$ and $V_0 = 65 E_{\text{so}}$. Eigenstates corresponding to smaller energies, have mixed electron-hole character, while for large energies, the states have only a u component, i.e. they are not occupied.

Self-consistent chemical potential

We describe the wire in the BdG formalism, Eq. (2.229), and self-consistently determine the chemical potential such that the average particle number in the wire

$$\langle \hat{N}(\mu) \rangle = \int dy \sum_{\sigma} \langle N_0, \{n_j\} | \psi_{\sigma}^{\dagger}(y) \psi_{\sigma}(y) | N_0, \{n_j\} \rangle \quad (2.266)$$

$$= \int dy \sum_{\substack{j \\ \mathcal{E}_j > 0}} \left[n_j |u_{j\sigma}(y, \mu)|^2 + (1 - n_j) |v_{j\sigma}(y, \mu)|^2 \right] \quad (2.267)$$

is given by $N_w \simeq N_0/20$, where we used Eq. (2.232), and N_0 is the particle number in the dot consisting of wire and superconductor [304]. This yields a good approximation for a fixed particle number description of the superconducting wire, even for small N_0 [408]. For a MZM, where $|\mathbf{u}|^2 = |\mathbf{v}|^2$, the average particle number does not depend on the occupation of the Majorana level, and also for other low energy states, the difference $\int dy (|\mathbf{u}_i|^2 - |\mathbf{v}_i|^2)$ is much smaller than one (see Fig. 2.29). As the self-consistency calculations are resource intensive, we therefore only determine μ self-consistently for the ground state and then use the same μ as an approximation for all higher excitations in the thermal average (details below). In Fig. 2.30, we show self-consistently determined chemical potentials in a Majorana wire as a function of the external Zeeman field such that the total particle number in the dot is fixed to $N_0 = 700$ ($N_w = 35$). Panel (a) depicts the chemical potentials for a ground state with N_0 electrons, intermediate states of $N_0 \pm 1$ electrons, and the excited state $N_0 - 1, \{n_2 = 1, n_{i \neq 2} = 0\}$, which shows that the differences are minor. This is because most of the charge lives in the superconductor such that the wire's chemical potential is almost unchanged for $N_0 \rightarrow N_0 + 1$, i.e. $\mu(N_0) \approx \mu(N_0 \pm 1)$. This allows us to only use the chemical potential of the intermediate

state when computing effective energy levels and couplings.

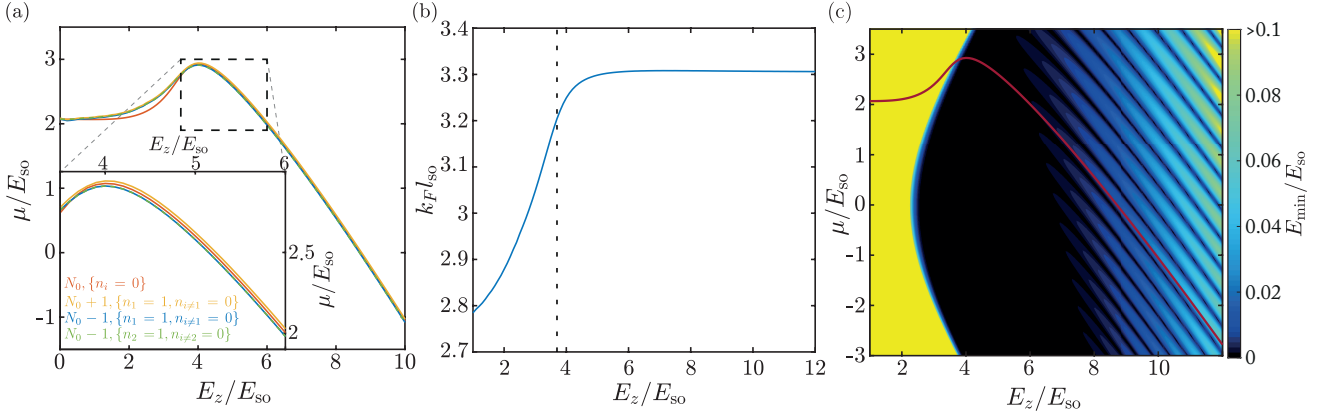


Figure 2.30: Self-consistently determined chemical potential to fix the particle number in the dot to $N_0 = 20 \cdot 35 \equiv 20N_w$ as a function of the applied Zeeman field E_z . (a) Chemical potential for even particle number N_0 in the dot with only empty Bogoliubov levels (solid red). For comparison, we also show the self-consistent chemical potential for $N_0 \pm 1$ where the additional electron resides in the first level (solid yellow and solid blue lines), and for $N_0 - 1$ with occupied second level (dashed green). All shown chemical potentials are close to each other (see zoom into marked region in the inset). (b) Fermi momentum $k_F l_{so} = (2 + \mu/E_{so} + \sqrt{(E_z/E_{so})^2 + 4 + 4\mu/E_{so}})^{1/2}$ as a function of the Zeeman field using the self-consistent chemical potential for N_0 electrons and empty levels. The dotted vertical line shows the Zeeman energy at which the transition to the topological regime occurs. Shortly after this transition, fixing N_0 also k_F becomes constant as a function of E_z . (c) Lowest energy level of the wire Hamiltonian as a function of both Zeeman field E_z and chemical potential μ . The red line is the self-consistent chemical potential for $N_0 = 20 \cdot 35$ electrons. After entering the topological phase, this chemical potential stays orthogonal to the finite size oscillations of the Majorana energy such that \mathcal{E}_1 stays approximately constant.

From Fig. 2.30(b) it also becomes apparent that fixing the particle number while changing the external field fixes the Fermi momentum k_F in the Majorana wire, Eq. (2.140), in the topological regime (for $E_z > \sqrt{\Delta^2 - \mu^2}$ indicated by the vertical line). In addition, μ follows the energy of the first Bogoliubov level parallel to the finite size oscillations (Fig. 2.30(c)) such that \mathcal{E}_1 is approximately constant as a function of the Zeeman field in the topological regime. Therefore, the finite size oscillations of the Majorana energy described in Sec. 2.2.3 are not visible. This may also allow for obtaining an approximation for $\mu(N_0)$ from the energy of the first level only, which would be much easier than obtaining the full set of eigenstates.

Excited dot states in the thermal average

We allow the dot to be in excited states $\{n_i\}$ with N_0 electrons when the elastic co-tunneling process occurs and perform a thermal average when computing the scattering matrix in Eq. (2.222). Because the condensate is made of Cooper pairs, it can only hold an even number of electrons. Therefore, the number parity of N_0 determines the number parity of the allowed occupation numbers, i.e. $\sum_i n_i \bmod 2 = N_0 \bmod 2$. Hence, for even N_0 , in the ground state, all electrons are in the condensate, $\{n_i = 0\}$, and for odd N_0 one electron has to occupy the first level,

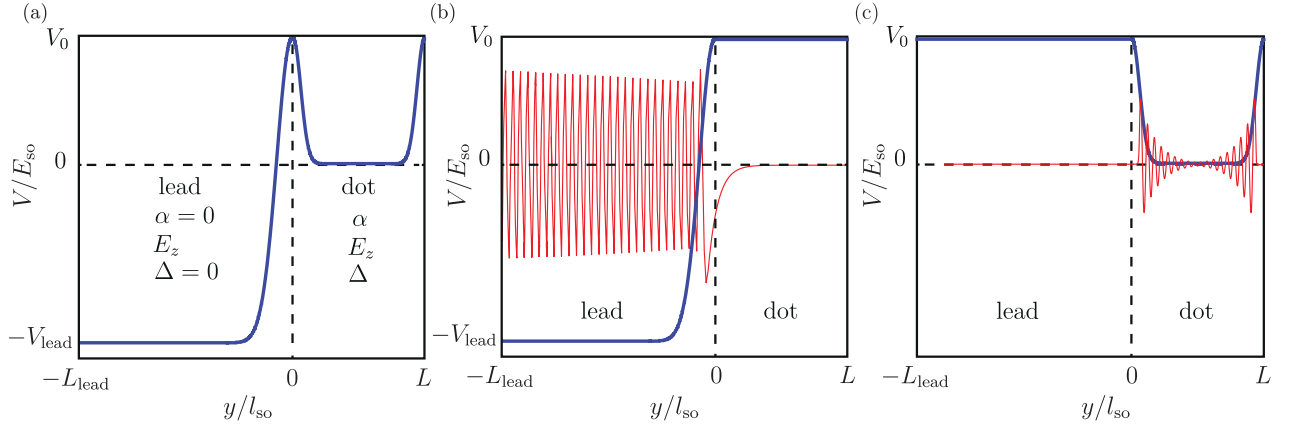


Figure 2.31: Microscopic model for the dot-lead couplings. (a) The confinement potential Eq. (2.270) separating the leads from the dot, where the lead-dot interface is marked by a vertical, black line. We consider normal-conducting leads of length $L_{lead} = 130 l_{so}$ without spin-orbit coupling, and a wire of length L . The potential in the leads is lowered by $V_{lead} = 100 E_{so}$ such that both spin species are present at the Fermi level. (b) Sketch of a wave function at the Fermi level, localized in the left lead (red) and auxiliary potential used to obtain the wave function (blue). (c) Majorana wave function in the wire (red) and auxiliary potential (blue).

$\{n_{i>1} = 0, n_1 = 1\}$. The excitation energy for occupation numbers $\{n_j\}$ is then given by

$$E_{ex}(N_0 \in 2\mathbb{N}, \{n_j\}) = \sum_j n_j \mathcal{E}_j \quad (2.268)$$

$$E_{ex}(N_0 \in 2\mathbb{N} + 1, \{n_j\}) = -\mathcal{E}_1 + \sum_j n_j \mathcal{E}_j, \quad (2.269)$$

such that the corresponding Boltzmann weight for the initial state with N_0 electrons and occupation numbers $\{n_j\}$ in the thermal average is $\exp(-\beta E_{ex}(N_0, \{n_j\}))$ with inverse temperature $\beta = (k_B T)^{-1}$. For small temperatures (we later use $\beta = 18 E_{so}$) the Boltzmann factors make states involving higher levels negligible such that we do not include excited states with any $n_{j>10} = 1$ and only consider initial states with up to four occupied levels, i.e. $\sum_j n_j \leq 4$.

Before we show numerical results for transport through Majorana wires obtained with the scattering matrix formalism, we discuss the microscopic models for the dot-lead couplings and explain their relations to the simpler model discussed in this section.

2.3.4 Microscopic model for dot-lead couplings

So far, we considered dot lead couplings that were defined as overlaps of an exponentially decaying phenomenological lead wave function and the dot eigenfunctions inside the tunneling barrier. In this section, we extend this idea by microscopically modeling parts of the lead of length L_{lead} and the transition region between lead and dot with a confinement potential (see Fig. 2.31) of Gaussian shape $V_{\sigma_i, V_0}(y) = V_0 \exp[-y^2/(2\sigma_i^2)]$ defined as [131]

$$V(y) = \begin{cases} V_{\sigma_1, V_0+V_{lead}}(y - L_{lead}) - V_{lead} & y \leq L_{lead} \\ V_{\sigma_1, V_0}(y - L_{lead}) & L_{lead} < y < y_1 \\ V_{\sigma_2, V_0}(y - y_1 + y_2 - L_{lead}) & y \geq y_1 \end{cases} \quad (2.270)$$

$$y_j = \sqrt{2\sigma_j^2 \ln(V_0/E_s)}. \quad (2.271)$$

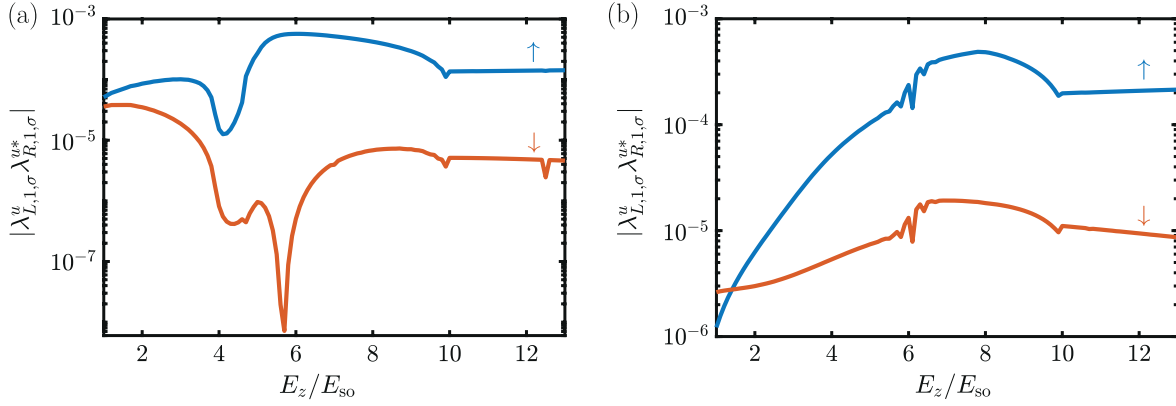


Figure 2.32: Dot-lead couplings in the microscopic model for a Majorana wire of length $L = 45.5 l_{so}$, Zeeman field dependent gap $\Delta(E_z) = \Delta(0)\sqrt{1 - (E_z/E_{z,c})^2}$, with critical Zeeman energy $E_{z,c} = 10 E_{so}$, $\Delta(4.5 E_{so}) = 2 E_{so}$, and self-consistent chemical potential to fix the average particle number. (a) For a steep confinement $\sigma_1 = \sigma_2 = 0.1 l_{so}$, $E_s = 10 E_{so}$, and wire particle number $N_w = 47L/(39 l_{so})$ and (b) for a smooth confinement with $\sigma_1 = 0.1 l_{so}$, $\sigma_2 = 6 l_{so}$, $E_s = 10 E_{so}$, and $N_w = 53L/(39 l_{so})$ (adapted from [131]).

We lower the potential in the lead by $V_{lead} = 100 E_{so}$, to ensure that both spin species are preset at the Fermi level in the leads. The confinement potential initially decays with width σ_1 and transitions smoothly into a decay with width σ_2 at (y_2, E_s) [131]. This allows us to consider both steep confinements, e.g. for $\sigma_1 = \sigma_2 = 0.1 l_{so}$, and shallow confinements, e.g. $\sigma_1 = 0.1 E_{so}$, $\sigma_2 = 6 l_{so}$, $E_s = 10 E_{so}$, which yields ABS in the trivial phase (see Sec. 2.2.4).

The main improvement of the microscopic model in comparison to the simpler model of the previous section is that the spin dependence for tunneling through the barrier in presence of a Zeeman field can be taken into account. To obtain the couplings, we first discretize the Hamiltonian H consisting of lead and dot sectors, as well as a standard hopping $-E_{so}(l_{so}/a)^2 \sigma_0$ at the lead dot interface between both sectors, and compute couplings as overlaps [131]

$$\begin{aligned} \lambda_{i\alpha\sigma}^u &= \langle \Phi_{\alpha\sigma}^u | H | \Psi_i \rangle \\ \lambda_{i\alpha\sigma}^v &= \langle \Phi_{\alpha\sigma}^v | H | \Psi_i \rangle. \end{aligned} \quad (2.272)$$

Here Ψ_i is the i -th BdG level in the wire, the wave function in the normal-conducting leads without spin-orbit coupling have the form $\Phi_{\alpha\sigma}^u = (\varphi_{\alpha\uparrow}^{(\varepsilon_F, \sigma)}, \varphi_{\alpha\downarrow}^{(\varepsilon_F, \sigma)}, 0, 0)$ and $\Phi_{\alpha\sigma}^v = (0, 0, \varphi_{\alpha\downarrow}^{(\varepsilon_F, \sigma)*}, -\varphi_{\alpha\uparrow}^{(\varepsilon_F, \sigma)*})$ where the wave function closest to the Fermi level localized in lead α with spin σ is given by $(\varphi_{\alpha\uparrow}^{(\varepsilon_F, \sigma)}, \varphi_{\alpha\downarrow}^{(\varepsilon_F, \sigma)})$ [131]. To numerically compute the wave function localized in one region, we continue the confinement potential at height V_0 in the remaining regions (Fig. 2.31b and c) [131], which is justified if the considered wave functions decay fully in the potential barrier.

Due to the direction of the Zeeman field that prefers the spin- \uparrow direction, tunneling of \downarrow -electrons through the barrier is also greatly suppressed [124]. In Fig. 2.32, we show the product of couplings $|\lambda_{1L\sigma}^u \lambda_{1R\sigma}^{u*}|$, which determines the dominant electron-like term in the scattering matrix, for spin- \uparrow (blue) and spin- \downarrow (red) electrons as a function of the external Zeeman field. In both cases steep confinement (panel a) and smooth confinement (panel b), we see that for a large enough Zeeman field $E_z \gtrsim 2 E_{so}$, coupling of \uparrow -electrons is stronger – even more than an order of magnitude in the topological phase (between the dashed vertical lines).

For the more paradigmatic model of the couplings from the previous section, where this difference in the couplings of different spin is not reflected in the effective lead wave function Eq. (2.242), we will therefore approximate $|\langle T_{\uparrow\uparrow} + T_{\downarrow\downarrow} \rangle| \approx |\langle T_{\uparrow\uparrow} \rangle|$. We then use the microscopic model to validate the results of the easier model.

2.3.5 Extracting Weidenmüller formula parameters from KWANT

For the second microscopic model, we use the python package KWANT [389], which allows us to define the system via hoppings and on-site terms, to connect translationally invariant leads to the system, to extract system quantities from the model (Fig. 2.33), and even to compute transport properties like scattering matrices. Unfortunately, KWANT does not support Coulomb-blockade. We can still use it to set up the Hamiltonian H_w , integrate out leads, extract the corresponding self-energies Σ_α , and compute propagating modes of the leads φ_α at the dot-lead interfaces. In the absence of Coulomb-blockade, the block of the scattering matrix for transmission from left to right lead can be obtained via [413]

$$T_{LR} = i\varphi_R^\dagger \Gamma_R G_w^r \Gamma_L \varphi_L, \quad (2.273)$$

with $\Gamma_\alpha = i(\Sigma_\alpha - \Sigma_\alpha^\dagger)$, $\Sigma = \Sigma_L + \Sigma_R$ and propagator

$$G_w^r = [\varepsilon \mathbb{1} - \mathcal{H}_w - \Sigma]^{-1}. \quad (2.274)$$

This is a numerical algorithm called wave function matching [413] used in KWANT for computing the scattering matrix [389] and can be brought into the form of a Weidenmüller formula [414]. In order to introduce Coulomb-blockade in the truncated Hilbert space picture, we first extract the wire Hamiltonian matrix \mathcal{H}_w using KWANT and diagonalize it which yields the eigenstates U_w and the energy eigenvalues \mathcal{E}_i . Inserting representations of the identity $\mathbb{1} = U_w^\dagger U_w = U_w U_w^\dagger$ into Eq. (2.273), we find the Weidenmüller formula

$$T_{LR} = \langle i\varphi_R^\dagger \Gamma_R U_w U_w^\dagger G_w^r U_w U_w^\dagger \Gamma_L \varphi_L \rangle \quad (2.275)$$

$$= \left\langle iW_R \frac{1}{\varepsilon \mathbb{1} - \text{diag}(\mathcal{E}_i) - U_w^\dagger \Sigma U_w} W_L^\dagger \right\rangle, \quad (2.276)$$

where the dot-lead couplings are $W^\dagger = (W_L, W_R)^\dagger$ with

$$W_L^\dagger = U_w^\dagger \Gamma_L \varphi_L \quad (2.277)$$

$$W_R = \varphi_R^\dagger \Gamma_R U_w. \quad (2.278)$$

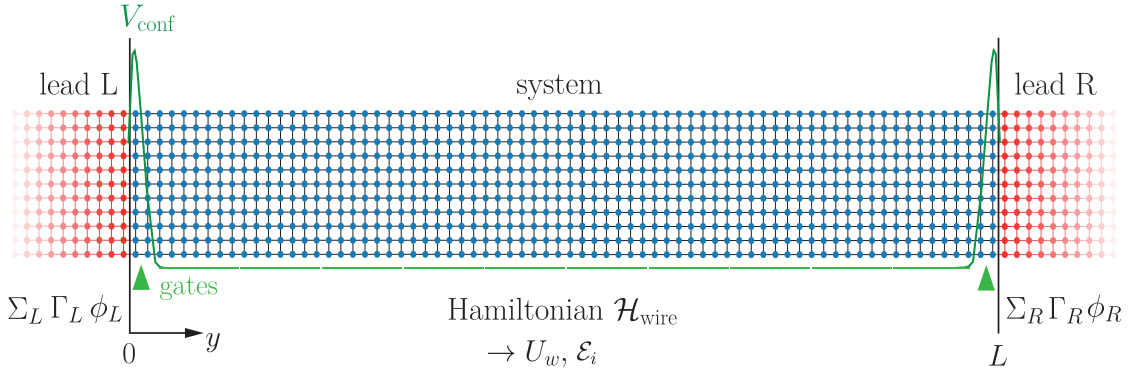


Figure 2.33: Schematic of a two-dimensional Majorana wire and leads modeled with KWANT [389]. Translationally invariant leads (red) and system, i.e. the Majorana wire (blue), are separated by a confinement potential induced by pinch-off gates (green). The python package KWANT allows us to define the setup in terms of onsite energies and hoppings, and computes lead self energies Σ_α , Γ_α , propagating lead modes ϕ_α , and the wire Hamiltonian matrix \mathcal{H}_w from which we extract eigenstates U_w and energy eigenvalues \mathcal{E}_i (taken from the supplement of the publication [303]).

This is the same as the expression inside the thermal average in Eq. (2.222) by replacing $U_w^\dagger \Sigma U_w \rightarrow \frac{-i}{2} W^\dagger W = U_w^\dagger i \text{Im}(\Sigma) U_w$. This can be done, as the missing real part $\text{Re}(\Sigma)$ has the form of a potential at the dot-lead interfaces (the only positions where Σ is non-zero) in Eq. (2.274) and can therefore be absorbed into the confinement potential due to the pinch-off gates. In order to add Coulomb-blockade to the Weidenmüller formula, we know from the previous section (Eqs. (2.258)-(2.265)) how to adjust the energy levels and where we need to swap electron-like u_j and hole-like wave functions v_j to determine the couplings. For the chemical potential in the wire, we use $\mu(N_0)$ self-consistently determined to have N_0 electrons in the dot. As we argued before (Sec. 2.3), this is a good approximation as $\mu(N_0) \approx \mu(N_0 \pm 1)$.

As the self-energies Σ_α and hence Γ_α are only non-zero at the dot-lead interfaces $y = y_\alpha$, the couplings $\varphi_\alpha^\dagger \Gamma_\alpha U_w$ are proportional to the wire wave functions evaluated at y_α which provides a justification for the choice in the more paradigmatic model Eq. (2.242). The advantage of using KWANT is that it correctly covers the band structure in the semi-infinite, translationally invariant leads and accurately treats dot-lead couplings via self-energies after integrating out the leads. In addition, KWANT simplifies defining two-dimensional systems even with more complex geometries by only requiring a boolean *shape* function that takes a position in space and returns whether this position is in the system, a function which returns onsite energies for each given position, and a function that takes two positions and returns the hopping energy between the points in space. Based on this information, KWANT can set up the grid, attach leads, and return the wire Hamiltonian and the dot-lead interfaces automatically.

In the following, we will first use the paradigmatic model (Sec. 2.3.3) and the first microscopic model (Sec. 2.3.4) to describe the Zeeman field and length dependence of the transmission amplitude in Sec. 2.4, and later the KWANT model to perform machine learning optimizations for Majorana wires in the presence of disorder in Sec. 2.5.

2.4 Coherent transmission amplitude through a Coulomb-blockaded nanowire

Major parts of this section closely follow the publication: Matthias Thamm and Bernd Rosenow, *Transmission amplitude through a Coulomb blockaded Majorana wire*, Physical Review Research 3, 023221 (2021) [131].

In this section, we study the amplitude of coherent transmission from which one can derive signatures for the presence of MZMs that are based on the non-locality of the Majorana wave functions and therefore hard to mimic by disorder or trivial zero modes such as ABSs. Such a signature has been studied in a recent experiment by Whiticar et al. [130], where they consider coherent transport through an interferometer with a Majorana wire embedded into one of the arms as a function of an external magnetic field. They find that for small Zeeman fields, the amplitude of coherent transmission is small, for larger fields – which they attribute to the topological regime – an amplitude maximum is observed, and after superconductivity is destroyed by a large field, the amplitude is small again. We use the scattering matrix formalism of Sec. 2.3 to study this transmission amplitude, i.e, the oscillation amplitude of the current as a function of the flux through an Aharonov-Bohm interferometer where the Majorana wire is in the Coulomb-blockade regime (Fig. 2.34). In this setup, we know from the discussion of the Weidenmüller formula for the scattering matrix Eq. (2.222) that the contribution of tunneling through a single dominant level of the Majorana wire with wave function $\varphi(y)$ to the transmission amplitude is roughly given by

$$T \sim \frac{\varphi(y_L)\varphi(y_R)}{E_c}, \quad (2.279)$$

where E_c is the large charging energy, and y_α is the positions of the dot-lead interface with lead α . Hence, for a MZM, where the wave function is exponentially localized at both ends with a correlation length ξ , the weight of the wave function at the wire ends scales like $\varphi(y_\alpha) \sim 1/\sqrt{\xi}$ and thus the transmission amplitude is approximately

$$T_{\text{MZM}} \sim (E_c \xi)^{-1}. \quad (2.280)$$

In contrast, for tunneling through a single extended state with $\varphi(y_\alpha) \sim 1/\sqrt{L}$, the transmission amplitude scales as

$$T_{\text{triv}} \sim (E_c L)^{-1}. \quad (2.281)$$

Eq. (2.279) implies that the transmission amplitude is large for well localized MZMs and small in the case of a trivial state. In addition, the occurrence of the maximum at the onset of the topological phase can already be qualitatively understood in this simple picture: When increasing the Zeeman field to bring the system into the topological phase, ξ is divergent at the topological phase transition such that the amplitude is small. Then MZMs become sharply localized resulting in a large transmission amplitude, and when further increasing the Zeeman field, the size of the effective p-wave gap $\Delta_{\text{ind},p}$ decreases such the MZMs become less localized ($\xi \propto \hbar v_F / \Delta_{\text{ind},p}$), and the transmission amplitude decreases again.

In addition, the charging energy $E_c \propto L^{-1}$ decreases with the system size L such that the amplitude for transport through MZMs is approximately linear in the wire length $T_{\text{MZM}} \propto L$, while it is independent of L for the trivial extended state $T_{\text{triv}} \sim 1$. Hence, in addition to

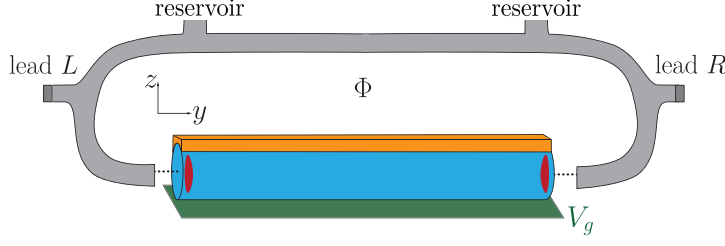


Figure 2.34: Schematic sketch of the Majorana interferometer setup. The lower arm contains the quantum dot consisting of a one dimensional Rashba wire (blue), superconductor (orange), and gate (green). The upper arm of the interferometer is the reference arm which only contains a wire. By varying a flux Φ , the transmission amplitude of electrons tunneling through the dot as a function of the gate voltage can be observed [304]. When the wire is tuned to the topological regime by an external Zeeman field, Majorana zero modes (red) are present at the ends. A reservoir is needed to avoid the phase rigidity effect [390].

the magnetic field dependence, the wire length dependence of the transmission amplitude also provides a signature for MZMs based on the non-locality.

Furthermore, it is also interesting to study the coherent transmission amplitude at finite temperatures, where a ground state degeneracy due to a pair of zero energy ABSs yields two contributions with equal magnitude but opposite sign to the thermal average [132], which indicates that the transmission amplitude is able to distinguish MZMs from ABSs. To confirm these qualitative statements, we combine numerical and analytical results in the following.

2.4.1 Magnetic field and wire length dependence of the transmission amplitude

We first study the Zeeman field dependence of the transmission amplitude using the effective model for the dot-lead couplings of Sec. 2.3.3. We begin with a Majorana wire of length $L = 32.5 l_{so}$ with a Zeeman field independent proximity s-wave gap Δ to understand how the effective p-wave gap influences the transmission amplitude. In the one-dimensional case, by increasing the strength of the Zeeman field, the system transitions into the topological regime and remains topological for all higher Zeeman fields. However, the effective p-wave gap, $\Delta_{ind,p} \sim E_z^{-1}$, decreases with the Zeeman energy such that the localization of the MZMs and hence the transmission amplitude is expected to decrease the deeper the system is in the topological phase.

In a next step, we then allow the magnetic field to destroy the induced s-wave proximity gap $\Delta(E_z)$ such that the system becomes trivial again for fields above a critical field strength. This more realistic situation allows for a comparison to be made with the experiment by Whiticar et al. [130].

We then further study the wire length dependence of the transmission amplitude where, according to the discussion above, we expect the transmission amplitude due to tunneling through MZMs to increase proportionally to the wire length.

Magnetic field independent induced gap

We study the transmission amplitude $|\langle T_{\uparrow\uparrow}(V_{g,mid}) \rangle|$ as a function of Zeeman energy E_z , computed at a gate voltage $V_{g,mid}$ in the middle between the two conductance resonances for a fixed

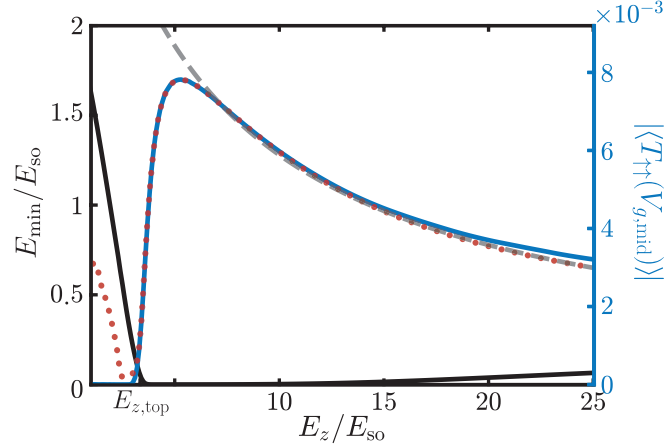


Figure 2.35: Transmission amplitude $|\langle T_{\uparrow\uparrow}(V_{g,\text{mid}}) \rangle|$ (blue y-axis) for transmission through the lowest effective level only (dotted, red), and $j_{\text{max}} = 200$ effective levels (solid, blue), together with the lowest BdG energy (black y-axis, solid black line) of the wire Hamiltonian as a function of the magnetic field. Here, $V_{g,\text{mid}}$ is the center between amplitude resonances corresponding to a particle number $N_w = 35$. The dashed gray line shows the decay of the amplitude according to Eq. (2.284), where the constant factor is obtained from a fit. We use $\Delta = 2 E_{\text{so}}$ and $L = 32.5 l_{\text{so}}$.

particle number $N_w = (L/l_{\text{so}})(14/13)$, such that the particle density is the same for all wire lengths.

We first consider a magnetic field independent proximity gap $\Delta = 2 E_{\text{so}}$. By increasing E_z , the transition to the topological phase takes place, in which an eigenstate close to zero energy is formed, separated from the second level by the topological gap (see Fig. 2.35). The transmission amplitude strongly increases when entering the topological phase at $E_{z,\text{top}}$, reaches a peak value, and then decreases. In the topological regime the tunneling matrix element for a Majorana wave function is $\propto 1/\sqrt{\xi}$, where the correlation length

$$\xi = \frac{\hbar v_F}{\Delta_{p,\text{ind}}} \quad (2.282)$$

with $v_F = \hbar k_F(1/m^* - \alpha_R^2(E_z^2 + \hbar^2 \alpha_R^2 k_F^2)^{-1/2})$ is determined by the induced effective p-wave gap at the Fermi points in the hybrid wire [86, 87, 366]

$$\Delta_{p,\text{ind}} = \frac{\hbar k_F \alpha_R \Delta}{\sqrt{E_z^2 + \alpha_R^2 \hbar^2 k_F^2}}. \quad (2.283)$$

With this, we obtain the Zeeman field dependence of the transmission amplitude as

$$|T_{\uparrow\uparrow}| \sim \frac{m^* \alpha_R \Delta}{\hbar} \frac{1}{\sqrt{E_z^2 + \alpha_R^2 \hbar^2 k_F^2 - \alpha_R^2 m^*}}, \quad (2.284)$$

proportional to the inverse field strength for large E_z (dashed gray line in Fig. 2.35, in very good agreement with the numerical result taking a single level into account). When comparing the result for transmission through $j_{\text{max}} = 200$ levels (solid blue line) with that for a single level (Fig. 2.35, dotted red line) it becomes apparent that the amplitude at the beginning of

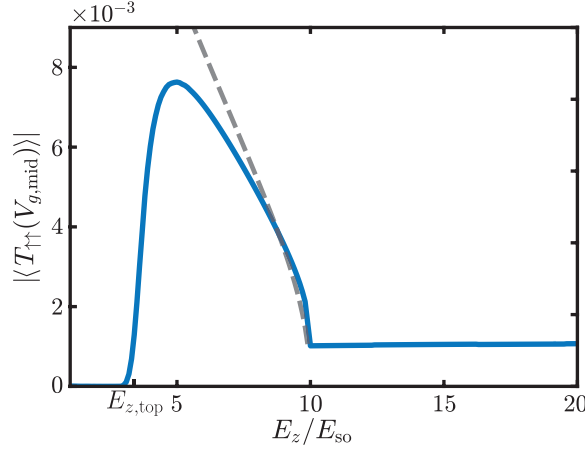


Figure 2.36: Transmission amplitude $|\langle T_{\uparrow\uparrow}(V_{g,\text{mid}}) \rangle|$ for $j_{\text{max}} = 200$ effective levels as a function of the magnetic field. Here, $V_{g,\text{mid}}$ is the center between amplitude resonances corresponding to a particle number $N_w = 35$. We use a wire length of $L = 32.5 l_{\text{so}}$ and a field dependent induced gap, Eq. (2.285), where $\Delta(4.5 E_{\text{so}}) = 2 E_{\text{so}}$ and $E_{z,c} = 10 E_{\text{so}}$ (solid, blue). The dashed gray line shows the decay of the amplitude according to Eq. (2.286), where the constant factor is obtained from a fit.

the topological range is mostly determined by the lowest level, i.e. the MZMs. For very large Zeeman energy, the Majorana modes are split more strongly, and there is a small correction due to taking into account many higher levels. In the trivial regime for $E_z < E_{z,\text{top}}$ however, where the spacing between the lowest energy Bogoliubons is small, many levels contribute to the transmission amplitude, and interfere destructively.

Magnetic field dependent induced gap

For a thin superconductor subject to a parallel field, we describe the suppression of the induced s-wave superconducting gap by the magnetic field via [412]

$$\Delta(E_z) = \Delta(0) \left[1 - \left(\frac{E_z}{E_{z,c}} \right)^2 \right]^{1/2}, \quad (2.285)$$

where $E_{z,c}$ is the critical Zeeman energy at which superconductivity is destroyed. Entering the topological region at $E_{z,\text{top}}$ is again accompanied by an increase in transmission amplitude (see Fig. 2.36). Further, within the topological regime, the proximity gap Δ is reduced, and the correlation length $\xi \propto 1/|\Delta_{p,\text{ind}}|$ increases, i.e. the Majorana wave function delocalizes. Therefore, the amplitude drops to the normal-conducting value over a relatively narrow range of magnetic field values. Using Eq. (2.285) in Eq. (2.283), we find an amplitude dependence

$$|T_{\uparrow\uparrow}| \sim \frac{m^* \alpha_R \Delta(0)}{\hbar} \frac{\sqrt{1 - (E_z/E_{z,c})^2}}{\sqrt{E_z^2 + \alpha_R^2 \hbar^2 k_F^2 - \alpha_R^2 m^*}}. \quad (2.286)$$

This dependence is depicted by the dashed gray line and fits well in the region where the amplitude decays to the normal-conducting value (see Fig. 2.36). For $E_z > E_{z,c}$, the wire is normal-conducting, and the amplitude is approximately constant. These results for the amplitude are in agreement with the recent experiment [130].

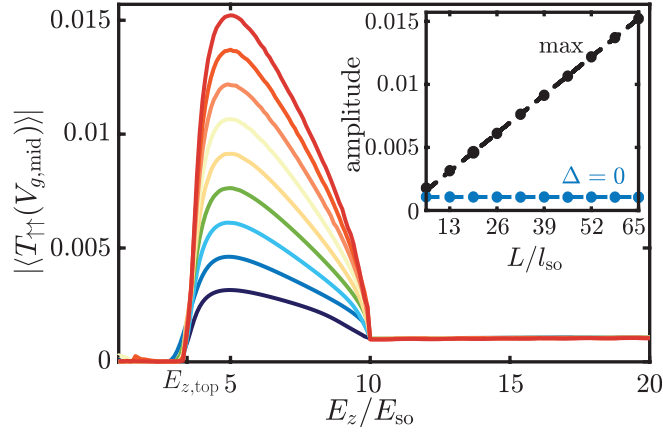


Figure 2.37: Transmission amplitude $|\langle T_{\uparrow\uparrow}(V_{g,\text{mid}}) \rangle|$ as a function of the magnetic field for different wire lengths $L = 13 l_{\text{so}}, 19.5 l_{\text{so}}, 26 l_{\text{so}}, 32.5 l_{\text{so}}, 39 l_{\text{so}}, 45.5 l_{\text{so}}, 52 l_{\text{so}}, 58.5 l_{\text{so}}$ and $65 l_{\text{so}}$. Here, $V_{g,\text{mid}}$ is the center between amplitude resonances for a particle number $N_w = 35L/(32.5 l_{\text{so}})$. We assume a Zeeman field dependent gap parameter Eq. (2.285) with $\Delta(4.5 E_{\text{so}}) = 2 E_{\text{so}}$ and a critical field $E_{z,c} = 10 E_{\text{so}}$. The inset shows the value at the maximum of the amplitude in the topological region (black circles) and the value of the amplitude in the normal-conducting region (blue circles) as a function of the wire length L .

Wire length dependence

The non-locality of MZMs is expected to have a profound consequence when considering wires of varying lengths. In the inset of Fig. 2.37, the value of the amplitudes at the maximum and in the normal-conducting region are depicted as a function of the wire length L . From our scattering matrix analysis, using a charging energy that is proportional to the inverse of the wire length, we find that the transmission amplitude is indeed proportional to the wire length in the topological region, while it is independent of the wire length in the normal-conducting range (see Fig. 2.37).

2.4.2 Disorder in the wire

The proposed experiment for establishing the wire length dependence of the transmission amplitude in the presence or absence of MZMs requires the comparison of different wires. Since these wires may differ from each other in terms of their detailed composition, we study how robust our results for the transmission amplitude are in the presence of on-site disorder. We use a Gaussian disorder distribution with zero mean and standard deviation σ_{dis} . Disorder is strong when the elastic scattering rate \hbar/τ from the impurities is on the order of the induced effective gap $\Delta_{p,\text{ind}}$ in the wire [415–422]. We define a critical disorder strength σ_c such that the effect of disorder on the amplitude is negligible for $\sigma_{\text{dis}} \ll \sigma_c$. For $\sigma_{\text{dis}} \approx \sigma_c$ disorder has noticeable effects on the amplitude and for $\sigma_{\text{dis}} \gg \sigma_c$ pair breaking sets in and destroys the superconducting properties and the amplitude vanishes. Using Fermi's golden rule, we estimate the elastic scattering rate for the case of a scatterer at each lattice site as

$$\frac{\hbar/\tau}{E_{\text{so}}} = \left(\frac{\sigma_{\text{dis}}}{E_{\text{so}}} \right)^2 \frac{a}{l_{\text{so}}} \frac{1}{k_F l_{\text{so}}} . \quad (2.287)$$

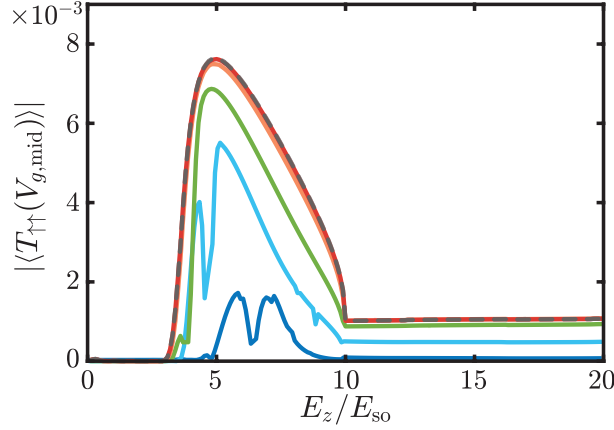


Figure 2.38: Transmission amplitude $|\langle T_{\uparrow\uparrow}(V_{g,\text{mid}}) \rangle|$ as a function of the magnetic field for various strengths of the on-site disorder. We use a wire length of $L = 32.5 l_{\text{so}}$, and compute the amplitude between resonances corresponding to particle number $N_w = 35$. The gray dashed line is for reference without disorder. The colored lines (from top to bottom) are numerically computed for Gaussian disorder with standard deviation $\sigma_{\text{dis}} = 0.1 E_{\text{so}} \ll \sigma_c$, $\sigma_{\text{dis}} = 1 E_{\text{so}}$, $\sigma_{\text{dis}} = 5 E_{\text{so}}$, $\sigma_{\text{dis}} = 10 E_{\text{so}} \sim \sigma_c$, and $\sigma_{\text{dis}} = 20 E_{\text{so}}$.

The induced gap at the Fermi momentum $k_F l_{\text{so}} = (2 + \mu/E_{\text{so}} + [(E_z/E_{\text{so}})^2 + 4 + 4\mu/E_{\text{so}}]^{1/2})^{1/2}$ is given by

$$\frac{\Delta_{p,\text{ind}}}{E_{\text{so}}} = 2 \frac{\Delta}{E_{\text{so}}} \frac{k_F l_{\text{so}}}{\sqrt{(E_z/E_{\text{so}})^2 + 4(k_F l_{\text{so}})^2}}. \quad (2.288)$$

We define σ_c such that $\hbar/\tau = \Delta_{p,\text{ind}}$ for $\sigma_{\text{dis}} = \sigma_c$, i.e.

$$\frac{\sigma_c}{E_{\text{so}}} = \sqrt{2 \frac{l_{\text{so}}}{a} \frac{\Delta}{E_{\text{so}}} \frac{k_F l_{\text{so}}}{[(E_z/E_{\text{so}})^2 + 4(k_F l_{\text{so}})^2]^{1/4}}}. \quad (2.289)$$

Numerical results of the amplitude for various disorder strengths are depicted in Fig. 2.38. When the disorder strength is smaller but of the order of σ_c , the transmission amplitude is reduced at its maximum. This reduction is however much smaller than the peak height such that the proposed experiment is robust against disorder $\sigma_{\text{dis}} < \sigma_c$. When using a disorder strength close to or larger than σ_c , the amplitude is significantly reduced.

2.4.3 Microscopic model and Andreev zero modes

With the effective model for the couplings considered so far, we are limited to the transmission amplitude of spin- \uparrow particles, which give the dominant contribution in the presence of a Zeeman field [124] (see also Fig. 2.32). To overcome this limitation, we study transport with the microscopic model for the dot-lead couplings as described in Sec. 2.3.4. For this, we explicitly model parts of the leads in addition to the wire which are separated by a confinement potential, Eq. (2.270), with Gaussian shape, height V_0 , initial decrease with standard deviation σ_1 , and final width σ_2 after the potential has decreased to height E_s . We then obtain the couplings from matrix elements of the tunneling Hamiltonian between states localized in lead α and in the dot, Eq. (2.272). This does not only allow us to compute the full transmission amplitude $|\langle T_{\uparrow\uparrow} + T_{\downarrow\downarrow} \rangle|$ but also to consider transport through topologically trivial ABSs for a shallow confinement potential with $\sigma_2 \gg \sigma_1$ as discussed in Sec. 2.2.4.

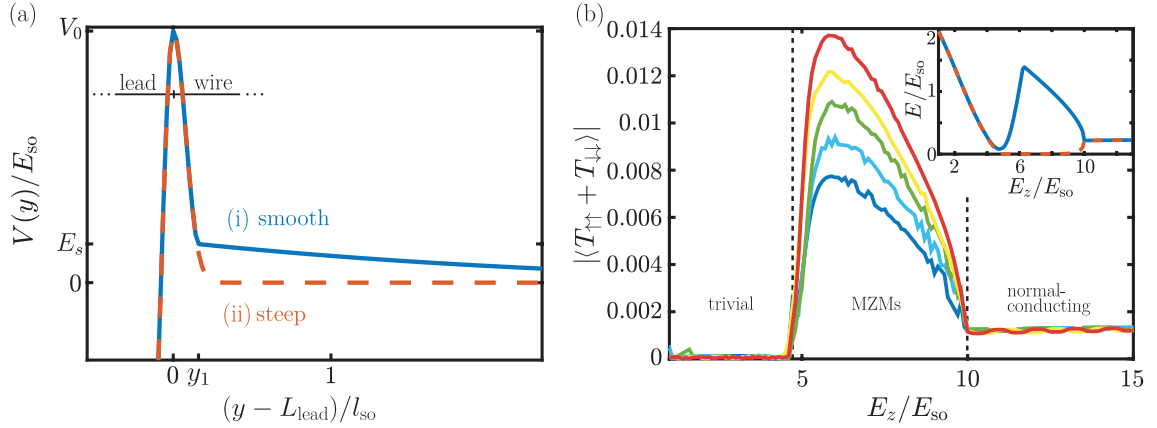


Figure 2.39: (a) Barrier potential used to compute the lead-wire couplings in the microscopic model. The leads of length L_{lead} are normal-conducting and without spin-orbit coupling. At position y_1 and energy E_s the narrow Gaussian peak transitions continuously into the wide peak in the case of the smooth potential. The height of the peak is given by $V_0 = 65 E_{\text{so}}$. (b) Numerical results for the amplitude $|\langle T_{\uparrow\uparrow}(V_{g,\text{mid}}) + T_{\downarrow\downarrow}(V_{g,\text{mid}}) \rangle|$ with microscopic couplings as a function of the Zeeman field using a steep confinement potential and ground state particle number $N_w = 47L/(39 l_{\text{so}})$. We consider wires of length $L = 32.5 l_{\text{so}}, 39 l_{\text{so}}, 45.5 l_{\text{so}}, 52 l_{\text{so}},$ and $58.5 l_{\text{so}}$. The results are in excellent agreement with Fig. 2.37 where we used the more paradigmatic model for the couplings.

We distinguish two types of barrier potentials (i) a single narrow Gaussian peak and (ii) a narrow Gaussian peak together with a potential decaying smoothly into the wire (see Fig. 2.39(a)). Parameters for (i) the steep potential are given by $\sigma_1 = \sigma_2 = 0.1 l_{\text{so}}$, $E_s = V_0$ and $V_0 = 65 E_{\text{so}}$, and for (ii) the smooth confinement has $\sigma_1 = 0.1 l_{\text{so}}$, $\sigma_2 = 6 l_{\text{so}}$, $E_s = 10 E_{\text{so}}$ and $V_0 = 65 E_{\text{so}}$. In case (i) there are zero-energy states only in the topological region, which are the MZMs (see inset of Fig. 2.39(b)). In case (ii), the Fourier decomposition of the smooth potential does not contain large momenta, so that in the trivial region each of the two bands contributes a pair of MZMs, which however are not coupled among each other by the potential [112, 113, 124]. Therefore, in addition to the MZMs in the topological region, two quasi-degenerate, quasi-zero energy Andreev bound states (also called pseudo-MZMs) occur in the trivial region for $5 E_{\text{so}} < E_z < 7.6 E_{\text{so}}$ (see inset of Fig. 2.39(b)). Since they are nearly degenerate, there are two ground states with equal Boltzmann weight in the thermal average. For even N_0 the degenerate ground states for $\mathcal{E}_1 = \mathcal{E}_2 = 0$ are states where either all N_0 electrons are in the condensate or $N_0 - 2$ electrons form the condensate and both pseudo-MZMs occupied. In the case of odd N_0 there are $N_0 - 1$ electrons in the condensate and either the first or the second pseudo-Majorana level is occupied. In both cases the thermally averaged amplitude is proportional to $\sum_{j=1}^2 (\lambda_{L,j,\uparrow}^u \lambda_{R,j,\uparrow}^{u*} + \lambda_{L,j,\uparrow}^v \lambda_{R,j,\uparrow}^{v*}) \approx 0$. The anti-unitary reflection symmetry $\tilde{\Pi} \varphi_j(y) = K \varphi_j(L - y)$ (where φ_j are eigenfunctions of H_{wire}) ensures that both terms are real and $\text{sgn}(\lambda_{L,j,\uparrow}^u \lambda_{R,j,\uparrow}^{u*}) = -\text{sgn}(\lambda_{L,j,\uparrow}^v \lambda_{R,j,\uparrow}^{v*})$ [304]. Due to the Majorana condition for zero energy states $|u_{j\sigma}| = |v_{j\sigma}|$, the terms cancel each other. Hence, the ground state degeneracy gives rise to a vanishing amplitude upon thermal averaging [132]: Forming a Cooper pair or occupying the two zero-energy pseudo-MZMs requires the same energy, but yields contributions with opposite signs and equal magnitude to the transmission amplitude.

In fact, the reflection symmetry is sufficient but not necessary for the relation $u_{j\sigma}(y_L) u_{j\sigma}^*(y_R) = -v_{j\sigma}^*(y_L) v_{j\sigma}(y_R)$. It can for example also be understood in a situation where the two ABS are

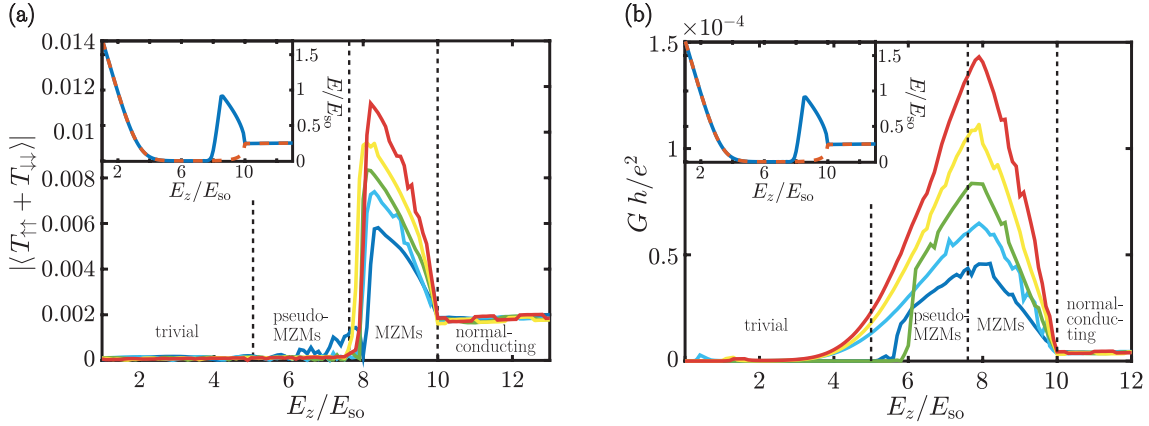


Figure 2.40: (a) Numerical results for the transmission amplitude $|\langle T_{\uparrow\uparrow}(V_{g,\text{mid}}) + T_{\downarrow\downarrow}(V_{g,\text{mid}}) \rangle|$ using tunnel couplings obtained for the smooth barrier potential as a function of the Zeeman energy. (b) Conductance through the wire without an interferometer in the case of a smooth potential for $N_w = 53L/(39l_{\text{so}})$. A comparison with (a) shows that the interferometer is crucial to distinguish MZMs from pseudo-MZMs. We consider wires of length $L = 32.5l_{\text{so}}, 39l_{\text{so}}, 45.5l_{\text{so}}, 52l_{\text{so}},$ and $58.5l_{\text{so}}$. We use a Zeeman field dependent induced gap Eq. (2.285) with a critical field $E_{z,c} = 10E_{\text{so}}$ and a particle number $N_w = 53L/(39l_{\text{so}})$. The insets depict the lowest two energy eigenvalues of the wire Hamiltonian for $L = 45.5l_{\text{so}}$.

eigenstates of the same spin operator, which for convenience we choose as the z -direction, such that only one spin component of the wave function is nonzero (allowing us to suppress the spin index in the following). For any Bogoliubon localized at both ends y_L, y_R of the wire, the quasiparticle operator is expressed in terms of the fermion operators $\psi(y), \psi^\dagger(y)$ as

$$\beta = \int dy \left[u(y)\psi(y) + v(y)\psi^\dagger(y) \right] \approx u(y_L)\psi(y_L) + u(y_R)\psi(y_R) + v(y_L)\psi^\dagger(y_L) + v(y_R)\psi^\dagger(y_R). \quad (2.290)$$

The anti-commutation relation $0 = \{\beta, \beta\}$ then ensures that

$$u(y_L)v(y_L) = -u(y_R)v(y_R), \quad (2.291)$$

and multiplying the equation with the complex conjugate of the left-hand side yields

$$|u(y_L)|^2|v(y_L)|^2 = -u^*(y_L)u(y_R)v(y_R)v^*(y_L). \quad (2.292)$$

As the left-hand side is real, we can parametrize

$$u^*(y_L)u(y_R) = +|u(y_L)u^*(y_R)|e^{-i\alpha} \quad (2.293)$$

$$v^*(y_L)v(y_R) = -|v^*(y_L)v(y_R)|e^{i\alpha}. \quad (2.294)$$

Then the relation for zero-energy states $|u(y)| = |v(y)|$ implies that indeed $u_{j\sigma}(y_L)u_{j\sigma}^*(y_R) = -v_{j\sigma}^*(y_L)v_{j\sigma}(y_R)$.

For a wire of finite length, the pseudo Majorana modes do not lie exactly at zero energy and a finite amplitude is observed. This is the case for the smallest wire length in Fig. 2.40(a). Here the terms in the thermal average with $(0, 0)$ and $(1, 1)$ occupation have Boltzmann weight 1 and $e^{-\beta(\mathcal{E}_1 + \mathcal{E}_2)}$ such that the amplitude is still suppressed by a factor $(1 - e^{-\beta(\mathcal{E}_1 + \mathcal{E}_2)})$. As long as the

two levels are nearly degenerate and nearly at zero energy, the amplitude is well below the $\Delta = 0$ value and no pronounced maximum is formed. In addition, the amplitude is not proportional to L in the pseudo-MZM regime.

In comparison with the more paradigmatic model considered before, we find that for a steep potential (i), all qualitative features of the amplitude remain unchanged (Fig. 2.39(b)). Importantly, the suppression of the transmission amplitude in the trivial regime occurs even when pseudo-MZMs are present (see Fig. 2.40(b)).

2.4.4 Comparison with direct conductance measurement

In this section, we compare signatures from the interferometer setup (Fig. 2.34), with an easier to implement direct conductance measurement through the dot, without interferometer. In the calculation of the transmission amplitude through the dot (interferometer case) or the transmission probability (direct conductance), the main difference is how the thermal average is performed. For the calculation of the amplitude of conductance oscillations through the interferometer, the thermal average is performed over the complex transmission amplitude (see Eq. (2.222)), so that the transmission phase contributes to such an average. In the case of a direct conductance measurement, the squared absolute value of the transmission amplitude is averaged, and the phase information does not contribute. Fig. 2.40(b) depicts the direct conductance through the dot for a smooth confinement potential, analogous to Fig. 2.40(a). For MZMs we also find a maximum at the beginning of the topological region, whose height scales with the wire length. The crucial difference is that the conductance is not suppressed for pseudo-MZMs and thus this maximum is not a unique signature for the presence of MZMs.

2.4.5 Comparison to approximate analytical solution and connection to experiment

In a recent experiment by Whiticar et al. [130], the transmission amplitude through a Coulomb blockaded Majorana wire was measured as a function of the Zeeman field. The experimental transmission amplitude shows a rapid growth upon entering the topological regime, followed by a pronounced maximum. Here, we discuss in detail how these features are explained by the localization properties of MZMs, which determine the transmission amplitude in the topological regime.

In the case of sufficiently long wires, in which the Majorana wave functions of opposite wire ends have negligible overlap, an analytical solution for the MZM wave functions can be found (see Sec. 2.2.2). Moreover, since the transmission amplitude in the topological region is determined almost exclusively by transport through MZMs, the transmission amplitude can directly be obtained from the Majorana wave functions. For large Zeeman fields, deep in the topological regime, the spatial decay of MZMs is characterized by the p-wave localization length $\xi = \hbar v_F / \Delta_{p,\text{ind}}$, Eq. (2.282). However, from the full analytic solution it becomes apparent that there is a second localization length

$$\begin{aligned} \xi_s &= \left(-\xi^{-1} + \sqrt{\xi^{-2} - \frac{\mu^2 + \Delta^2 - E_z^2}{(\xi^{-2} + k_F^2) E_{\text{so}}^2}} \right)^{-1} \\ &\propto \frac{1}{E_z - \sqrt{\Delta^2 + \mu^2}}, \end{aligned} \quad (2.295)$$

which describes the localization properties of MZMs for Zeeman fields $E_z \gtrsim E_{z,\text{top}}$ close to the topological phase transition. We can approximate the envelope of the Majorana wave function

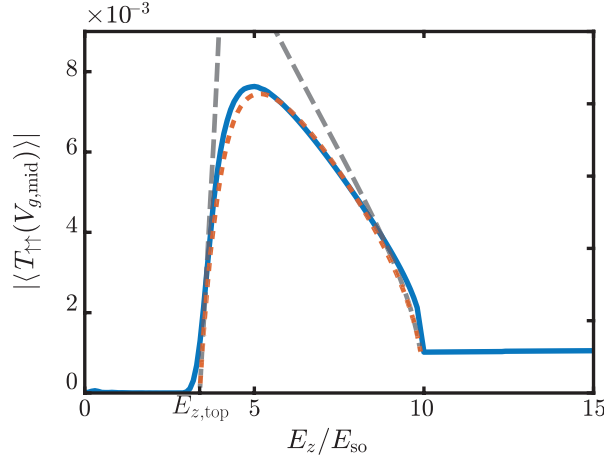


Figure 2.41: Transmission amplitude $|\langle T_{\uparrow\uparrow}(V_{g,\text{mid}}) \rangle|$ for $j_{\text{max}} = 200$ effective levels (solid, blue) as a function of the magnetic field with parameters as in Fig. 2.36. The dashed gray lines depict the ξ_s approximation at the beginning of the topological regime and the ξ approximation in the region where superconductivity is destroyed by the magnetic field. The dotted red line depicts the approximation Eq. (2.297), where both correlation lengths are taken into account. The maximum of the transmission amplitude arises due to the interplay of both terms. We use the same proportionality constant for all three approximations obtained by a fit.

by a sum of two exponentially decaying terms (for details see Sec. 2.2.2)

$$\chi_{L,\uparrow} \approx \frac{1}{\sqrt{\xi + \xi_s}} \left(e^{-y/\xi} + e^{-y/\xi_s} \right). \quad (2.296)$$

The corresponding Majorana wave function at the right end is then given by $\chi_{R,\uparrow}(y) \propto \chi_{L,\uparrow}(L - y)$. This yields for the transmission amplitude

$$|T_{\uparrow\uparrow}(V_{g,\text{mid}})| \propto \frac{1}{\xi + \xi_s}. \quad (2.297)$$

A comparison shows that the approximated transmission amplitude (dotted, red line in Fig. 2.41) is in very good agreement with the numerical results for transmission through 200 levels (solid, blue line in Fig. 2.41). Thus, the competition of the two correlation lengths ξ and ξ_s explains the occurrence of the maximum in the transmission amplitude.

In addition, in Fig. 2.41, we compare the numerically obtained transmission amplitude to approximations taking into account the larger of the two localization lengths: The behavior of the transmission amplitude at the beginning of the topological region can be understood by the localization length ξ_s alone, i.e. $|T_{\uparrow\uparrow}| \propto 1/\xi_s$ (dashed gray line in the beginning of the topological regime). On the other hand, the behavior near the transition into the normal-conducting region, is due to the p-wave localization length ξ , i.e. $|T_{\uparrow\uparrow}| \propto 1/\xi$ (dashed gray line at the end of the topological regime). The maximum occurs where the magnitude of the localization lengths is roughly comparable.

The picture described above allows to explain the magnetic field dependence of the transmission amplitude found by Whiticar et al. [130]. In the experiment, the transmission amplitude depends only weakly on the magnetic field in the region of small Zeeman fields, as predicted for the trivial phase. Above a device-specific value of the magnetic field, a rapid increase of the

transmission amplitude is observed, which can be explained by the magnetic field dependence of $1/\xi_s$ at the beginning of the topological phase. Due to the divergence of ξ_s at the phase transition $E_z = E_{z,\text{top}}$, the transmission amplitude increases linearly $|T_{\uparrow\uparrow}| \propto E_z - E_{z,\text{top}}$ in the topological regime. For larger Zeeman fields, a well-defined maximum of the amplitude arises in the experiment, which can be understood in terms of the concurrence of both correlation lengths ξ and ξ_s . When superconductivity is destroyed by the magnetic field, Whiticar et al. observe a rapid decline of the transmission amplitude. This decrease can be explained in our model by the divergence of the coherence length ξ due to the vanishing of the induced p-wave gap when approaching the critical magnetic field.

Since the amplitude of coherent transmission does not exhibit a maximum in the case of pseudo-MZMs, we believe that it is very likely that genuine topological MZMs were observed in the experiment. This is further supported by the observation that together with the appearance of the maximum also the even-odd splitting of the conductance resonances is suppressed. While the behavior of the transmission amplitude in the topological regime can be understood with our one-dimensional model, it is currently not possible to explain the large ratio between the value of the transmission amplitude at the maximum and the value in the normal-conducting regime for Device 2 measured by Whiticar et al. This could be because the amplitude in the experiment is not corrected for the influence of the transmission through the reference arm. On the other hand, it might be necessary to include the influence of orbital effects and several transverse subbands in the theoretical calculations for quantitative agreement between theory and experiment.

2.4.6 Conclusion

In this section, we studied the Zeeman field and wire length dependence of the amplitude of coherent transmission through Majorana wires in the Coulomb-blockade regime. We showed that the transmission amplitude has a maximum as a function of the Zeeman field at the beginning of the topological regime in agreement with a recent experiment [130]. This maximum is due to tunneling through MZMs: its height increases linearly with the wire length, and its occurrence can be understood from the concurrence between two correlation lengths present in the topological phase. These correlation lengths were found from the analytical MZM solution in agreement with the numerical simulations. In addition, we showed that Zeeman field and wire length dependence provide unique signatures for MZMs that allow us to distinguish them from trivial ABSs at finite temperatures [411].

Furthermore, we showed that these results are stable in presence of small disorder in the wire. In the following section, we consider strong disorder that is able to destroy MZMs and the topological phase.

2.5 Using machine learning to cancel disorder effects

Major parts of this section closely follow the publication: Matthias Thamm and Bernd Rosenow, *Machine learning optimization of Majorana hybrid nanowires*, arXiv preprint arXiv:2208.02182 (accepted for publication in Physical Review Letters) [303].

In recent years, increasingly complex quantum systems have been proposed, and their implementation has become tangible [56–60]. Usually, such systems require very complicated tuning of many parameters in order to function as intended [58, 64, 76, 423]. This can even become so complex that there is no way around automating this tuning [61–63], and machine learning algorithms in particular have proven to be exceptionally flexible and robust for this purpose [60, 62–67]. Here, we consider the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [242, 243, 246, 424] machine learning algorithm for adjusting the gate voltages of an array of gates in proximity to a strongly disordered Majorana wire. As discussed in the previous section, due to the presence of a topological gap, MZMs are quite stable against disorder. However, for the experimental realization of hybrid wires, different materials need to be contacted which makes it challenging, if not impossible, to keep disorder weak in such systems [102–105]. By inducing ABSs or other zero modes in the superconducting gap, strong disorder can mimic many signatures of MZMs in the trivial phase [102, 104, 108–111] or destroy the gap altogether [107]. This is why disorder still poses a major challenge in experiments [102–105]. In a recent study [106] implementing the so-called topological gap protocol [302] which is based on multi-terminal measurements, it was shown that due to disorder in 80% of Majorana devices no topological phase can be found.

We present a case study of automatic tuning of a gate array in proximity to a strongly disordered Majorana wire using the CMA-ES [242, 243] algorithm. CMA-ES is a machine learning algorithm that does not need system specific information to operate, and is widely applicable for high dimensional optimization problems [246, 249–252]. Crucially, a good metric allows improving desirable system properties during optimization. For example, signatures of MZMs can be mimicked by topologically trivial Andreev bound states (ABSs) [111–128, 376], which one would like to avoid. We therefore use the amplitude of coherent transmission [129–132] through a Coulomb-blocked Majorana wire as a metric, which has been measured by placing the wire in an arm of an electron interferometer [130], and which allows distinguishing MZMs from ABSs [131, 132]. We find that already 100 to some 1000 amplitude measurements are sufficient to tune the gate array, such that (i) both the localization of the MZMs and the transmission amplitude are significantly improved, and (ii) strong potential disorder is compensated. Experimental implementation of such automation could open the route to robust MZMs, for which scalable braiding protocols have been proposed [79, 293, 425] that are already based on gate arrays. Furthermore, we discuss important theoretical considerations on the implementation and performance of CMA-ES optimization, potentially relevant for many systems and situations.

2.5.1 Setup and CMA-ES algorithm

Fig. 2.42 depicts a schematic illustration of the CMA-ES [242, 243] algorithm. After being initialized by specifying an initial step size $\sigma^{(0)} = 0.1 E_{\text{so}}$, covariance matrix $C^{(0)} = 1$, evolution paths $\mathbf{s}_C^{(0)} = 0$ and $\mathbf{s}_\sigma^{(0)} = 0$, initial configuration $\mathbf{V}_g^{(0)} = 0$, and seed for the random generator, in each iteration t a generation of n_{pop} candidate solutions is drawn from a multivariate normal distribution $\mathcal{N}[\mathbf{V}_g^{(t)}, (\sigma^{(t)})^2 C^{(t)}]$. These candidates then have to be evaluated using the metric,

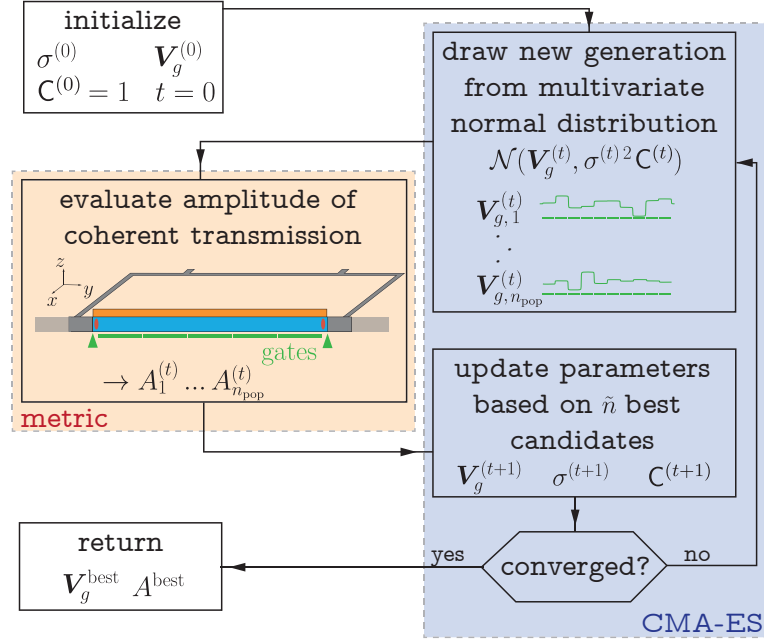


Figure 2.42: Schematic diagram of the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) algorithm [242, 243] that we use to learn an optimal gate array configuration that maximizes the amplitude of coherent transmission A through a Majorana wire. Initially, one sets a step size $\sigma^{(0)}$, a covariance matrix $C^{(0)}$, and a starting gate configuration $\mathbf{V}_g^{(0)}$. In each iteration t , n_{pop} gate voltage configurations are drawn from a multivariate normal distribution with mean $\mathbf{V}_g^{(t)}$ and covariance $C^{(t)}$. As a metric, we use the transmission amplitude for coherent transport through a Majorana hybrid wire consisting of a semiconductor (blue) with Rashba spin-orbit coupling and an s-wave superconductor (orange), placed in proximity to an array of gates (green) that allows tuning of a potential profile along the wire. By inserting the hybrid wire into one arm of an Aharonov-Bohm interferometer, the transmission amplitude can be measured [130–132, 304] and based on the magnitude of the amplitude A_i for the proposed gate voltage configurations $\mathbf{V}_{g,i}^{(t)}$, the parameters for the next iteration are adjusted. The new mean value $\mathbf{V}_g^{(t+1)}$ is determined from the \tilde{n} best candidates and step size $\sigma^{(t+1)}$ and covariance $C^{(t+1)}$ are updated based on the previous iterations. The goal of this machine learning algorithm is to contract the search region in which the randomly drawn values lie around the location of the optimum of the metric with subsequent iterations.

i.e. a measurement of the system with the suggested gate voltage configurations has to be performed, and the resulting values are passed back to the CMA-ES algorithm. Based on the \tilde{n} best candidates and the evolution of $C^{(t)}$ and $\mathbf{V}_g^{(t)}$, new parameters for drawing the subsequent generation are then determined (for details see Sec. 1.2.3 and App. A.5). If the optimization is successful, the search area in which the most candidates lie contracts around the optimum of the metric.

For the numerical computations, we use the Python implementation `pycma` [424] to optimize Fourier components a_k, b_k of an array of N_g gates of identical size with gate voltages

$$\tilde{V}_j = \frac{b_0}{2} + \sum_{k=1}^{\lfloor (N_g-1)/2 \rfloor} a_k \sin\left(\frac{2\pi}{N_g} kj\right) + \sum_{k=1}^{\lfloor N_g/2 \rfloor} b_k \cos\left(\frac{2\pi}{N_g} kj\right). \quad (2.298)$$

Optimization of Fourier components is more robust than directly optimizing the voltages: since a

Fourier component affects all voltages, the algorithm is not distracted by first lowering the gates at the ends which would slightly improve the amplitude, but corresponds to a local optimum only. In the following, we set the average gate voltage $b_0 = 0$, and consider the case of non-zero average gate voltage later.

To obtain the spatial potential profile $V_g(y)$ in the system, we assume that the wire is located a distance z_{sys} away from the gate surface and then approximate

$$V_g(y) = \mathcal{F}^{-1} \left[e^{-|q|z_{\text{sys}}} \mathcal{F} \left[\sum_{j=1}^{N_g} \tilde{V}_j \chi_j(y) \right] \right] , \quad (2.299)$$

where \mathcal{F} and \mathcal{F}^{-1} are Fourier transform and inverse Fourier transform in the variables y and q , respectively, and $\chi_j(y)$ is the characteristic function of gate j , i.e., $\chi_j(y)$ is 1 if y lies in gate j and 0 otherwise. For small z_{sys} , this produces a smooth potential profile that is approximately \tilde{V}_j over gate j .

To enhance desired features of a system during the optimization, the choice of an appropriate metric is crucial (see discussion later). We consider here as a case study the optimization of the coherent transmission amplitude $A = |\langle T_{\uparrow\uparrow} + T_{\downarrow\downarrow} \rangle|$ through a Majorana wire in Coulomb blockade, which can be measured by embedding the wire in the arm of an electron interferometer (panel titled “metric” in Fig. 2.42) and determining the amplitude of the current oscillations when changing the flux through the interferometer loop [130]. We first consider a strictly one-dimension model for the hybrid wire, and generalize to a more realistic two-dimensional model later. The Majorana wire consisting of a semiconductor with Rashba spin-orbit coupling α_R and a superconductor with s-wave gap Δ is described in Nambu basis $(d_{\uparrow}^{\dagger}(y), d_{\downarrow}^{\dagger}(y), d_{\downarrow}(y), -d_{\uparrow}(y))$ using the Hamiltonian

$$\begin{aligned} \mathcal{H}_{\text{wire}} = \tau_z \left[-\frac{\hbar^2 \partial_y^2}{2m^*} \sigma_0 - \mu \sigma_0 - i\hbar \alpha_R \sigma_x \partial_y + \delta_{\text{dis}}(y) \sigma_0 + V_g(y) \sigma_0 + V_{\text{conf}}(y) \sigma_0 \right] \\ - E_z \tau_0 \sigma_z + \Delta \tau_x \sigma_0 , \end{aligned} \quad (2.300)$$

with disorder potential δ_{dis} , confinement potential V_{conf} , gate potential V_g (see Eq. (2.299)), and Pauli matrices σ_i and τ_i acting in spin and particle-hole space, respectively. Rashba spin-orbit coupling defines a characteristic energy scale $E_{\text{so}} = \alpha_R^2 m^* / 2 = 0.05 \text{ meV}$ and length scale $l_{\text{so}} = \hbar / (\alpha_R m^*) = 0.19 \mu\text{m}$ of the system, where $\hbar \alpha_R = 0.2 \text{ eV \AA}$ and $m^* = 0.02 m_e$ are realistic values for InAs [84, 90]. Throughout this paper, we consider wires of length $L = 13 l_{\text{so}}$ on a grid with spacing $a = 0.026 l_{\text{so}}$. For the one dimensional wire, we use a chemical potential $\mu = 1 E_{\text{so}}$, a Zeeman energy $E_z = 6 E_{\text{so}}$, and proximity s-wave gap $\Delta = 2 E_{\text{so}}$, such that the system in the absence of disorder and gate voltages is in the topological regime, where exponentially localized near zero energy states, the MZMs, are present at the ends of the wire enhancing the transmission amplitude [129]. We consider an electron temperature of $T = 183 \text{ mK}$ when performing the thermal average.

We describe disorder in the wire by first drawing random numbers δ with standard deviation σ_{dis} from a normal distribution and then allowing correlations with correlation length λ_{dis} by damping high Fourier modes

$$\delta_{\text{dis}}(y) = \mathcal{F}^{-1} \left[e^{-|q|\lambda_{\text{dis}}} \mathcal{F}[\delta(y)] \right] . \quad (2.301)$$

Here the case $\lambda_{\text{dis}} = 0$ corresponds to onsite disorder.

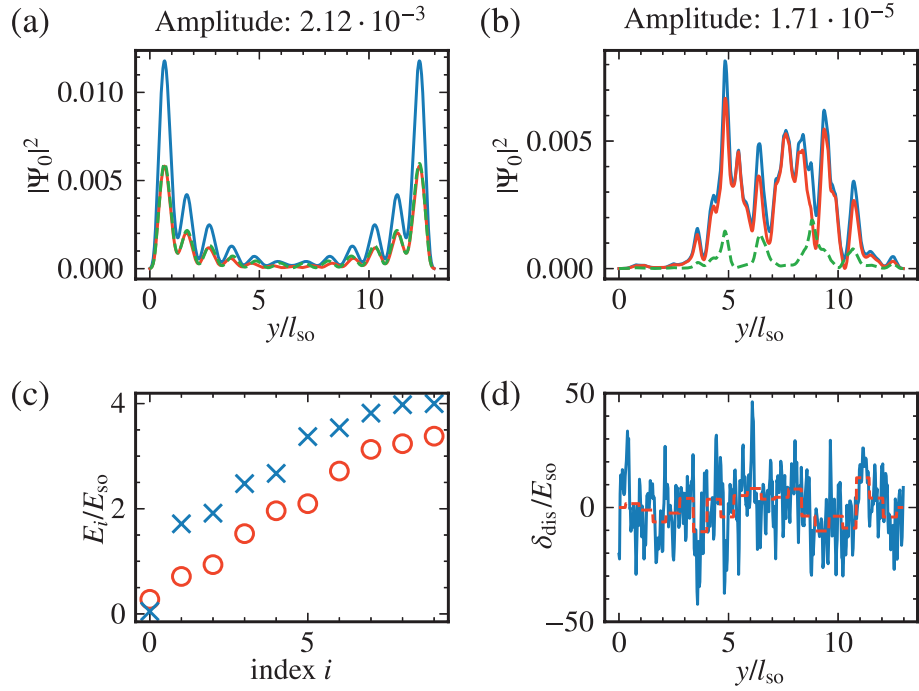


Figure 2.43: Reference results where all gate voltages are set to zero. Wave function $|\Psi_0|^2$ of the lowest level (blue), and corresponding hole wave function $|\mathbf{v}_0|^2$ (orange) and electron wave function $|\mathbf{u}_0|^2$ (green) for the case (a) without disorder, (b) with disorder strength $\sigma_{dis} = 50 E_{so}$ and a coherence length $\lambda_{dis} = 0.052 l_{so}$. (c) Energies of the lowest ten Bogoliubov levels for the case without disorder (blue crosses) and with disorder (red circles). (d) Disorder potential along the wire (blue) and average of the disorder over each gate (dashed, red). In the presence of the strong disorder, the transmission amplitude is diminished, the topological gap is destroyed, and there is no Majorana zero mode present.

To separate wire and leads, we consider steep tunnel barriers V_{conf} of shape $V_{\sigma, V_0}(y) = V_0 \exp[-y^2/(2\sigma^2)]$ with $\sigma = 0.1 l_{so}$ and $V_0 = 65 E_{so}$ given by

$$V_{conf}(y) = V_{\sigma, V_0}(y - x_0) + V_{\sigma, V_0}(y - L + x_0) \quad (2.302)$$

$$x_0 = \sqrt{2}\sigma \sqrt{\ln \left(2 \frac{V_0 + V_{lead}}{V_0 + 2V_{lead}} \right)}, \quad (2.303)$$

such that the maxima are located close to the ends of the wire at x_0 and $L - x_0$ where x_0 is chosen such that the potential has decayed to $V_0/2$ at the ends of the wire. Here, the confinement potential is created by separate gates that are not included in the optimization. The outermost gates used for optimization, start at a distance $0.3 l_{so}$ from the ends of the wire.

For simplicity, we assume that leads are normal conducting and without spin orbit coupling. We treat Coulomb blockade in the Majorana wire using a mean-field approximation such that adding an electron to the system of N_0 electrons costs an additional charging energy E_c . We consider a large charging energy $E_c = 8 E_{so}$ which allows to truncate the Hilbert space and only consider co-tunneling processes via effective energy levels $\varepsilon_{eff, i}$ containing both charging energy and single particle energy levels. Furthermore, we consider the system to be tuned to the center between the conductance resonances for a fixed particle number N_0 in the Majorana wire.

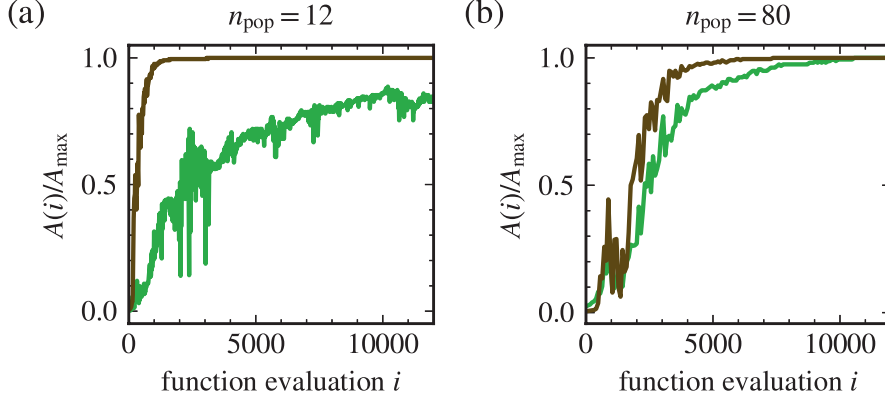


Figure 2.44: Convergence speed of the CMA-ES algorithm for optimizing the transmission amplitude A in presence of disorder using 20 gates. We consider runs of the CMA-ES algorithm with population sizes (a) $n_{\text{pop}} = 12$ and (b) $n_{\text{pop}} = 80$. In both cases, we consider a “hard” problem (green) with onsite disorder, and an “easier” problem with short range disorder correlations $\lambda_{\text{dis}} = 0.052 l_{\text{so}}$. The panels depict the rate of convergence $A(i)/A_{\text{max}}$ as a function of the number of function evaluations i (number of amplitude measurements) where A_{max} is the value to which the amplitude converges. We find that for a hard problem, it can be advantageous to use a large population size while for an easier problem a small population size yields convergences after a small number of function evaluations. In the case of onsite disorder we find convergence after a few thousand function evaluations while the algorithm often requires less than 1000 evaluations for reasonable convergence if there are correlations present in the disorder potential.

For numerically computing the amplitude of coherent transmission $A = |\langle T_{\uparrow\uparrow} + T_{\downarrow\downarrow} \rangle|$, we use the microscopic model described in Sec. 2.3.5 based on the python package KWANT [389]. As described in Eq. (2.276), the transmission amplitude in the middle between conductance resonances is obtained from

$$\mathbf{T} = \left\langle i\boldsymbol{\varphi}_R^\dagger \Gamma_R U_w \frac{1}{\varepsilon - \text{diag}(\varepsilon_{\text{eff}}) - U_w^\dagger \Sigma U_w} U_w^\dagger \Gamma_L \boldsymbol{\varphi}_L \right\rangle, \quad (2.304)$$

with eigenvectors U_w of the wire Hamiltonian Eq. (2.300), we can determine the transmission amplitude $A = |T_{\uparrow\uparrow} + T_{\downarrow\downarrow}|$ in the middle between conductance resonances. Here, ε is the energy of incoming electrons in the lead, and to conveniently obtain the tight-binding Hamiltonian, self-energies Σ_α , $\Gamma_\alpha = i(\Sigma_\alpha - \Sigma_\alpha^\dagger)$, $\Sigma = \sum_\alpha \Sigma_\alpha$, and propagating modes $\boldsymbol{\varphi}_\alpha$ of lead α , we use KWANT [389].

For transport through MZMs with wave function $\chi_{\sigma,\alpha}(y)$ exponentially localized at the wire ends y_α , the transmission amplitude is approximately $|T_{\uparrow\uparrow}| \sim |\chi_{\sigma,L}(y_L)\chi_{\sigma,R}^*(y_R)|/(E_c/2)$ and can thus be enhanced by increasing the localization of the Majorana wave functions. A finite temperature can be considered by computing the scattering matrix for different thermally excited states and averaging the transmission amplitude $A = |\langle T_{\uparrow\uparrow} + T_{\downarrow\downarrow} \rangle|$. Denoting the energies of the ABSs by ϵ_1, ϵ_2 , the contribution of trivial ABSs to A is suppressed by a factor $1 - \exp[-\beta(\epsilon_1 + \epsilon_2)]$, as there is destructive interference between the transmission through states with empty and occupied ABSs, respectively (for more details see also Sec. 2.4.3 and Refs. [131, 132]). When setting $k_B T = 0.3 E_{\text{so}}$, we find that local optima of A due to ABSs are reliably ignored, and we did not encounter false positives in our optimization runs. For the optimizations, we consider

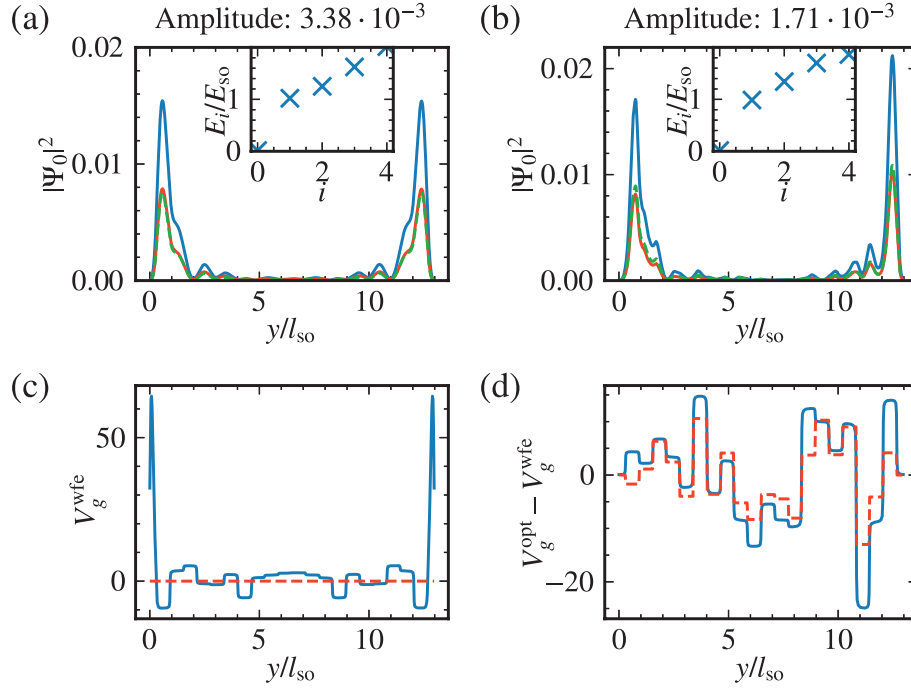


Figure 2.45: Optimization of transmission via tuning of 20 gates for a one-dimensional wire in the topological regime with $n_{\text{pop}} = 80$. Wave function $|\Psi_0|^2$ of the Majorana level (blue), and the corresponding hole and electron wave functions $|\mathbf{v}_0|^2$ (orange), $|\mathbf{u}_0|^2$ (green) for (a) no disorder in the wire and (b) disorder strength $\sigma_{\text{dis}} = 50 E_{\text{so}}$ and a correlation length $\lambda_{\text{dis}} = 0.052 l_{\text{so}}$. The insets depict the energies of the lowest five Bogoliubov levels. Optimized gate potentials in the absence of disorder are shown in (c), and in (d) the difference between optimized potential obtained with and without disorder is shown. The dashed, red line shows the negative average $-V_{\text{dis}}^{\text{avg}}$ of the disorder potential over the gates, indicating that the algorithm has learned the shape of the disorder potential. The Majorana zero mode, topological gap, and transmission amplitude are restored by the optimized gates (compare Fig. 2.43 for reference). We find that the optimal gate voltage on top of wave function engineering also tries to reduce the disorder by approximately canceling the average disorder over the gates.

transport through the first 10 levels (see also Sec. 2.5.6) and verify the final results by taking into account 50 levels for the plots. We use the `pycma` [424] python implementation of the CMA-ES [242, 243] algorithm, with an initial configuration $\mathbf{V}_g^{(0)} = 0$, the starting step size $\sigma^{(0)} = 0.1 E_{\text{so}}$, population sizes of 80 or $4 + 3 \ln(n_{\text{gates}})$, and a seed of the pseudo random number generator of 12345678, if not specified otherwise. As algorithm termination conditions, we use $\text{topfun} = 10^{-15}$, $\text{tolfunhist} = 10^{-8}$, and $\text{tolx} = 10^{-5} E_{\text{so}}$. We note, however, that the potentials do not change significantly anymore much earlier to meeting these conditions, such that one can stop the optimization earlier in an experimental situation.

In the absence of disorder and with zero voltage at all gates, the lowest level of the wire is approximately at zero energy in the middle of the topological gap (blue crosses in Fig. 2.43c) and the associated wave function $\Psi_0 = (\mathbf{u}_0, \mathbf{v}_0)$ is localized at the wire ends and satisfies the Majorana condition $|\mathbf{u}_0(y)| = |\mathbf{v}_0(y)|$ (Fig. 2.43a). However, if one adds strong disorder (Fig. 2.43d) with $\sigma_{\text{dis}} = 50 E_{\text{so}}$ and correlation length $\lambda_{\text{dis}} = 0.052 l_{\text{so}}$, both topological gap (red circles, Fig. 2.43c) and MZMs (Fig. 2.43b) are destroyed. As a result, the associated transmission amplitude is two orders of magnitude smaller compared to the clean wire.

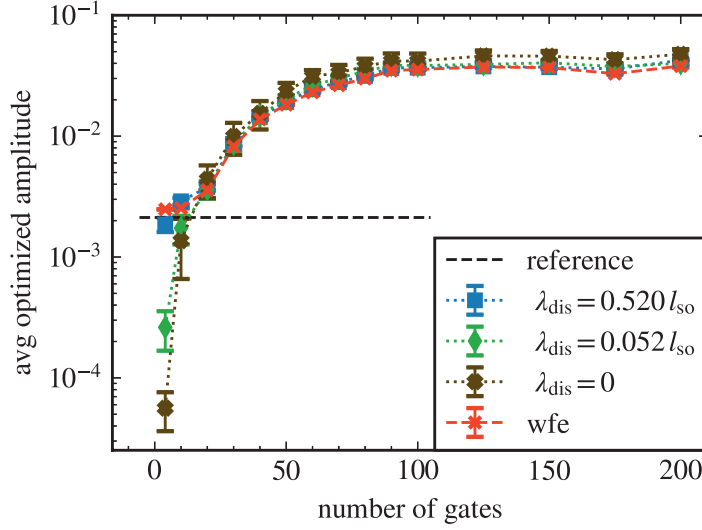


Figure 2.46: Optimized transmission amplitude as a function of the number of equally sized gates along the wire. For all calculations, we use disorder with strength $\sigma_{\text{dis}} = 50 E_{\text{so}}$ and various correlation lengths $\lambda_{\text{dis}} = 0.52 l_{\text{so}}$ (blue squares), $0.052 l_{\text{so}}$ (green diamonds), 0 (brown crosses), where zero means on site disorder. Red crosses indicate the wave function engineering result obtained for optimization without any disorder in the wire. We show averages over 10 seeds for each disorder strength and average over 10 seeds of the CMA-ES algorithm for wave function engineering. The black dashed line shows the reference value obtained without optimization and without disorder. We observe that a smaller correlation length of the disorder requires more gates to compensate disorder effects, but already 20 gates are sufficient to obtain results similar to wave function engineering.

2.5.2 Optimization results for one dimensional wires

To understand the convergence behavior and the influence of the population size n_{pop} on the CMA-ES algorithm, we consider two scenarios: (i) a finite correlation length $\lambda_{\text{dis}} = 0.052 l_{\text{so}}$ and (ii) onsite disorder, and optimize 20 gates with population sizes $n_{\text{pop}} = 12$ and $n_{\text{pop}} = 80$. In the easier case (i) already the smaller population size is sufficient to achieve fast convergence after less than 1000 function evaluations (brown line Fig. 2.44a), whereas for $n_{\text{pop}} = 80$ about five times as many evaluations are necessary (brown line Fig. 2.44b). In contrast, we find that the more difficult problem (ii) converges poorly in the case of small population sizes, but converges almost as fast as the correlated disorder case for $n_{\text{pop}} = 80$. Thus, if the primary time effort is to perform a function evaluation, we recommend deviating from the standard value $n_{\text{pop}} = 4 + 3 \ln(N_g - 1)$ [254] for the case of a small disorder correlation length.

We consider two different types of optimizations in the following: (i) wave function engineering, i.e., optimization in the absence of disorder to determine what shape a potential should have to improve the localization of the MZMs, and (ii) optimization with disorder in the wire. Comparing the resulting potentials for both cases allows to separately understand the effect of disorder on the optimization. We show results for $n_{\text{pop}} = 80$ here, but have verified all results with $N_g < 100$ for $n_{\text{pop}} = 4 + 3 \ln(N_g - 1)$ as well. In the case of wave function engineering for 20 gates, we find that the transmission amplitude is enhanced by a factor of about 1.6 due to the improved localization of the MZMs (Fig. 2.45a) while keeping a sizable topological gap (inset). To achieve this, the optimization lowered potentials of the outermost gates to draw more weight of the wave functions to the wire ends (Fig. 2.45c).

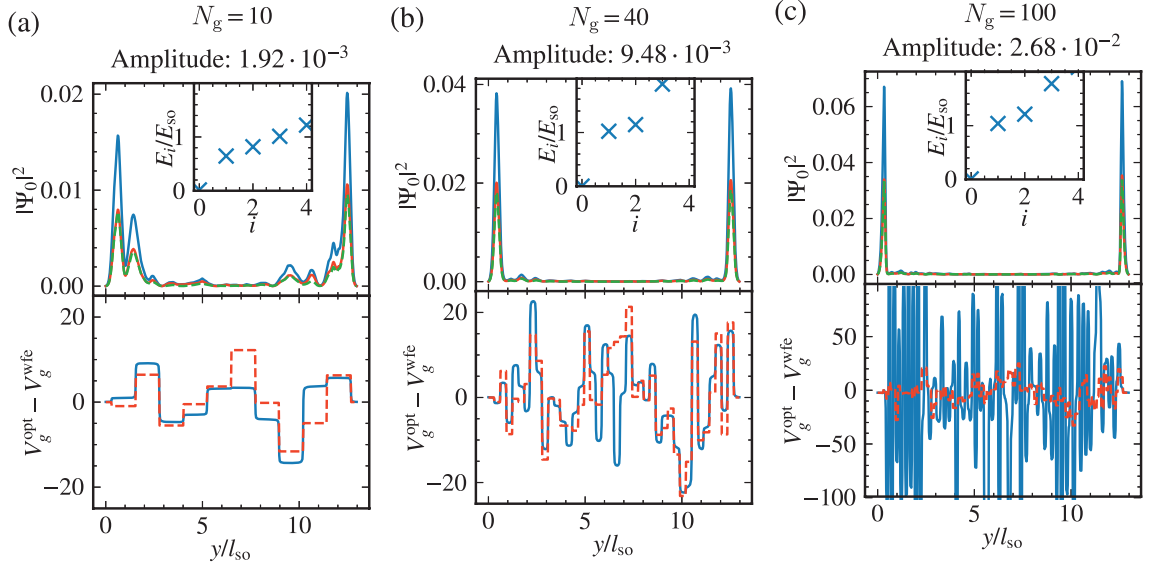


Figure 2.47: Optimization results for (a) $N_g = 10$ gates, (b) 50 gates, and (c) 100 gates. In all plots, the same disorder potential with strength $\sigma_{dis} = 50 E_{so}$ and correlation length $\lambda_{dis} = 0.052 l_{so}$ is chosen. The left panels show the wave function $|\Psi_0|^2$ of the lowest level (blue), and corresponding hole wave function $|v_0|^2$ (orange) and electron wave function $|u_0|^2$ (green). The insets depict the energies of the lowest Bogoliubov levels. The right panels show CMA-ES optimization results for the gate potentials that maximize the transmission amplitude subtracted the corresponding gate potential obtained for wave function engineering without disorder (blue). The negative averages $-V_{dis}^{avg}$ of the disorder potentials over the gates are shown by the dashed red lines. We observe an increasing amplitude with an increasing number of gates. While 10 gates are not sufficient to cancel the disorder effects, for 40 gates, the negative average disorder is in good agreement with the optimized gate potential. For 100 gates, adjustments can be made on such small scale that the optimized gate voltages can deviate from the average disorder and still cancel disorder effects.

In case (ii) with disorder (compare with Fig. 2.43b), the optimization almost completely restores the MZMs and the topological gap, increasing the amplitude by two orders of magnitude (Fig. 2.45b). The optimized potential compensates the average disorder (dashed red line Fig. 2.45d), in addition to the zero disorder optimal values (Fig. 2.45c). We emphasize again that the CMA-ES algorithm has no knowledge about the system, but only suggests gate configurations based on corresponding transmission amplitudes.

We next consider how reliable the optimization is for different disorder correlation lengths, how many gates are necessary, and how strong the dependence on the seed of the CMA-ES random number generator is. For this, we consider 15 different values for the number of gates, from $N_g = 4$ to $N_g = 200$, and three types of disorder, onsite ($\lambda_{dis} = 0$), $\lambda_{dis} = 0.052 l_{so}$ and $\lambda_{dis} = 0.52 l_{so}$, as well as wave function engineering without disorder. We consider ten different disorder realizations and average the resulting amplitudes, while in the absence of disorder we average over ten different seeds. We find that for at least 20 gates all disorder profiles can be compensated reliably (see Fig. 2.46). For too few gates $N_g \leq 10$, disorder with very small correlation length cannot be removed anymore. For many gates, $N_g \approx 100$, the amplitude saturates, having increased by one order of magnitude as compared to $N_g = 20$,

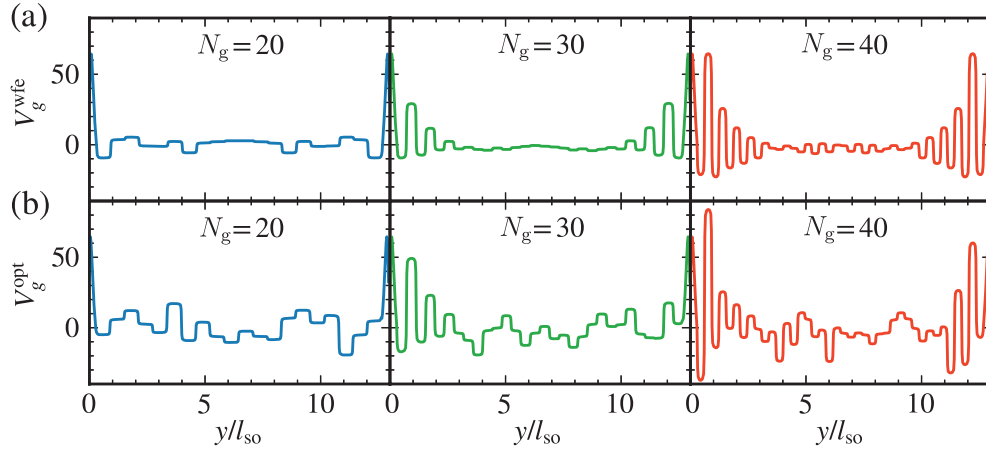


Figure 2.48: Optimized gate voltage potentials for various numbers of gates N_g , (a) for wave function engineering without disorder in the wire, and (b) in the presence of disorder with strength $\sigma_{\text{dis}} = 50 E_{\text{so}}$ and correlation length $\lambda_{\text{dis}} = 0.052 l_{\text{so}}$.

but with the drawback that up to 10^5 function evaluations are needed. We observe a sweet spot $20 \leq N_g \leq 50$, where the number of necessary function evaluations is acceptable and still significant improvements of the amplitude and complete compensation of disorder are possible.

With standard electron beam lithography and Al gates isolated by native oxide it is currently possible to construct such gate arrays of 20 to 50 gates in proximity to Majorana wires of length $L = 13 l_{\text{so}} \approx 2.6 \mu\text{m}$ [426, 427] considered here.

In Fig. 2.47, we show the resulting optimized wave functions of the first level (top panel) and corresponding optimized gate voltages relative to the voltages for wave function engineering (bottom panel) for $\sigma_{\text{dis}} = 50 E_{\text{so}}$, $\lambda_{\text{dis}} = 0.052 l_{\text{so}}$, and different numbers of gates (a) $N_g = 10$, (b) $N_g = 40$, and (c) $N_g = 100$. It becomes apparent that a larger number of gates helps to localize the MZMs; however, for $N_g = 100$ the potential can change on very short scales such that even if it strongly deviates from the average disorder over the gates it still cancels the effective disorder effects. For $N_g = 10$ there are not enough gates to fully counter the disorder, and for $N_g = 40$ the MZMs can be fully restored by a potential that, in addition to wave function engineering, follows the negative average disorder over the gates (dashed, red lines).

In addition, in Fig. 2.48, we show the gate potentials for various numbers of gates N_g optimized without disorder (panel a) and in the presence of disorder (panel b) separately. In the wave function engineering case, the potential is approximately symmetric, it is small in the middle of the wire, and it is large towards the ends, where the MZMs are present. At the ends there are oscillations in the optimized potential, such that the gates closest to the ends are lowered to attract weight of the MZM wave functions and thus increase the dot-lead couplings to the Majorana level. For a larger number of gates, the size of individual gates is smaller such that a finer optimization of the MZM region is possible resulting in a stronger localization and hence a larger transmission amplitude. When disorder is present, the potential additionally cancels the disorder and is therefore not symmetric anymore. In addition, the features observed for wave function engineering are still visible.

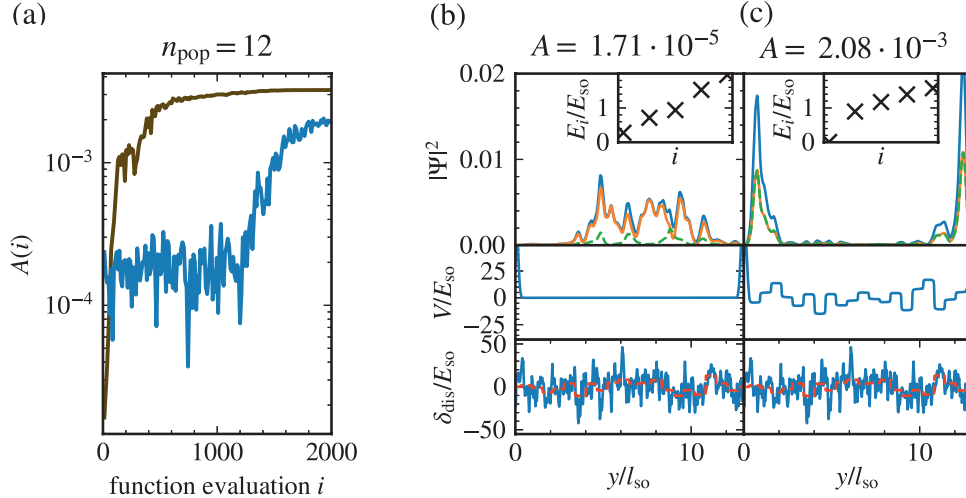


Figure 2.49: Optimization with normally distributed noise of zero mean and standard deviation 10^{-4} in the amplitude. Panel (a) shows the amplitude of coherent transmission as a function of the number of metric evaluations during the optimization of the noisy amplitude (blue) and for reference in the case without noise (brown). The right panels show the Majorana wave functions (top panels, blue: $|\Psi_0|^2$, orange: $|\mathbf{v}_0|^2$, green: $|\mathbf{u}_0|^2$), energy levels (insets), potentials of confinement and gates $V = V_{\text{conf}} + V_g$ (center panel), and random potential δ_{dis} (lower panel) for (b) the disordered wire before optimization and (c) the system optimized with noise in the amplitude. Shortly after 1000 function evaluations, the CMA-ES optimizer escapes the noisy region and restores the MZMs that were destroyed by disorder.

2.5.3 Optimization with noisy signal

Because the CMA-ES algorithm is not gradient based, the optimization can even escape a noisy region. Here, we consider again the disordered wire of Fig. 2.43, where the amplitude without optimized gates is reduced to $1.7 \cdot 10^{-5}$ due to disorder. We assume that the current measurement is noisy by adding a random number (normally distributed with mean 0 and standard deviation 10^{-4}) to each amplitude before handing it to back the optimizer. In Fig. 2.49, we show the amplitude during optimization as a function of the number of function evaluations for the noisy case (blue) and for reference without noise in the amplitude (brown). As the noise is roughly an order of magnitude larger than the initial amplitude, the signal is completely hidden in the noise. However, after about 1000 function evaluations the algorithm escapes the noisy region, and is still able to cancel the disorder and to restore the MZMs (right panel). Based on these observations, we estimate that the noise can be about an order of magnitude larger than the signal in the presence of strong disorder, but the noise should be at least an order of magnitude smaller than the signal from MZMs.

2.5.4 Optimization of average voltage

So far, we considered the case of zero average gate voltage, $b_0 = 0$, to be able to compare initial to optimized configurations at the same position in the phase diagram ($-b_0/2$ directly affects the effective chemical potential) to demonstrate that the algorithm is capable of canceling disorder effects. Here, we start initially in the trivial phase with $E_z = 6 E_{\text{so}}$, $\mu = 8 E_{\text{so}}$, and $\Delta = 2 E_{\text{so}}$

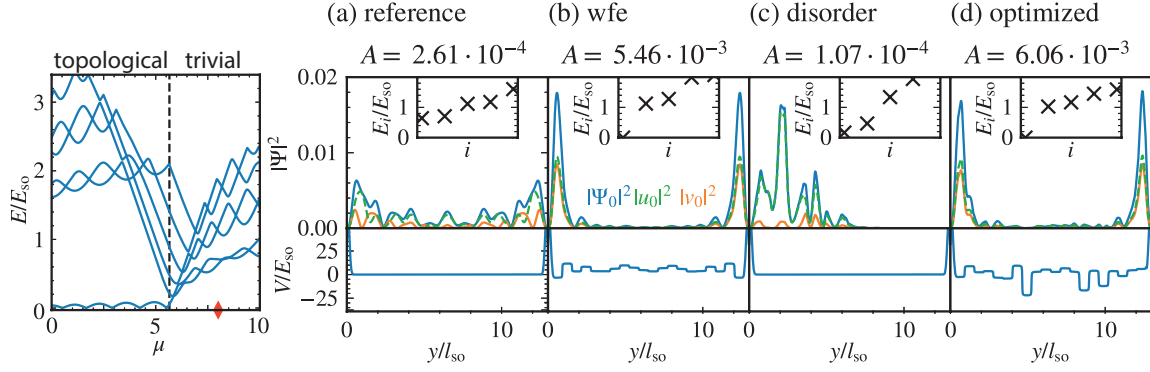


Figure 2.50: Optimization including the zero-order Fourier component b_0 for a wire starting in the trivial phase. The left panel depicts the first six energy levels as a function of the chemical potential in a clean wire where the red marker indicates the location of the initial system in the trivial phase. The right panel shows the wave functions for (a) the reference case without disorder, (b) optimized gate voltages in a clean wire, (c) the disordered wire ($\sigma_{\text{dis}} = 50 E_{so}$ and $\lambda_{\text{dis}} = 0.052 l_{so}$) with zero voltage on all gates, and (d) the optimized, disordered wire. The lower panels depict the voltages $V = V_g + V_{\text{conf}}$ of the gates and the confinement potential.

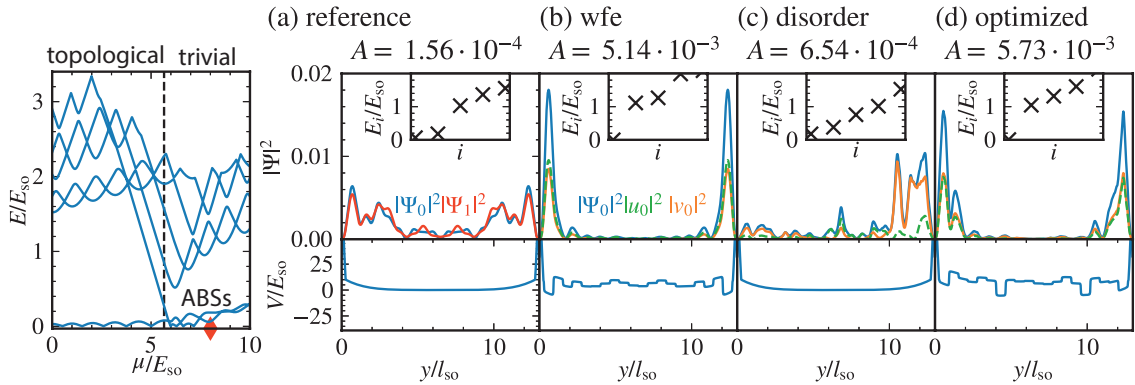


Figure 2.51: Optimization including the zero-order Fourier component b_0 for a wire starting in the trivial phase with ABSs. The left panel depicts the first six energy levels as a function of the chemical potential where the red marker indicates the location of the initial system in the trivial phase. The right panel shows the wave functions for (a) the reference case without disorder, (b) optimized gate voltages in a clean wire, (c) the disordered wire ($\sigma_{\text{dis}} = 50 E_{so}$ and $\lambda_{\text{dis}} = 0.052 l_{so}$) with zero voltage on all gates, and (d) the optimized, disordered wire. The lower panels depict the voltages $V = V_g + V_{\text{conf}}$ of the gates and the confinement potential.

(Fig. 2.50 left panel). We find (right panel) that the algorithm is capable to tune the wire into the topological phase by adjusting b_0 (panel b), and even to additionally cancel strong disorder at the same time (panel d). In Fig. 2.51, we show a similar optimization for a wire with smooth confinement [112, 131], such that there are ABSs in the initial configuration without disorder. We find that the optimization is also capable to overcome the ABSs, while tuning the wire into the topological phase.

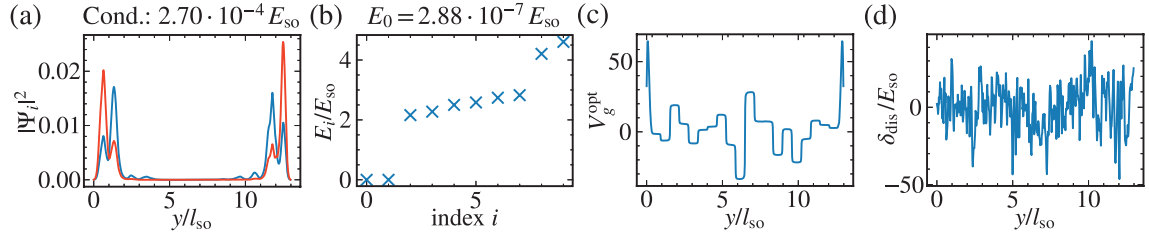


Figure 2.52: Results for direct conductance optimization in a one dimensional wire in presence of disorder. The direct conductance cannot distinguish between MZMs and ABSs. We use 20 gates of equal size along the wire. (a) Wave function $|\Psi_i|^2$ of the lowest level $i = 0$ (blue) and the second level $i = 1$ (red). (b) Energies of the lowest ten Bogoliubov levels. (c) CMA-ES optimization result that maximizes the direct conductance. (d) Disorder potential for $\sigma_{dis} = 50 E_{so}$ and $\lambda_{dis} = 0.052 l_{so}$. With the optimized gates, we observe two near zero energy ABSs and no MZMs.

2.5.5 Choice of the metric

A particular importance for automatic optimization lies in the choice of an appropriate metric that reinforces the desired features. One has to be careful that the metric does not depend excessively on unwanted features. Therefore, in the case of the Majorana wire, our choice falls on the coherent transmission amplitude, since it distinguishes ABS from MZMs [131, 132], benefits from enhanced localization of MZMs, and therefore makes MZMs more robust in the optimization. Nevertheless, one must be careful with the choice of parameters. For example, if the temperature is too small, the thermal average no longer sufficiently penalizes ABSs with small finite energy, causing these solutions to steer the CMA-ES algorithm away from Majorana solutions. Furthermore, optimization in Fourier components is very useful, since by choosing $b_0 = 0$ the mean gate potential over the wire vanishes allowing us to compare improvements from the optimization at the same position in the phase diagram. These particular decisions in the choice of the metric can have a decisive influence on the speed of convergence and the success of the optimization. CMA-ES optimization can therefore also be used as a tool to reveal weaknesses and potential alternative explanations for a signature of a desired feature in complex realistic setups.

For the case of a Majorana wire, one might be tempted to choose other metrics that are easier to obtain in an experiment by not requiring the interferometer setup. In the following, we discuss the drawbacks of three other potential metrics and additionally provide corresponding optimization results in Majorana wires with strong disorder:

First, we consider the direct conductance through the wire without an interferometer, which is easier to measure experimentally, but has the disadvantage of not being able to distinguish between ABSs and MZMs. This manifests itself in the optimization by yielding a pair of trivial near-zero energy levels (Fig. 2.52b), the ABSs, both of which are localized at the wire ends (Fig. 2.52a).

Another potential metric is the topological gap $|\varepsilon_1 - \varepsilon_0|$, which, however, does not depend on the localization of the MZMs, nor does it rely on the presence of MZMs and the topological phase either. An optimization shows large $|\varepsilon_1 - \varepsilon_0|$ (Fig. 2.53b), but the associated lowest level is not a Majorana state (Fig. 2.53a). In addition, $\varepsilon_0 = 1.9 \cdot 10^{-8} E_{so}$ is also strongly reduced, which also shows that minimizing ε_0 does not favor MZMs and can even be realized with ABSs.

Furthermore, we mentioned problems related to suboptimal parameters. For the thermal av-

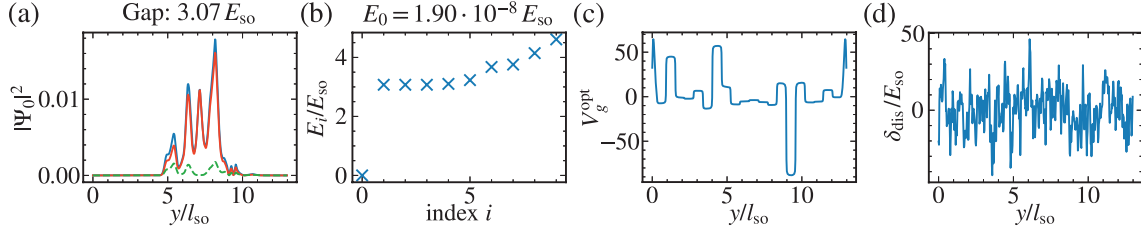


Figure 2.53: Results for gap optimization in a one dimensional wire in presence of disorder. We define the gap as the difference between the first two energy eigenvalues. We use 20 gates of equal size along the wire. (a) Wave function $|\Psi_0|^2$ of the lowest level and corresponding hole and electron wave functions $|v_0|^2$ (orange) and $|u_0|^2$ (green). (b) Energies of the lowest ten Bogoliubov levels. (c) CMA-ES optimization result that maximizes the gap. (d) Disorder potential for $\sigma_{dis} = 50 E_{so}$ and $\lambda_{dis} = 0.052 l_{so}$. With the optimized gates, we observe an increased gap, however despite the lowest level being close to zero energy it is in fact not a Majorana level as $|u_0| \neq |v_0|$.

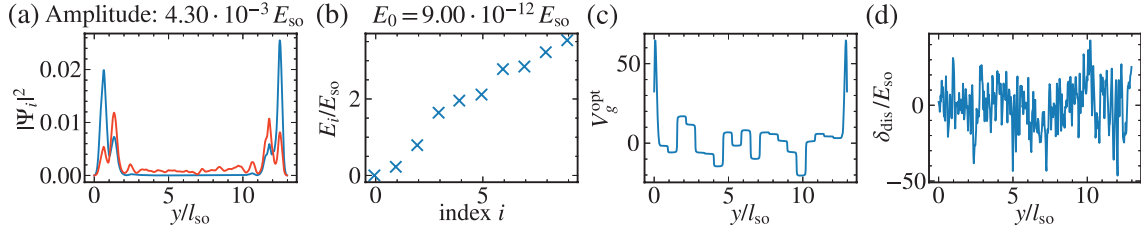


Figure 2.54: Results for transmission amplitude optimization in a one dimensional wire in presence of disorder but at very low temperature $\beta = 18 E_{so}^{-1}$. At very small temperatures an ABSs with low but finite energy has very different weight than a zero energy ABSs in the thermal average, such that the amplitude does not cancel when there are two ABSs with slightly split energy. We use 20 gates of equal size along the wire. (a) Wave function $|\Psi_i|^2$ of the lowest level $i = 0$ (blue) and the second level $i = 1$ (red). (b) Energies of the lowest ten Bogoliubov levels. (c) CMA-ES optimization result that maximizes the gap. (d) Disorder potential for $\sigma_{dis} = 50 E_{so}$ and $\lambda_{dis} = 0.052 l_{so}$. With the optimized gates, we observe a diminished gap and a pair of ABSs near zero energy which are split by a small energy difference.

erage to reliably penalize ABSs, the temperature should be sufficiently high, and at the same time, of course, the temperature must be below the critical temperature for preserving superconductivity. If one chooses a too small temperature, for example $T = 34$ mK, the optimization favors ABSs with energy slightly larger than zero (Fig. 2.54) and also the gap above the ABS levels can be strongly reduced.

2.5.6 Number of levels included in the scattering calculation

In the optimization, we include $j_{max} = 10$ effective levels when computing the scattering matrix, Eq. (2.276), Eq. (2.224), and Eq. (2.225). In the absence of disorder and if all gates are tuned to zero voltage, we found in Sec. 2.4 that the transmission amplitude is fully determined by the contribution of the MZMs in the topological regime, as contributions from consecutive higher levels have alternating signs and thus cancel each other. For the same reason, in the trivial regime, only a few levels are sufficient to obtain the transmission amplitude. As the amplitude

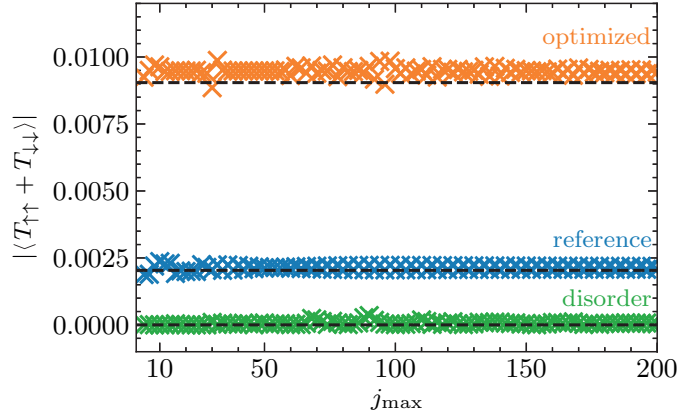


Figure 2.55: Dependence of the transmission amplitude $A = |\langle T_{\uparrow\uparrow} + T_{\downarrow\downarrow} \rangle|$ on the number of effective levels j_{\max} considered in the scattering matrix calculation. We show the transmission amplitudes for reference with strong disorder (green), for reference without disorder (blue), and for optimized gates in presence of disorder (orange). The dashed, vertical lines show the values of the transmission amplitude for taking all levels into account. Because contributions from consecutive higher levels cancel each other, we plot amplitudes only for even level numbers j_{\max} . Numerical results are shown for $N_g = 40$ gates and disorder $\sigma_{\text{dis}} = 50 E_{\text{so}}$, $\lambda_{\text{dis}} = 0.052 l_{\text{so}}$ which is also considered in Fig. 2.46, Fig. 2.47 and Fig. 2.48. It becomes apparent that the amplitude only weakly depends on j_{\max} and that reference and optimized amplitudes are clearly separated for any $j_{\max} > 4$.

cancellations are due to symmetries of the BdG wave functions, it is a priori not clear that they have to carry over to the case of strong disorder or to non-trivial choices of the gate voltages.

In Fig. 2.55, we compute transmission amplitudes $|\langle T_{\uparrow\uparrow} + T_{\downarrow\downarrow} \rangle|$ in the middle between conductance resonances as a function of the number of effective levels j_{\max} . We consider the reference case without disorder (blue) and compare to the case with strong disorder (green) and with optimized gates for countering disorder (orange). This comparison shows that the cancellations of contributions from higher levels are present in all cases such that only a few $j_{\max} \sim 10$ levels are sufficient to accurately approximate the transmission amplitude when all levels are taken into account (dashed, black lines). In addition, only including few effective levels with small single particle energies $\mathcal{E} < E_c$ is consistent with the truncation of the Hilbert space described in Sec. 2.3.3 for treating Coulomb-blockade.

In order to rule out that the optimization makes use of the cancellation effects between higher levels by fine-tuning contributions from levels j_{\max} and $j_{\max} + 1$ such that level j_{\max} gives a large contribution that would be cancelled by level $j_{\max} + 1$, we evaluate the final amplitudes after optimization taking 50 levels into account. This way, we can be sure that the optimization results are really due to optimizing the dot-lead couplings of MZMs and not due to flaws in the truncation. When using the CMA-ES algorithm it is important to be careful with such details, as it has no information about the system and can exploit all weaknesses of a model to find a maximum.

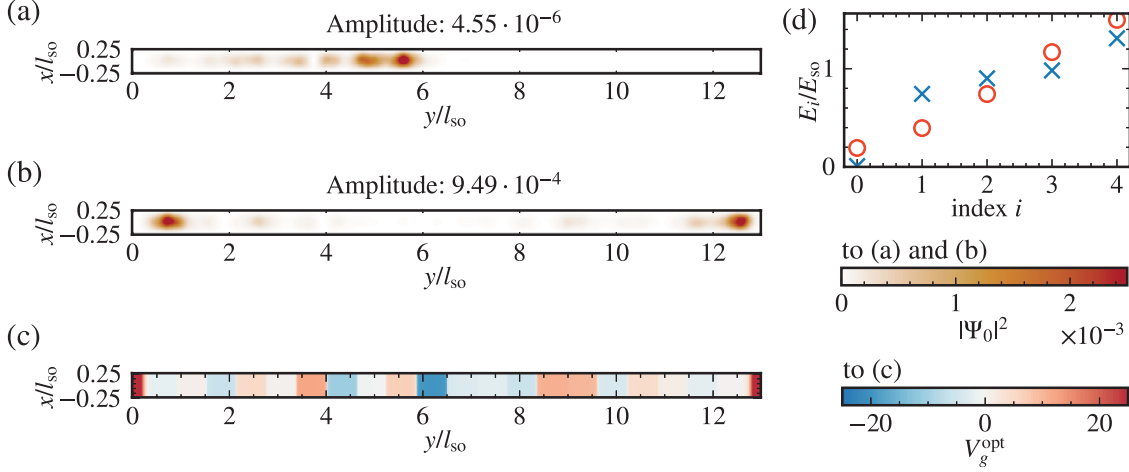


Figure 2.56: Optimization of 20 gates along a two-dimensional wire in the topological regime. Wave function $|\Psi_0|^2$ of the lowest level for (a) disorder with $\sigma_{\text{dis}} = 150 E_{\text{so}}$ and $\lambda_{\text{dis}} = 0.052 l_{\text{so}}$ where all gates are set to zero for reference and (b) optimized gates in presence of disorder. (c) CMA-ES optimization result for the gate potential that maximizes the transmission amplitude. (d) Energies of the lowest ten Bogoliubov levels for the reference case with disorder (red circles) and for the optimized gate potential (blue crosses). Similarly to the one dimensional case, the Majorana zero modes, topological gap, and transmission amplitude are restored by the optimized gates.

2.5.7 Optimization for two-dimensional wires

To consider further effects in realistic Majorana wires, we also study a two-dimensional wire with length $L_y = 13 l_{\text{so}}$ and width $L_x = 0.39 l_{\text{so}}$. The full 2d Hamiltonian is given by

$$\mathcal{H}_{\text{wire}}^{2d} = \tau_z \left[-\frac{\hbar^2}{2m^*} (\partial_x^2 + \partial_y^2) \sigma_0 - \mu \sigma_0 - i\hbar \alpha_R (\sigma_x \partial_y - \sigma_y \partial_x) + \delta_{\text{dis}}(x, y) \sigma_0 \right. \\ \left. + V_g(x, y) \sigma_0 + V_{\text{conf}}(y) \sigma_0 \right] + \frac{\mu_B g B_z}{2} \tau_0 \sigma_z + \Delta \tau_x \sigma_0, \quad (2.305)$$

with Lande factor $g = -14.9$ [381], and we take into account the orbital effect of the magnetic field by adding a Peierls phase $e^{-ie/\hbar \int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{A} \cdot d\mathbf{r}}$ to the hoppings from site \mathbf{r}_1 to site \mathbf{r}_2 . We choose the gauge for vector potential and superconducting phase as described in Sec. 2.2.5 to include the orbital effect of the magnetic field.

We first choose a chemical potential $\mu = 63 E_{\text{so}}$ and Zeeman energy $E_z = 6 E_{\text{so}}$ such that the wire in the absence of disorder and gates is in the topological regime with one occupied subband. In the presence of strong disorder, the MZMs are destroyed (Fig. 2.56a) and the gap collapses (red circles in Fig. 2.56d), but again optimization with only 20 gates along the wire can restore the MZMs (Fig. 2.56b) as well as the gap (blue crosses in Fig. 2.56d), similarly to the one dimensional case. For reference, we further show the amplitude in the absence of disorder and zero gate voltage on all gates in Fig. 2.57a, as well as the result for wave function engineering in Fig. 2.57b.

Gate optimization can also be fruitful for higher subbands, as we show in Fig. 2.58 where we consider the second topological phase for $\mu = 144.5 E_{\text{so}}$, $E_z = 6 E_{\text{so}}$, $\Delta = 2 E_{\text{so}}$. However, in

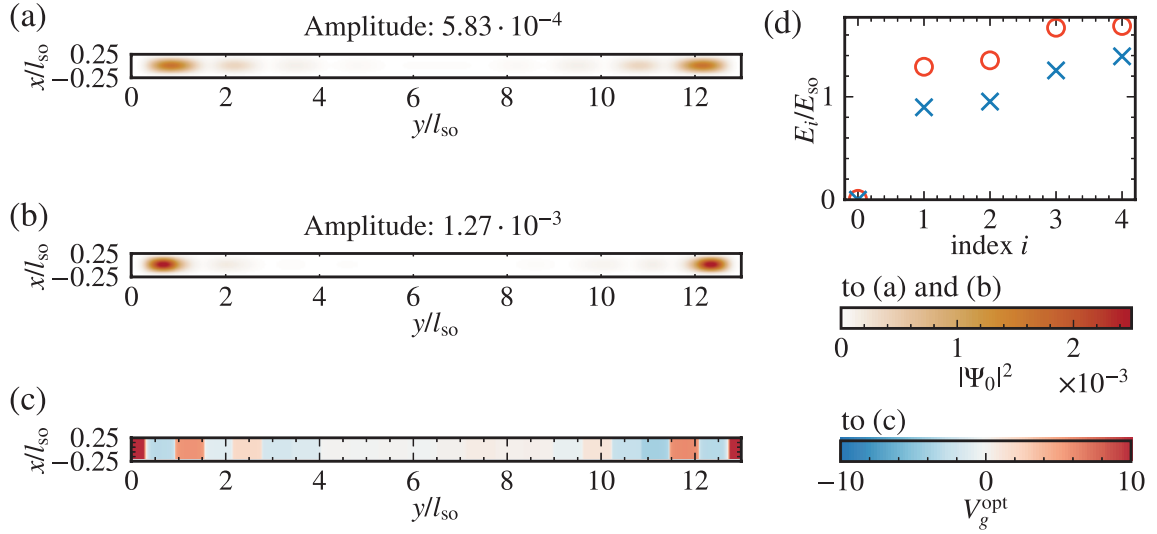


Figure 2.57: Results for optimization of a two-dimensional wire. We use 20 gates of equal size along the wire. We show the wave function $|\Psi_0|^2$ of the lowest level for (a) the reference case without disorder and with zero gate voltage on all gates, and (b) optimized gates without disorder (wave function engineering). (c) CMA-ES optimization result for the gate potential that maximizes the transmission amplitude. (d) Energies of the lowest five Bogoliubov levels for the reference case (red circles) and for the optimized gate potential (blue crosses).

the presence of levels from different subbands near the Fermi level, many subtleties arise that can distract the CMA-ES optimization, such that we regularly observe that the optimization does not restore MZMs in presence of disorder. Here, different subbands have very different coupling strengths to the leads, e.g. MZMs might have smaller couplings than states from other subbands [428]. In order to mitigate this problem, we move the superconductor and the first and last gate a distance $1.04 l_{\text{so}}$ away from the ends of the wire and add on-site disorder with strength $\delta_{\text{dis}} = 100 E_{\text{so}}$ to the superconductor-free region [428]. Using the modified setup, we find MZMs, which in the reference case without bulk disorder (Fig. 2.58a) couple about one order of magnitude stronger to the leads than other low energy levels. When adding bulk disorder (Fig. 2.58b), they are destroyed and low energy levels couple with similar strength to the leads. After optimization (Fig. 2.58c), the MZMs are restored with a coupling about twice as strong as other low energy levels. Even with these modifications, in presence of higher subbands at the Fermi level, the occurrence of Andreev bound states and other strongly coupling non-topological low energy states cannot reliably be excluded making the optimization overall more fragile. On the other hand, this also shows that CMA-ES optimization helps with identifying weaknesses in setups, such that it can also be used as a tool to test ways to stabilize desired features in the system.

2.5.8 Conclusion

We studied machine learning optimization of a gate array using the CMA-ES algorithm. Using the coherent transmission amplitude through a Coulomb blocked Majorana wire as metric, we find: (i) optimization in absence of disorder improves localization of MZMs significantly and (ii) optimization even restores MZMs fully in the case of strong disorder that otherwise destroys the topological phase. We discussed the importance of the choice of an appropriate metric, showed

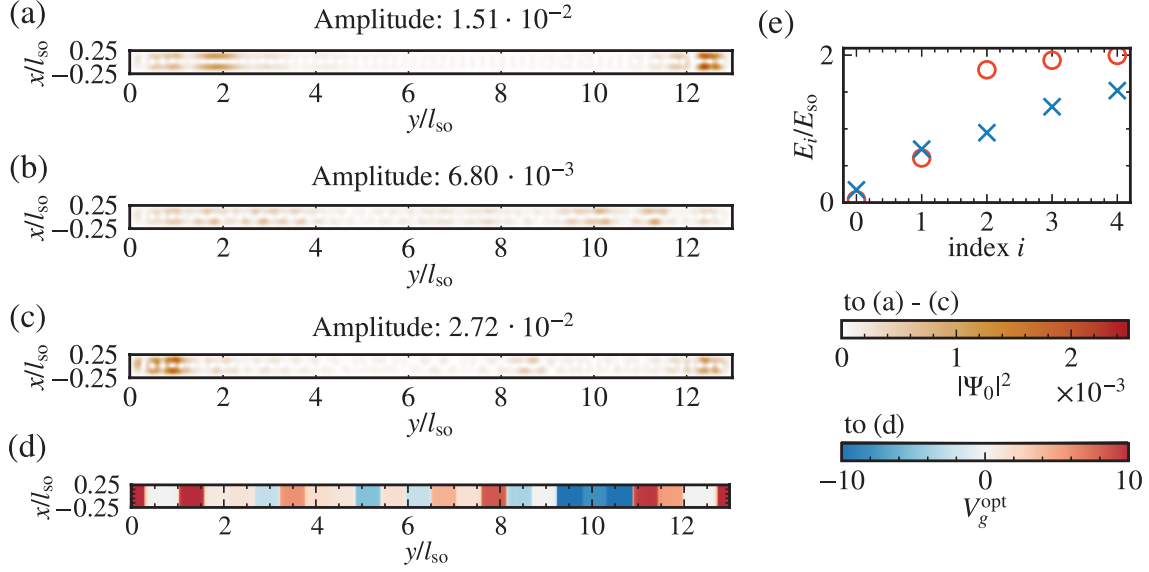


Figure 2.58: Results for optimization of a two-dimensional wire in the second topological phase. Gates and superconductor are moved a distance $1.04 l_{so}$ away from the wire ends and onsite disorder with strength $100 E_{so}$ is added in the normal-conducting region to improve coupling of MZMs to the leads. We use 20 gates of equal size along the wire with $\mu = 244.5 E_{so}$, $E_z = 6 E_{so}$, and $\Delta = 2 E_{so}$. We show the wave function $|\Psi_0|^2$ of the lowest level for (a) the reference case without disorder and with zero gate voltage on all gates, (b) with bulk disorder ($\lambda_{dis} = 0.052 l_{so}$, $\delta_{dis} = 90 E_{so}$) before gate voltage optimization, and (c) with optimized gate voltages in the presence of disorder. Panel (d) depicts the CMA-ES optimization result for the gate potential that maximizes the transmission amplitude. Panel (e) shows the energies of the lowest five Bogoliubov levels for the reference case (red circles) and for the optimized gate potential (blue crosses).

that the number of necessary function evaluations would be experimentally feasible, and that a moderate number of gates is sufficient for restoration of MZMs in the presence of disorder. While tuning scalable topological qubits remains a major challenge due to the required number of individually tunable gates, we argue that automated tuning of individual wires and smaller systems can enable proof-of-principle experiments and advance the understanding of disorder effects, with the ultimate goal of scalability in improved materials with lower disorder.

3. Random matrix theory of deep neural network weight matrices

Neural networks have been applied very successfully in a large variety of different fields [42–44, 44–48, 228]. However, the understanding of many aspects of their learning behavior remains incomplete. Because modern neural networks are often highly overparameterized [190, 204–214], they are also capable of memorizing large amounts of random noise, but at the same time they are able to learn rules from a dataset and generalize to new data [39, 429]. Recently, some insight was gained into networks in the highly overparametrized limit: from ultrawide networks it was found that learning is biased towards simple functions [47, 430–432], which might allow them to learn an underlying rule first, instead of directly overfitting to the data. The huge number of parameters and the complexity of modern networks makes it impossible to microscopically track the learning behavior of individual weights [238, 429]. Therefore, random matrix theory (RMT) has been used to study neural networks by considering predictions of the bulk behavior of the weight matrices instead of the individual components [35, 36, 433]. In Ref. [37] it has been shown that the bulk of singular value spectra follows a Marcenko-Pastur distribution [184] – the RMT prediction for random matrices – and that large singular values have heavy tailed distributions whose decay constant allows us to draw conclusions about the training state and generalization performance of the networks [238].

In this chapter, we study how information and noise are stored in the weights of neural networks by using RMT as a zero information hypothesis: because weights of neural networks are initialized randomly, they perfectly agree with RMT predictions before training. During training the networks are presented with a training dataset, and the weights are adjusted such that the network correctly classifies the images of this dataset. After training, we therefore attribute deviations from RMT predictions to learned information. We study a variety of different networks and network architectures as described in Sec. 3.1. In Sec. 3.2, we analyze these networks using RMT and find that the weights of trained networks are predominately random with only low rank parts containing the learned information. We then study in Sec. 3.3 how noise in the training dataset is stored in the weights, where we find that for adequate training, information and noise are separated in the spectra of the weights. Based on this, we introduce an algorithm [329] to remove effects of the noise from the trained weights, which can significantly improve the generalization performance.

Parts of the results presented in Sec. 3.1 and 3.2 have been published in Ref. [328] and Ref. [329], and were obtained in close collaboration with Max Staats, who also included parts of the results in his Master’s thesis [186].

3.1 Neural networks

In this section, we describe the neural networks we are going to study in this chapter. We first introduce the network architectures, datasets, how data is processed by the networks, and the training algorithm. In addition, we provide details on the hyperparameters used for training in Sec. 3.1.2. For the comparison between random and trained weights, we consider singular value decomposition of the weight matrices which we explain in detail in Sec. 3.1.3.

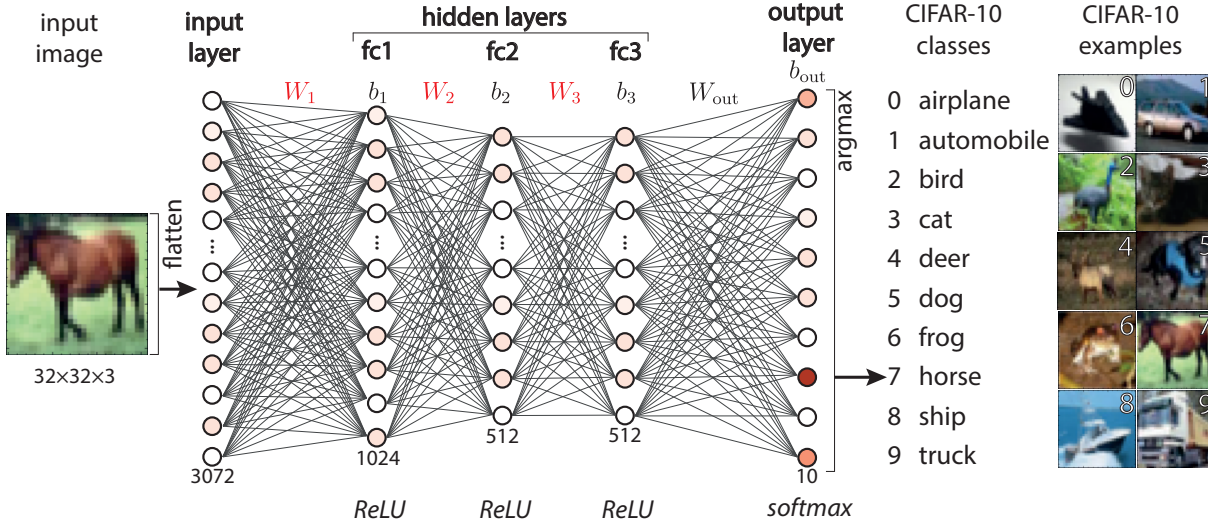


Figure 3.1: Architecture of an MLP1024 [102, 128] fully connected network for training on the CIFAR-10 dataset. The network consists of an input layer with 3072 neurons (circles) that accepts images of size 32×32 with 3 color channels, an output layer of 10 neurons, each representing a class of the CIFAR-10 dataset, and three hidden layers with 1024, 512, and 512 neurons, respectively. Every neuron in layer l is connected with each neuron in the previous layer $l - 1$ by weights $W^{(l)}$, and each layer (except for the input layer) has an associated bias vector $b^{(l)}$. For hidden layers we consider ReLU activations, and we use softmax as the output activation function. An image is flattened, and then activates the input layer according to its pixel color values. Then these activations are propagated through the network to the output layer according to Eq. (3.2), where the index of the neuron with the highest output activation defines the class predicted by the network. Before training, weights are initialized randomly (from a Glorot-uniform [189] distribution), and biases are set to zero. During training, the network is presented the training subset of the CIFAR-10 data set, and weights and biases are adjusted such that the network predictions match the labels. On the right-hand side, we show one example image from each of the ten classes and the corresponding label.

3.1.1 Network architectures and performance for image classification

To cover a large variety of networks for image classification, we consider both networks we train ourselves and large pre-trained networks that are available via the `torch` [434] and `tensorflow` [188] python packages. For the former, we train the networks on the CIFAR-10 [190] dataset which consists of 60000 images $\{\mathbf{x}_k\}$ of shape $32 \times 32 \times 3$ (with 3 color channels) from ten different classes $C_k \in \{0, 1, \dots, 9\}$ with the corresponding labels $\{\mathbf{y}_k\}$. The label of image k is a vector of length ten with a 1 at the position of the class C_k and zeros in all other entries, $[\mathbf{y}_k]_j = \delta_{j,C_k}$. We split the dataset into three subsets: i) a training set of 50000 images that is presented to the network during training, ii) a validation set of 2000 images that is not directly shown to the network during training but which is used to optimize hyperparameters, and iii) a test dataset of 8000 images that is only shown to the fully trained and finished network to estimate its generalization abilities to new, unseen data. It is important to never optimize hyperparameters on the test set. Otherwise, the accuracy obtained would not be a measure for the generalization performance.

Fully connected networks

A type of networks that we train are fully connected feedforward networks that consist of an input layer, three hidden layers (fc1, fc2, fc3), and an output layer as depicted in Fig. 3.1. Each layer consists of n_i neurons whose activations are real numbers denoted as $\mathbf{a}^{(l)} \in \mathbb{R}^{n_l}$. In addition, the hidden layers and the output layer have a weight matrix $W^{(l)} \in \mathcal{M}_{n_l \times n_{l-1}}$, a bias vector $\mathbf{b}^{(l)}$, and an activation function $f^{(l)}: \mathbb{R}^{n_l} \rightarrow \mathbb{R}^{n_l}$. As discussed in the introduction, an image from the CIFAR-10 dataset with size $32 \times 32 \times 3$ that is presented to the network is first flattened to a single vector \mathbf{x}_k with 3072 entries that activates the neurons in the input layer, i.e., the activations of the input layer are set to the pixel color values of the image

$$\mathbf{a}^{(0)}(\mathbf{x}_k) = \mathbf{x}_k . \quad (3.1)$$

Then these input activations are propagated through the network such that the activations of layer l are given by

$$\mathbf{a}^{(l)}(\mathbf{x}_k) = f^{(l)} \left(W^{(l)} \mathbf{a}^{(l-1)}(\mathbf{x}_k) + \mathbf{b}^{(l)} \right) . \quad (3.2)$$

The prediction of the network is then obtained from the activations of the output layer \mathbf{a}_L : the predicted class $\hat{C}^{(k)}$ is given by the position of the maximum activation in \mathbf{a}_L ,

$$\hat{C}_k = \operatorname{argmax} \left(\mathbf{a}^{(L)}(\mathbf{x}_k) \right) . \quad (3.3)$$

Initially the weights $W^{(l)}$ are randomly drawn from a Glorot-uniform distribution [189], and biases $\mathbf{b}^{(l)}$ are initialized with zeros. During training the network is shown the image-label pairs of the training dataset, and the weights and biases are adjusted such that the network predictions agree with the labels (see Sec. 3.1.2).

We consider fully connected networks with layer sizes $\mathbf{n}_{\text{MLP1024}} = [3072, 1024, 512, 512, 10]$ denoted as MLP1024 and networks with $\mathbf{n}_{\text{MLP512}} = [3072, 512, 512, 512, 10]$ that we call MLP512. As activation functions, we choose ReLU for the hidden layers,

$$f^{(l)}(\mathbf{z}) = \sum_j \max(z_j, 0) \mathbf{e}_j , \text{ for } 1 \leq l \leq 3 \quad (3.4)$$

and softmax for the output layer

$$f^{(L)}(\mathbf{z}) = \sum_j \frac{\exp(z_j)}{\sum_k \exp(z_k)} \mathbf{e}_j . \quad (3.5)$$

Convolutional neural networks

Besides fully connected layers, we also consider networks with another type of trainable layer, called a convolutional layer. The weights of a convolutional layer with n_i filters of dimensions $d_1 \times d_2 \times c_{l-1}$ are 4D tensors $W_l \in \mathcal{M}_{(d_1 \times d_2 \times c_{l-1}) \times n_l}$ that act on a 3D tensor of neurons with shape $(n \times m) \times c_{l-1}$ to produce n_i feature maps, i.e. the activations of the convolutional layer, by performing a convolution. For this convolution, the filter – which is typically much smaller than the size of the neurons of the previous layer ($d_1 < n$, $d_2 < m$) – is moved over the neurons

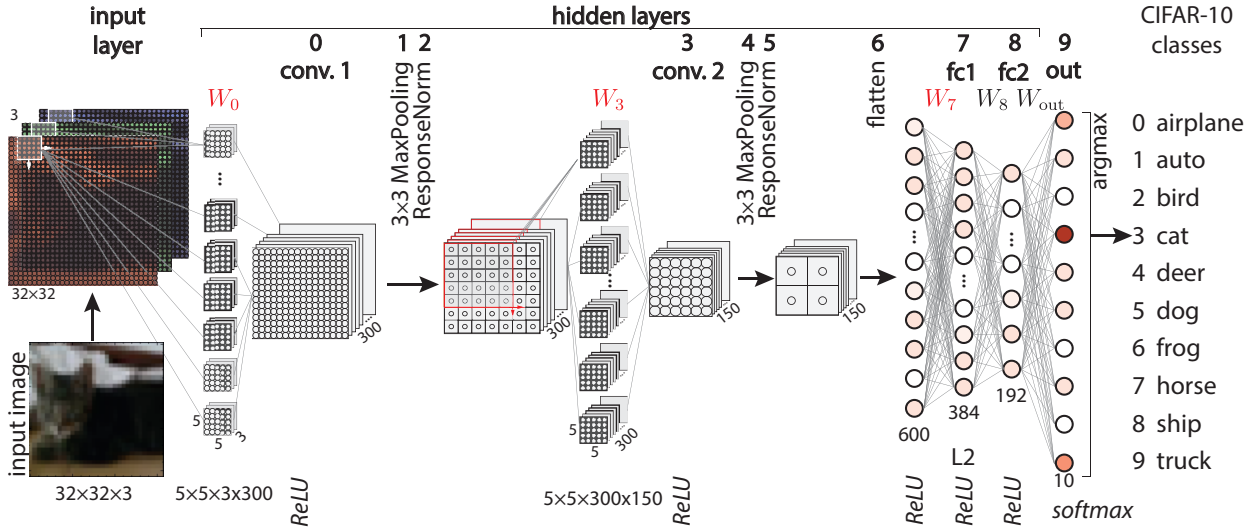


Figure 3.2: Architecture of a miniAlexNet [102, 128] convolutional neural network for training on the CIFAR-10 dataset. The network consists of an input layer with shape $32 \times 32 \times 3$, an output layer of 10 neurons, one for each class, and hidden layers. These are two convolutional layers each followed by a 3×3 max-pooling layer, a response normalization layer, and two fully connected layers with 384 and 192 neurons, respectively. Here, the weights of the convolutional layers are 4D tensors. In the first convolutional layer (conv. 1), there are 150 filters each of shape $5 \times 5 \times 3$ that act on the three color channels by sliding them over the input layer to produce 300 feature maps as the activations of the layer. On each of these feature maps, 3×3 regions are contracted to a single that has the activation of the neuron with the largest activation in this region, and then activations are normalized between neighboring neurons (local response normalization with radius 5, bias 1, $\alpha = 1$, and $\beta = 0.5$ [128]). To move activations to the next convolutional layer (conv. 2), 150 filters of shape $5 \times 5 \times 300$ are applied to the feature maps yielding 150 new feature maps which are again reduced by pooling and normalized by local response-normalization. All 600 activations of the second convolutional layer are then flattened and processed through the remaining dense layers as described for the fully connected network. We use ReLu activations for the hidden layers, softmax for the output layer, and we train the first dense layer using L_2 regularization.

of layer $l - 1$, and the activations of the neurons of the convolutional layer l are given by the inner product of the filter with the region of neurons it acts on [435]:

$$a_{i,j,r}^{(l)} = \sum_{i_1=1}^{d_1} \sum_{i_2=1}^{d_2} \sum_{k=1}^{c_{l-1}} W_{i_1,i_2,r}^{(l)} a_{i+i_1,j+i_2,k}^{(l-1)}, \quad (3.6)$$

with integer indices $0 \leq r \leq n_i$, $0 \leq i \leq n - d_1$, and $0 \leq j \leq m - d_2$. In practice, to cover all neurons of layer $l - 1$ evenly, one pads the previous layer with a stripe of neurons with zero-activations of width $d_1/2$ on top and bottom, and $d_2/2$ on the sides before performing the convolution.

The main advantages of a convolutional layer are that (i) information can be stored with a significantly smaller number of weight entries compared to fully connected layers and (ii) filters can learn features – like shapes and edges – such that the activations on the feature maps are large where the corresponding feature is located in the image (or in the neurons of the

previous hidden layer). Convolutional layers may therefore also be trained to locate objects in images [436], and convolutional networks are often superior to fully connected networks for image classification [44].

For the convolutional neural networks (CNNs) we consider here, each convolutional layer is directly followed by a max-pooling layer [437, 438] of size 3×3 which subdivides each feature map into patches of size 3×3 and replaces the 9 neurons in each patch by a single one whose activation is that of the largest activation in the patch. This is a coarse graining step to reduce the number of neurons that need to be processed by the next layer. To avoid diverging activations for unbound activation functions like ReLu, we additionally perform a local-response normalization after each max-pooling operation. For this the surrounding activations of each neuron are multiplied with a short ranged function that decreases with the distance to the center neuron, the resulting products are summed up, and the activation of the neuron in the center is then divided by this sum.

The CNN architecture we consider in the following is denoted as miniAlexNet (see Fig. 3.2) and was suggested by Zhang et al. [39]. Here, the activations of the input layer are directly given by the $32 \times 32 \times 3$ pixel values of an input image from the CIFAR-10 dataset. The input layer is followed by two sets of layers containing a convolutional layer, a 3×3 max-pooling layer, and a local response normalization layer (radius 5, bias 1, $\alpha = 1$, $\beta = 0.5$ [39]) each. The first convolutional layer (conv. 1) contains 300 filters of shapes $5 \times 5 \times 3$, and for the second convolutional layer, we consider 150 filters of shape $5 \times 5 \times 300$. After this, all 600 neurons of the 150 feature maps are flattened and connected to a dense layer with 384 neurons, followed by another fully connected hidden layer with 192 neurons that is further connected to the output layer of size 10. The predictions of the network are obtained from the output layer in the same way as for the MLP networks. We again use ReLU as activations for the hidden layers and softmax for the output layer.

Large pre-trained networks

The advantage of the self-trained networks, discussed so far, is that we have full control of the training process and can also adjust the training set, e.g. by replacing portions of the labels by wrong data to see how the networks react to this label noise. However, training of large modern networks can require huge amounts of computing resources and time [44] which is why we also consider large networks that were already trained on supercomputers available via `torch` [434] and `tensorflow` [188]. We here focus on alexnet [44] and vgg19 [213] trained on the much larger imagenet [439] dataset that contains about 3.2 million labeled images from 1000 classes. Both networks are convolutional neural networks with exceptional performance on the imagenet dataset (see also Tab. 3.1 below).

3.1.2 Details on preprocessing and training

Preprocessing

Preparing the data before training can be crucial for the learning performance [440]. To make the images more comparable, we standardize any image \mathbf{x} that is handed to the network by

$$\boldsymbol{\xi} = \frac{\mathbf{x} - \boldsymbol{\mu}}{\sigma}, \quad (3.7)$$

where $\boldsymbol{\mu}$ is the mean and σ the standard deviation of the pixel values in image \mathbf{x} . The same preprocessing also needs to be applied to validation and test data, when training networks with preprocessed images.

Training with gradient descent

So far, we described the architecture of the networks, how the networks predict the class of an image, and how the parameters, i.e., the weights and biases are initialized. During training, the network is presented the N standardized images $\{\mathbf{x}_k\}$ and labels $\{\mathbf{y}_k\}$ of the training set in batches of size $n_b = 32$, and the aim is to adjust the weights and biases such that the predictions \hat{C}_k , Eq. (3.3), agree with the labels $C_k = \text{argmax}(\mathbf{y}_k)$. This agreement is measured with a loss function ℓ for which we choose the cross-entropy cost [435]

$$\ell(\mathbf{x}_k; \mathbf{W}, \mathbf{b}) = -\frac{1}{n_b} \sum_{k=1}^{n_b} \mathbf{y}_k \cdot \ln \left(\mathbf{a}^{(L)}(\mathbf{x}_k) \right) = -\frac{1}{n_b} \sum_{k=1}^{n_b} \ln \left(a_{C_k}^{(L)}(\mathbf{x}_k) \right) . \quad (3.8)$$

Here, $a_j^{(l)}$ denotes the j -th entry of the activations $\mathbf{a}^{(l)}$ of layer l . Because we use softmax as the output activation function, the entries of \mathbf{a}_L form a probability distribution with $\sum_j a_j^{(L)} = 1$. Therefore, if the prediction $\mathbf{a}^{(L)}(\mathbf{x}_k)$ matches the label \mathbf{y}_k , i.e. $a_j^{(L)}(\mathbf{x}_k) = \delta_{j,C_k}$, the contribution for image k to the loss is zero. In the case where the network gives 0 probability for the given class, i.e. $a_{C_k}^{(L)}(\mathbf{x}_k) = 0$, the loss diverges to ∞ . If the probability for the correct class is in between 0 and 1, there is a positive contribution from image k to the loss. Therefore, minimizing the loss aligns the predictions of the network for the images of the training dataset with the labels.

To minimize the loss, we use mini-batch gradient descent, i.e, the training dataset is split into batches of size 32, the batches are shown to the network, and for each batch a gradient update

$$\begin{aligned} W_{ij}^{(l)} &\longrightarrow W_{ij}^{(l)} - \eta \frac{\partial \ell}{\partial W_{ij}^{(l)}} \\ b_i^{(l)} &\longrightarrow b_i^{(l)} - \eta \frac{\partial \ell}{\partial b_i^{(l)}} \end{aligned} \quad (3.9)$$

with learning rate η is performed. In general, computing many gradients can be computationally expensive; however, we can use an efficient algorithm called backpropagation [441] to compute the updates. Backpropagation is based on the chain rule for differentiation: using the pre-activations $\mathbf{z}^{(l)} = W^{(l)} \mathbf{a}^{(l-1)} + \mathbf{b}^{(l)}$ with $\mathbf{a}^{(l)} = \mathbf{f}^{(l)}(\mathbf{z}^{(l)})$, the partial derivatives of the loss with respect to the weights can be written as

$$\frac{\partial \ell}{\partial W_{ij}^{(l)}} = \sum_{i_1, i_2} \frac{\partial z_{i_2}^{(L)}}{\partial W_{ij}^{(l)}} \frac{\partial a_{i_1}^{(L)}}{\partial z_{i_2}^{(L)}} \frac{\partial \ell}{\partial a_{i_1}^{(L)}} \quad (3.10)$$

$$= \sum_{i_1, i_2, \dots, i_{2(L-l)+1}} \frac{\partial z_{i_{2(L-l)+1}}^{(l)}}{\partial W_{ij}^{(l)}} \frac{\partial a_{i_{2(L-l)}}^{(l)}}{\partial z_{i_{2(L-l)+1}}^{(l)}} \dots \frac{\partial z_{i_4}^{(L-1)}}{\partial a_{i_5}^{(L-2)}} \frac{\partial a_{i_3}^{(L-1)}}{\partial z_{i_4}^{(L-1)}} \frac{\partial z_{i_2}^{(L)}}{\partial a_{i_3}^{(L-1)}} \frac{\partial a_{i_1}^{(L)}}{\partial z_{i_2}^{(L)}} \frac{\partial \ell}{\partial a_{i_1}^{(L)}} , \quad (3.11)$$

where

$$\frac{\partial z_{i_{2(L-l)+1}}^{(l)}}{\partial W_{ij}^{(l)}} = \delta_{i, i_{2(L-l)+1}} a_j^{(l-1)} , \quad (3.12)$$

$$\frac{\partial z_i^{(m)}}{\partial a_j^{(m-1)}} = W_{ij}^{(m)} , \quad (3.13)$$

$$\frac{\partial a_i^{(m)}}{\partial z_j^{(m)}} = \frac{\partial [\mathbf{f}^{(m)}(\mathbf{z}^{(m)})]_i}{\partial z_j^{(m)}} \equiv F_{ij}'^{(m)} . \quad (3.14)$$

For ReLU activations, the last expression simplifies to $F'_{ij}{}^{(m)} = \theta(z_i^{(m)})\delta_{ij}$, and for the more complicated softmax activation, one obtains the non-diagonal expression $F'_{ij}{}^{(L)} = \delta_{ij}[\mathbf{f}^{(L)}(\mathbf{z}^{(L)})]_i(1 - [\mathbf{f}^{(L)}(\mathbf{z}^{(L)})]_j) + (1 - \delta_{ij})[\mathbf{f}^{(L)}(\mathbf{z}^{(L)})]_i[\mathbf{f}^{(L)}(\mathbf{z}^{(L)})]_j$. It is convenient to introduce vectors $\boldsymbol{\delta}^{(l)}$ [435] as

$$\begin{aligned}\delta_j^{(l)} &\equiv \frac{\partial \ell}{\partial z_j^{(l)}} = \sum_{i_1, i_2, \dots, i_{2(L-l)}} \frac{\partial a_{i_{2(L-l)}}^{(l)}}{\partial z_{i_{2(L-l)+1}}^{(l)}} \cdots \frac{\partial z_{i_4}^{(L-1)}}{\partial a_{i_5}^{(L-2)}} \frac{\partial a_{i_3}^{(L-1)}}{\partial z_{i_4}^{(L-1)}} \frac{\partial z_{i_2}^{(L)}}{\partial a_{i_3}^{(L-1)}} \frac{\partial a_{i_1}^{(L)}}{\partial z_{i_2}^{(L)}} \frac{\partial \ell}{\partial a_{i_1}^{(L)}} \\ &= \sum_{i_1, i_2, \dots, i_{2(L-l)}} F'_{i_{2(L-l)}j}{}^{(l)} W_{i_{2(L-l)-1}, i_{2(L-l)}}^{(l+1)} F'_{i_{2(L-l)-1}, i_{2(L-l)}}{}^{(l+1)} \cdots W_{i_2, i_3}^{(L)} F'_{i_1 i_2}{}^{(L)} \frac{\partial \ell}{\partial a_{i_1}^{(L)}} ,\end{aligned}$$

such that

$$\frac{\partial \ell}{\partial W_{ij}^{(l)}} = a_j^{(l-1)} \delta_i^{(l)} = [\boldsymbol{\delta}^{(l)} (\mathbf{a}^{(l-1)})^T]_{ij} . \quad (3.15)$$

Hence, for obtaining the gradient updates of all weights only two passes through the network are necessary, one usual forward pass of the mini-batch to obtain all $\mathbf{z}^{(m)}$ and $\mathbf{a}^{(m)}$, and the $\boldsymbol{\delta}^{(m)}$ are obtained from a single backward pass

$$\delta_i^{(L)} = \sum_j F'_{ji}{}^{(L)} \frac{\partial \ell}{\partial a_j^{(L)}} \quad (3.16)$$

$$\delta_i^{(l)} = \sum_{jn} F'_{ni}{}^{(l)} W_{jn}^{(l+1)} \delta_j^{(l+1)} . \quad (3.17)$$

To summarize, the four equations of backpropagation are given by [435]

$$\boldsymbol{\delta}^{(L)} = [F'{}^{(L)}]^T \nabla_{\mathbf{a}^{(L)}} \ell \quad (3.18)$$

$$\boldsymbol{\delta}^{(l)} = [F'{}^{(l)}]^T \left([W^{(l+1)}]^T \boldsymbol{\delta}^{(l+1)} \right) \quad (3.19)$$

$$\nabla_{W^{(l)}} \ell = \boldsymbol{\delta}^{(l)} [\mathbf{a}^{(l-1)}]^T \quad (3.20)$$

$$\nabla_{\mathbf{b}^{(l)}} \ell = \boldsymbol{\delta}^{(l)} . \quad (3.21)$$

In the case of ReLU activations in layer l , where F' is a diagonal matrix, one may use an element wise (Hadamard) product with the diagonal entries of F' instead of matrix multiplication.

To minimize the loss, one gradient update according to Eq. (3.9) is performed for each mini-batch of the training data. One says that one epoch has passed after all images in the training dataset are shown to the network once.

L_2 regularization

Sometimes it can be useful to force the weight entries of a layer to have small magnitudes, which can also help to prevent diverging activations for unbound activation functions and in some cases even to reduce overfitting [435]. This is achieved by adding a regularization term to the loss function that penalizes large weights [188]

$$\ell_{L_2}(\mathbf{x}_k; \mathbf{W}, \mathbf{b}) = \ell(\mathbf{x}_k; \mathbf{W}, \mathbf{b}) + \sum_l \lambda^{(l)} \sum_{ij} |W_{ij}^{(l)}|^2 , \quad (3.22)$$

where the hyperparameter $\lambda^{(l)}$ is called the L_2 regularization strength that can be individually set for each layer. By default, we do not use L_2 regularization, i.e. $\lambda^{(l)} = 0$.

Gradient descent with momentum

So far, we described the standard gradient descent algorithm where updates of the weight (biases) are along the opposite direction of the gradient of the loss function with respect to the weights $-\nabla_W \ell$ (biases $-\nabla_b \ell$) rescaled by the learning rate η . As this gradient update only includes information at the current step, it can easily be slowed down or even get stuck in a small local minimum on the way to a better local (or even global) optimum of the loss landscape. One way to retain information about the previous steps is to apply the gradient updates to *velocities* (denoted by w for weights, v for biases) which are initially set to zero and then adjusting in each step according to

$$w_{ij}^{(l)} \longrightarrow \mu w_{ij}^{(l)} - \eta \frac{\partial \ell}{\partial W_{ij}^{(l)}} \quad (3.23)$$

$$v_i^{(l)} \longrightarrow \mu v_i^{(l)} - \eta \frac{\partial \ell}{\partial b_i^{(l)}}. \quad (3.24)$$

Here, $0 \leq \mu \leq 1$ is the so-called *momentum* hyperparameter. We then use the velocities to update weights and biases according to [435]

$$W_{ij}^{(l)} \longrightarrow W_{ij}^{(l)} + w_{ij}^{(l)} \quad (3.25)$$

$$b_i^{(l)} \longrightarrow b_i^{(l)} + v_i^{(l)}. \quad (3.26)$$

Hence, the case $\mu = 0$ corresponds to standard gradient descent, and the closer μ is to +1, the more the update step is influenced by the previous one. Training with momentum can help to overcome small local minima and often speeds up convergence [435].

Training parameters and learning rate schedule

For training our networks on the CIFAR-10 dataset, we use mini-batch stochastic gradient descent with a mini-batch size of $n_b = 32$ and momentum $\mu = 0.95$. We train without regularization, except of the first dense layer in miniAlexNet, where we use L_2 regularization with strength $\lambda = 10^{-4}$. In Sec. 3.2 and most of Sec. 3.3, we train networks for 100 epochs using the following learning rate schedule: we start with an initial learning rate $\eta_0 = 0.005$ at $t = 0$ and reduce the learning rate at each update step t using a decay rate $0 \leq r \leq 1$ [188]

$$\eta_t = \eta_0 \cdot r^{t/t_0}, \quad (3.27)$$

where we choose $r = 0.95$ and $t_0 = 50000/32$. Here, t_0 is the number of update steps per epoch. This means, every epoch, the learning rate is reduced by a factor 0.95. The idea behind using a variable learning rate is that initially a large learning rate allows the network to make large steps in the direction of the minimum and to escape smaller local minima, and towards the end of training, a small learning rate helps to get close to an optimal minimum without overshooting too much. In Sec. 3.3, we additionally train some MLP1024 networks with an overfitting schedule that keeps training for much longer than needed to achieve 100% training accuracy. For this schedule, we train for 500 epochs with a stepwise learning rate initially set to $\eta_0 = 0.001$ and reduced by a factor of 0.7 every 50 epochs.

The code for training the networks and for reproducing the results of the publication Ref. [328] is open source and has been made available online as a Zenodo archive [442].

Table 3.1: Performance of trained neural networks. For MLP architectures, we trained 10 networks with different seeds for the initialization. We report mean generalization accuracy and error of the mean. We use d to indicate a dense layer, c for a convolutional layer, p for max pooling, f for flattening, and r for response normalization layer (with a depth radius of 5, a bias of 1, $\alpha = 1$, and $\beta = 0.5$). We measured accuracies marked with * on the imagenet validation dataset. Accuracies for the pre-trained torch models are from the torch documentation [434] where they were obtained for the imagenet test dataset.

network	dataset	noise	train acc	test acc
MLP512 {d 3072, d 512, d 512, d 512, d 10}	CIFAR-10	0%	100%	(55.70 \pm 0.11)%
MLP1024 {d 3072, d 1024, d 512, d 512, d 10}	CIFAR-10	0%	100%	(56.99 \pm 0.11)%
		10%	100%	(52.10 \pm 0.12)%
		20%	100%	(48.26 \pm 0.13)%
		30%	100%	(44.55 \pm 0.15)%
		40%	100%	(40.57 \pm 0.11)%
		50%	100%	(36.69 \pm 0.11)%
		60%	100%	(32.46 \pm 0.23)%
		70%	100%	(27.85 \pm 0.15)%
		80%	100%	(23.13 \pm 0.19)%
		100%	100%	10.3%
MLP1024 {d 3072, d 1024, d 512, d 512, d 10} overfitting schedule	CIFAR-10	0%	100%	(55.99 \pm 0.11)%
		10%	100%	(51.72 \pm 0.10)%
		20%	100%	(47.77 \pm 0.11)%
		30%	100%	(43.60 \pm 0.13)%
		40%	100%	(38.84 \pm 0.15)%
		50%	100%	(34.04 \pm 0.14)%
		60%	100%	(29.01 \pm 0.08)%
		70%	100%	(23.94 \pm 0.08)%
		80%	100%	(19.06 \pm 0.09)%
miniAlexNet {c 300 5 \times 5, p 3 \times 3, r, c 150 5 \times 5, p 3 \times 3, r, f, d 384, d 192, d 10} [128]	CIFAR-10	0%	100%	78.5%
		20%	100%	66.4%
		40%	100%	49.8%
		100%	100%	10.2%
alexnet [44] (torch)	imagenet	0%		56.5%
alexnet [44] (Matlab)		0%		57.1%*
vgg19 [213] (tensorflow)	imagenet	0%		71.8%*
vgg19 [213] (torch)		0%		72.4%

Label noise

To study how noise is stored in neural networks in comparison to information about an underlying rule, we train several networks in Sec. 3.3 with label noise. For this, we assign random labels to a certain fraction of the training data such that the network has to overfit and memorize these images to still achieve 100% training accuracy. We observe that both MLP and miniAlexNet networks are capable of perfectly learning the training data even in the case of 100% label noise, in agreement with results by Zhang et al. [39]. Even when training with small amounts of label noise, the generalization performance is significantly worse than for training with a smaller dataset where the noisy image-label pairs are removed. This suggests that noise also negatively influences the trained weights, an effect we try to revert by noise filtering in Sec. 3.3.

Performance of neural networks

In Tab. 3.1, we show the training and generalization accuracies for networks trained as described above. For MLP networks, we average the accuracies for ten different realizations of the initial weights by using different seeds for the pseudo random number generator for initialization, and additionally provide the error of the mean. For the large, pre-trained networks accuracies are measured on the much larger imagenet dataset. We see that the convolutional miniAlexNet networks generalize significantly better than fully connected networks of the MLP architecture. In addition, label noise significantly reduces the performance for both architectures.

3.1.3 Singular value decomposition of weights

In the following, we study the singular values and vectors of neural network weight matrices: any real matrix $W^{(l)} \in \mathcal{M}_{n_l \times n_{l-1}}$ with $n = \min(n_l, n_{l-1})$ can be decomposed into two orthogonal matrices $U^{(l)} \in O_{n_l \times n_l}$, $V^{(l)} \in O_{n_{l-1} \times n_{l-1}}$, and a diagonal matrix $\Sigma^{(l)} \in D_{n_l \times n_{l-1}}$ with

$$\Sigma_{ij}^{(l)} = \begin{cases} \delta_{ij} \nu_i & \text{for } i \leq n \\ 0 & \text{else} \end{cases} \quad (3.28)$$

such that

$$W^{(l)} = U^{(l)} \Sigma^{(l)} V^{(l)T}. \quad (3.29)$$

Here, ν_i are the singular values that we can always choose as real, non-negative numbers, $\nu_i \geq 0$. In the case $n_l \leq n_{l-1}$, the positive semi-definite matrix $W^{(l)} W^{(l)T}$ has eigenvalues ν_i^2 and eigenvectors $U^{(l)}$ as

$$W^{(l)} W^{(l)T} = U^{(l)} \Sigma^{(l)} V^{(l)T} V^{(l)} \Sigma^{(l)} U^{(l)T} = U^{(l)} [\Sigma^{(l)}]^2 U^{(l)T}, \quad (3.30)$$

such that we focus on the left singular values $U^{(l)}$ in this case. In the opposite case, where $n_l > n_{l-1}$, $W^{(l)T} W^{(l)}$ has eigenvalues ν_i^2 and eigenvectors $V^{(l)}$, such that for $n_l > n_{l-1}$ we focus on the right singular vectors. With this choice, we make sure that there are no parts of the singular vectors that cannot contribute to the weight $W^{(l)}$ due to the zero entries in $\Sigma^{(l)}$. One further defines the rank of the weight $W^{(l)}$ as the number of non-zero singular values and says that $W^{(l)}$ has full rank if this number agrees with $n = \min(n_l, n_{l+1})$.

Because the weights of fully connected layers are already $n_l \times n_{l-1}$ matrices, we can straight forwardly compute the singular values and vectors as described above. The situation is different for convolutional layers, where weights are 4D tensors $W^{(l)} \in \mathcal{M}_{(d_1 \times d_2 \times c_{l-1}) \times n_l}$ such that we first need to reshape the weights before we can compute a singular value decomposition [443]. Other methods for computing a singular value decomposition are discussed by Sedghi et al. [444]. Here, we choose reshaping as this (i) results in a matrix that is large enough to perform a statistical analysis on the singular values and vectors, and (ii) can easily be inverted which allows us to modify the singular values, reconstruct the weights, and reinsert them into the network. There are many ways to reshape a 4D tensor into a 2D matrix, however, we like to keep weight entries that are close to each other in the tensor also close in the matrix [328]. As the extension of the filter d_1, d_2 is usually much smaller than the number of feature maps n_l , we first flatten along d_1 and d_2 and then along the smaller dimension of c_{l-1} and n_l . This allows us to keep the larger dimension $\max(c_{l-1}, n_l)$ separately.

For the miniAlexNet architecture, we hence reshape the first convolutional layer $l = 0$: $\mathcal{M}_{d_1 \times d_2 \times 3 \times n_0} \rightarrow \mathcal{M}_{(d_1 d_2 \cdot 3) \times n_0}$ according to

$$\tilde{W}_{i_1+i_2 d_1+i_3 d_1 d_2, i_4}^{(0)} = W_{i_1, i_2, i_3, i_4}^{(0)} , \quad (3.31)$$

and the second convolutional layer $l = 3$: $\mathcal{M}_{d_1 \times d_2 \times 300 \times n_3} \rightarrow \mathcal{M}_{(d_1 d_2 n_3) \times 300}$ according to

$$\tilde{W}_{i_1+i_2 d_1+i_4 d_1 d_2, i_3}^{(3)} = W_{i_1, i_2, i_3, i_4}^{(3)} . \quad (3.32)$$

3.2 Random matrix theory as a zero-information hypothesis

Parts of this section closely follow the publication: Matthias Thamm¹, Max Staats¹, and Bernd Rosenow, *Random matrix analysis of deep neural network weight matrices*, Physical Review E 106, 054124 (2022) [328].

In this section, we study weights of trained neural networks using predictions from random matrix theory (RMT). Because neural network weights are randomly initialized, they perfectly match RMT predictions before training, and after training, we attribute deviations from RMT predictions as indication of learned information. Following this *zero-information hypothesis* ansatz, we compare RMT predictions with empirical results for trained deep neural networks to locate stored information in the singular value decomposition of their weights: From the excellent agreement of the level spacing distribution and the level number variance of unfolded singular value spectra with RMT, we conclude that the bulk of the spectra is random. In addition to these universal predictions, we also directly compare singular values to the Marcenko-Pastur distribution and singular vector entries to the Porter-Thomas distribution, which shows that there are some large singular values and corresponding vectors that deviate from RMT predictions which we attribute to stored information in this part of the spectrum [37, 328].

We further study networks trained in different learning regimes. We find that the universal properties remain unchanged, indicating that the bulk of the spectrum is random in all cases. However, the non-universal RMT predictions allow for distinguishing lazy learning, where weights barely change during training, from rich learning for which training significantly updates the weights. We find the best generalization performance in between the lazy and rich learning regimes.

It has been suggested that the singular value spectra are heavy tailed, often with a power-law decay [37, 238]. We perform a careful analysis of the tails of singular value spectra for large pre-trained networks, however, we do not find power-laws in most of the cases.

3.2.1 Interpretation of Kolmogorov-Smirnov tests

It can be quite subtle to decide whether a sample of data points is from a given distribution – a problem we need to solve for comparing empirical results from trained weight matrices to RMT predictions, which are usually formulated in terms of the distribution of the random variables. The approach we choose here is to perform Kolmogorov-Smirnov tests, which measure the deviations of the empirical cumulative distribution function (cdf) distributed according to the test distribution and the probability p – the so-called p -value – for obtaining at least as much deviation when drawing randomly from the test distribution. For samples actually following the test distribution, the p -value is uniformly distributed in $[0, 1]$, and if samples from the unknown distribution tend to have small p -values, it is an indication that they are not from the test distribution. Of course for a single sample, a large p -value does not mean that the sample has to be from the distribution, and a small p -value does not tell with certainty that it is not from the test distribution. One therefore considers a confidence interval such that for $p \leq 0.05$ we can be 95% sure that the sample is not from the distribution and therefore reject the null hypothesis.

¹equal author contribution

Computation of Kolmogorov-Smirnov test statistics

We describe in the following how the test statistics that determine the p -value can be obtained given a sample $\boldsymbol{\xi}$ of length N to be tested for a distribution with cdf $C(\boldsymbol{x})$. In the case where we do not know a closed expression for this cumulative test distribution but where we can generate samples from the distribution, for example because they are computed from random matrices, we can approximate $C(\boldsymbol{x})$ using Monte-Carlo sampling:

- i) Generate a number of M (for example 50000) data points X_k that follow the test distribution.
- ii) Rank order the samples such that $X_{i_k} \leq X_{i_{k+1}}$.
- iii) Approximate the cumulative distribution with

$$C(X) = \sum_{X_k \leq X} \frac{1}{M} , \quad (3.33)$$

such that $C(X_k) = i_k/M$.

With the expression for the cumulative test distribution $C(x)$, we can then use the following algorithm to compute the Kolmogorov-Smirnov test statistics [445]:

- i) Draw a sample $\boldsymbol{X}^{(i)}$ of size N from the test distribution, i.e. the hypothesis distribution of the Kolmogorov-Smirnov test.
- ii) Get the empirical cdf for this sample i as

$$\hat{C}^{(i)}(X) = \sum_{X_k^{(i)} \leq X} \frac{1}{N} . \quad (3.34)$$

- iii) Compute the Kolmogorov-Smirnov distance for the sample

$$D_{\text{KS}}^{(i)} = \max |\hat{C}^{(i)}(X^{(i)}) - C(X^{(i)})| . \quad (3.35)$$

- iv) Repeat steps a-c about $M = 50000$ times to get a vector of Kolmogorov-Smirnov distances $\boldsymbol{D}_{\text{KS}}$.
- v) Compute the cumulative distribution of the Kolmogorov-Smirnov distance when drawing from the test distribution

$$\bar{C}_{\text{KS}}(D) = \sum_{D_{\text{KS}} < D} \frac{1}{M} . \quad (3.36)$$

- vi) Define the p -value as

$$p(D) = 1 - \bar{C}_{\text{KS}}(D) , \quad (3.37)$$

i.e. the probability to get a Kolmogorov-Smirnov distance at least as large as D when drawing from the test distribution.

Performing the Kolmogorov-Smirnov test

With the Kolmogorov-Smirnov statistics, we can compute the p -value for a sample ξ of length N and decide whether we accept or reject the null hypothesis that the sample is from the test distribution $C(X)$:

- i) Compute the empirical cdf of the sample $\hat{C}_\xi(X) = \sum_{\xi < X} 1/N$.
- ii) Compute the Kolmogorov-Smirnov distance for the sample $D_\xi = \max |\hat{C}_\xi(\xi) - C(\xi)|$.
- iii) Compute the p -value for the sample using the Kolmogorov-Smirnov statistics, $p = 1 - \bar{C}_{KS}(D_\xi)$.
- iv) If $p \leq 0.05$ reject the hypothesis, i.e., ξ is likely not from the test distribution; else accept the null hypothesis.

In the following, we will use Kolmogorov-Smirnov tests to compare level spacings of unfolded singular values to the Wigner surmise, singular vector entries to the Porter-Thomas distribution, and to test tails of singular value spectra for power-laws.

3.2.2 Universal level spacing statistics and level number variance

We first consider universal properties of random matrices that are independent of the particular realization such that no additional fit parameters are necessary. To obtain such a universal property, we compute the $n = \min(n_l, n_{l-1})$ singular values ν_i of a matrix $W^{(l)} \in \mathcal{M}_{n_l, n_{l-1}}$, and then normalize the mean density of the singular values to unity, yielding the unfolded singular values ξ_i [7, 11, 12, 19, 182]. For this unfolding, we first need to estimate the probability density of the singular values $P(\nu)$ which can be achieved by a technique called Gaussian broadening [182, 446]: We approximate $P(\nu)$ as a sum of Gaussian functions each centered around one of the singular values ν_k with width depending on the density of singular values around ν_k , given by $\sigma_k = (\nu_{k+a} - \nu_{k-a})/2$,

$$P(\nu) \approx \frac{1}{n} \sum_{k=1}^n \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{(\nu - \nu_k)^2}{2\sigma_k^2}\right). \quad (3.38)$$

Here, a is called the broadening window size, and we use $a = 15$ in the following [30]. From the estimate for the probability density, we obtain the cumulative distribution

$$F(\nu) = n \int_{-\infty}^{\nu} P(x) dx, \quad (3.39)$$

which then defines the unfolded spectrum as

$$\xi_i = F(\nu_i). \quad (3.40)$$

As discussed in the introduction, a universal property obtained from the unfolded and sorted spectrum, $\xi_{i+1} > \xi_i$, is the distribution of level spacings

$$s_i = \xi_{i+1} - \xi_i, \quad (3.41)$$

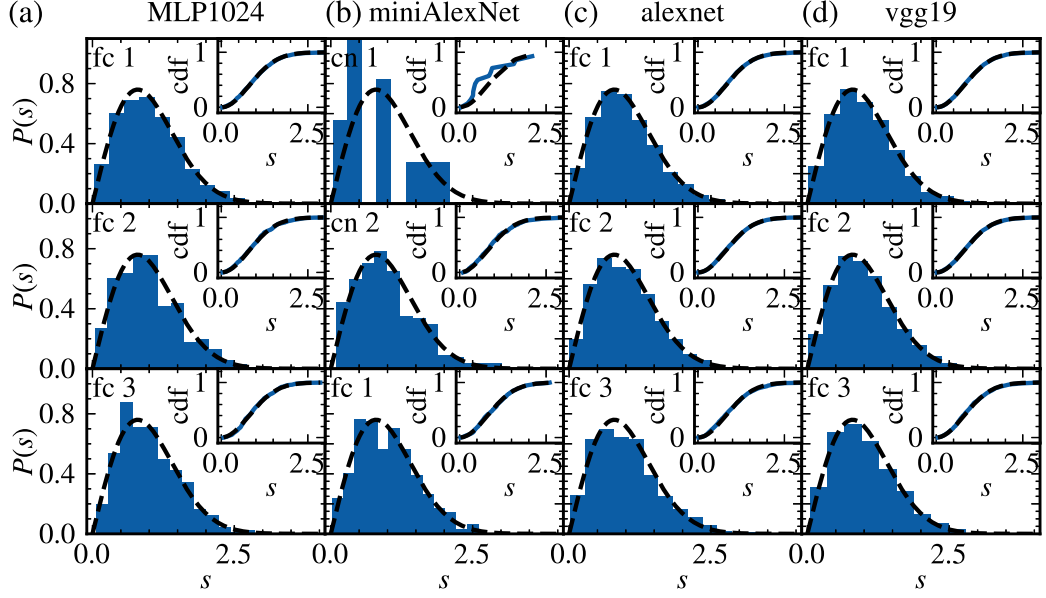


Figure 3.3: Nearest neighbor spacing distributions of unfolded singular values of weight matrices for (a) MLP1024, (b) miniAlexNet, (c) alexnet, and (d) vgg19. The main panels depict the probability density histograms and the insets show the cumulative distribution functions. In addition, we depict the Wigner surmise theory prediction Eq. (3.42) for random matrices with dashed, black lines. Except for the first convolutional layer of miniAlexNet which has only 75 singular values, we find excellent visual agreement with the RMT predictions in all classes. This is further supported by Kolmogorov-Smirnov tests which cannot reject the null hypothesis as shown in Tab. 3.2.

which for real, random matrices from the GOE universality class is well described by the Wigner surmise [1, 7, 11, 12, 19, 182]

$$P_{\text{WS}}(s) = \frac{\pi s}{2} \exp\left(-\frac{\pi}{4} s^2\right). \quad (3.42)$$

A more general form is given by the Brody distribution [11, 12, 19]

$$P_{\text{Br}}(s) = B(1 + \beta)s^\beta \exp(-Bs^{1+\beta}) \text{ with } B = \{\Gamma([\beta + 2]/[\beta + 1])\}^{1+\beta}, \quad (3.43)$$

which reduces to the Wigner surmise in the limit $\beta \rightarrow 1$. In Fig. 3.3, we show the level spacing distribution for trained weights of several layers from various networks of different architecture (blue histograms), and the corresponding empirical cdfs (insets, blue) together with the RMT prediction, i.e. the Wigner surmise (black). We find excellent visual agreement with the RMT predictions in all cases, except for the first convolutional layer in miniAlexNet (upper panel in b) where the spectrum consists of only 75 values. This visual agreement is also confirmed by Kolmogorov-Smirnov tests for the Wigner surmise and fits to the Brody distribution, where $p \geq 0.05$ and $\beta \sim 1$ (see Tab. 3.2). To obtain the values of β and the errors, we use bootstrap sampling from the spectra [447–449], fit the Brody distribution to each sample, and then calculate β as the mean of the fit results and the error of the mean from the standard deviations. Here, bootstrap sampling means drawing n values from the ν_i with equal probability and with repetition. We conclude that the bulk of the singular value spectra of trained weights – even for large, modern

Table 3.2: Kolmogorov-Smirnov test results of the distribution of unfolded singular value spacings of the weight matrices against the Brody distribution with $\beta = 1$. Rejection of the null hypothesis is based on the $\alpha = 0.05$ significance level. The p -value indicates how likely it is to obtain a distribution with at least as much cumulative density function deviation as the one tested for drawing random numbers from a Brody distribution with $\beta = 1$. In addition, we show the results of a fit of a Brody distribution with free parameter β to the cumulative density function of the computed level spacings. The error was determined by a bootstrap sampling method. We find excellent agreement with the Wigner surmise for a variety of network architectures. (Parts of the results have been published in [328].)

network	ks-test p-value			Brody β from fit		
	layer 1	layer 2	layer 3	layer 1	layer 2	layer 3
MLP512 (hidden layers)	0.347	0.401	0.812	0.79 ± 0.09	1.01 ± 0.10	1.04 ± 0.12
MLP1024 (hidden layers)	0.799	0.812	0.792	0.94 ± 0.07	0.91 ± 0.11	1.03 ± 0.10
miniAlexNet (conv. 1, conv. 2, fc1)	0.025	0.859	0.743	0.16 ± 0.21	0.85 ± 0.14	0.95 ± 0.16
alexnet (dense layers) – pytorch	0.670	0.229	0.160	0.96 ± 0.04	0.95 ± 0.04	0.83 ± 0.07
vgg19 (dense layers) – tensorflow	0.376	0.221	0.969	0.95 ± 0.04	0.92 ± 0.04	0.97 ± 0.08
vgg19 (dense layers) – pytorch	0.376	0.652	0.557	0.97 ± 0.04	0.95 ± 0.04	0.92 ± 0.07

networks – follows the RMT predictions. In addition, we consider a second universal property that probes the long-range correlations between singular values – the level number variance [7, 11, 12, 182]

$$\Sigma^2(l) = \langle (N_\xi(l) - l)^2 \rangle_\xi , \quad (3.44)$$

i.e. the variance in the number of unfolded singular values N_ξ in intervals of length l . This quantity is more sensitive to details of the singular value distribution than the level spacings that only probe the spectrum locally. To compute the level number variance from the unfolded spectrum, for each l , we repeatedly draw values $\xi_0 \in [\min \xi_i + l/2, \max \xi_i - l/2]$ and count the number N_{ξ_0} of singular values in the interval $[\xi_0 - l/2, \xi_0 + l/2]$. We update the running-variance for each drawn value ξ_0 and stop only after the variance has converged [450].

For random matrices from the GOE, for large l , the level number variance is given by [7, 11, 12, 182]

$$\Sigma_{\text{GOE}}^2(l) \approx \frac{2}{\pi^2} \left[\ln(2\pi l) + \gamma + 1 - \frac{\pi^2}{8} \right] , \quad (3.45)$$

where γ is the Euler-Mascheroni constant. It was shown that this is also a good approximation in the case of GOE matrices for smaller l [182], but the level number variance also depends on the window size a used for broadening [33]. For the empirical weights, we do not consider level number variances for $l \gg a$, as the unfolding protocol breaks down if the window size a is much larger than the averaging window l [328]. However, the important information is whether $\Sigma^2(L)$ initially grows linearly – as for random matrices – or faster, e.g. exponentially for the case of an uncorrelated, diagonal random matrix.

For all the networks considered here, we find that the level-number variance for trained weights grows only logarithmically (blue lines in Fig. 3.4), and in particular there are only small differences to the initialized, random weights (red lines) [328]. These results for the universal properties suggest that the weights are predominately random.

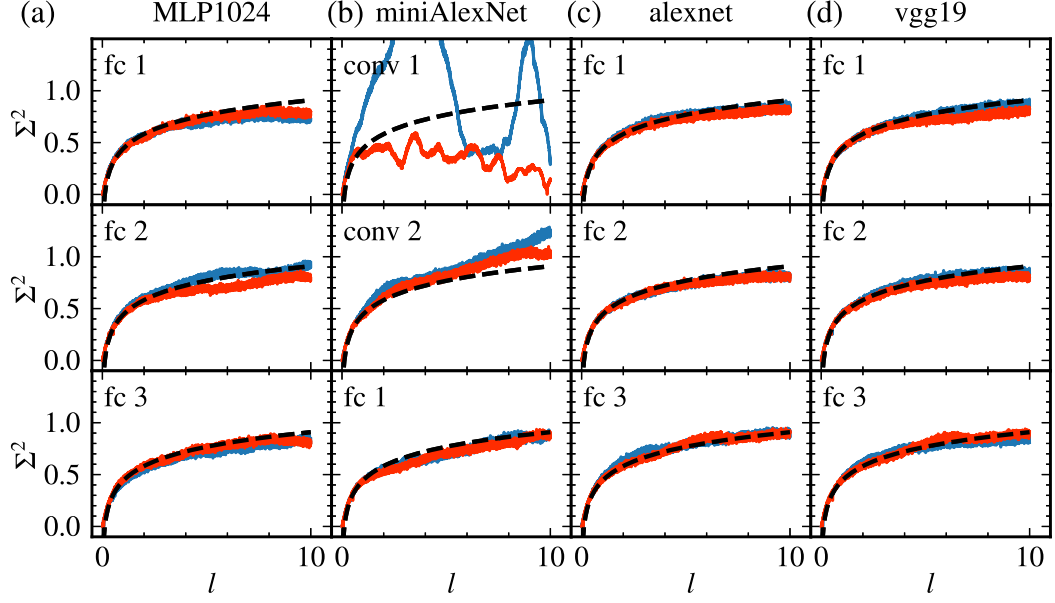


Figure 3.4: Level number variance of unfolded singular values of the weights for (a) MLP1024, (b) miniAlexNet, (c) alexnet, and (d) vgg19. Red curves show the results for initialized, random weights, and blue lines depict level number variances for fully trained weights. The dashed, black lines depict the theory prediction, Eq. (3.45), for the GOE. Except for the first convolutional layer of miniAlexNet which has only 75 singular values, the level number variance grows logarithmically in all cases – even after training. Particularly for large networks in (c) and (d), where the statistics are most reliable, deviations from the RMT prediction are small [328].

3.2.3 Singular values and vectors: Marcenko-Pastur and Porter-Thomas distribution

A comparison of trained weights with universal RMT predictions indicates randomness in the bulk of the weights. To study how the networks can nevertheless achieve good generalization accuracies, we now focus on non-universal properties to locate the parts of the weights that contain the relevant information about the underlying rule in the training data. For this purpose, we first consider the singular value spectra ν of the weight matrices, which in the case of a random weight $W \in \mathcal{M}_{m,n}$ (with $n < m$) of i.i.d. normal distributed entries with standard deviation σ is given by the Marcenko-Pastur distribution [184, 185, 451, 451]

$$P_{\text{MP}}(\nu) = \begin{cases} \frac{n/m}{\pi \tilde{\sigma}^2 \nu} \sqrt{(\nu_+^2 - \nu^2)(\nu^2 - \nu_-^2)} & \nu \in [\nu_-, \nu_+] \\ 0 & \text{else} \end{cases} \quad (3.46)$$

with $\nu_{\pm} = \tilde{\sigma}(1 \pm \sqrt{n/m})$ and $\tilde{\sigma} = \sqrt{m}\sigma$ (see also Sec. 1.1.4). In Fig. 3.5, we show the singular value spectra obtained with Gaussian broadening with window size $a = 15$ according to Eq. (3.38). For initialized, random weights (red), we find excellent agreement with the Marcenko-Pastur prediction (black, dashed) in all cases. Even after training (blue) significant parts of the spectrum remain Marcenko-Pastur like, and in addition there are larger singular values that form a heavy tail as reported in Refs. [37, 238] (for a detailed analysis of the tail distribution, see Sec. 3.2.5). In the case where weights are trained with L_2 regularization (panels c, d, and fc1 in panel b), the Marcenko-Pastur curve is slightly shifted towards smaller values during

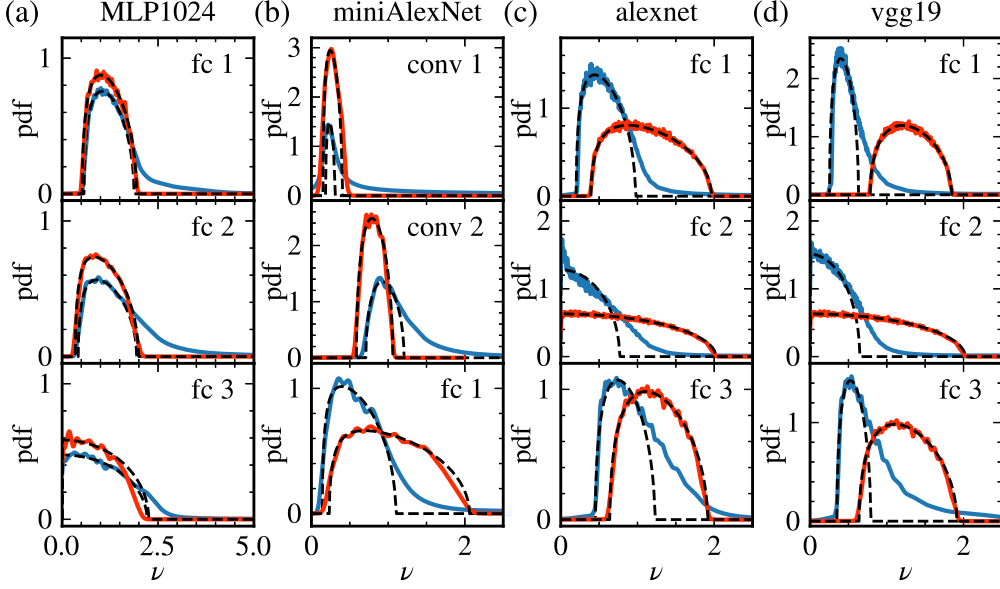


Figure 3.5: Distribution of the singular values ν of the weight matrices of (a) the fully connected three hidden layer network MLP1024, (b) the convolutional neural network miniAlexNet, (c) the pre-trained `pytorch` [434] model of alexnet, and (d) the `pytorch` model of vgg19. We show results for the dense, fully connected layers (fc) and in the case of miniAlexNet also for the two convolutional layers (conv). The spectral distributions are calculated by broadening with a window size of $a = 15$ singular values [30, 328]. The red curves show the distribution directly after random Glorot initialization, and the blue lines depict the result after fully training the networks. The dashed, black lines are fits to the Marcenko-Pastur distribution, Eq. (3.46). After random initialization, the spectra agree well with the RMT predictions, and even after training the bulk of the singular values still follows a two-parameter Marcenko-Pastur distribution. While for layers trained without L_2 regularization, the range of Marcenko-Pastur distributed singular values before and after training are comparable, when training with L_2 regularization (c, d, and fc1 in b) singular values of the bulk are additionally shifted towards smaller values.

training. We fit a two-parameter adjusted Marcenko-Pastur distribution (black, dashed) for the trained spectra by fixing ν_- as the smallest singular value and leaving ν_+ and the pre-factor in Eq. (3.46) as fit parameters to account for the fact that some weight moved into the tail such that the total weight in the bulk is less than one. When training without L_2 regularization, it becomes apparent that the Marcenko-Pastur region $[\nu_-, \nu_+]$ remains unchanged and only some large singular values develop without significantly influencing the random bulk. In all cases, most of the singular values remain within the Marcenko-Pastur region as found from the fits.

In addition, we consider the entries of singular vectors of length n which for random vectors would follow a Porter-Thomas distribution, i.e. a Gaussian with zero mean and a standard deviation of $1/\sqrt{m}$ fixed by the normalization of the vectors (see also Sec. 1.1.4). For a weight matrix $W \in \mathcal{M}_{m,n}$ with $m > n$, we consider the right singular vectors V , i.e. the eigenvectors of $W^T W$, and if $m < n$, we focus on the left singular values U , the eigenvectors of $W W^T$. This ensures that the considered matrix $W W^T$ or $W^T W$ has full rank.

To test the entries of singular vectors for a Porter-Thomas distribution, we order the vectors such that the corresponding singular values are rank ordered, $\nu_{i+1} > \nu_i$, and then perform

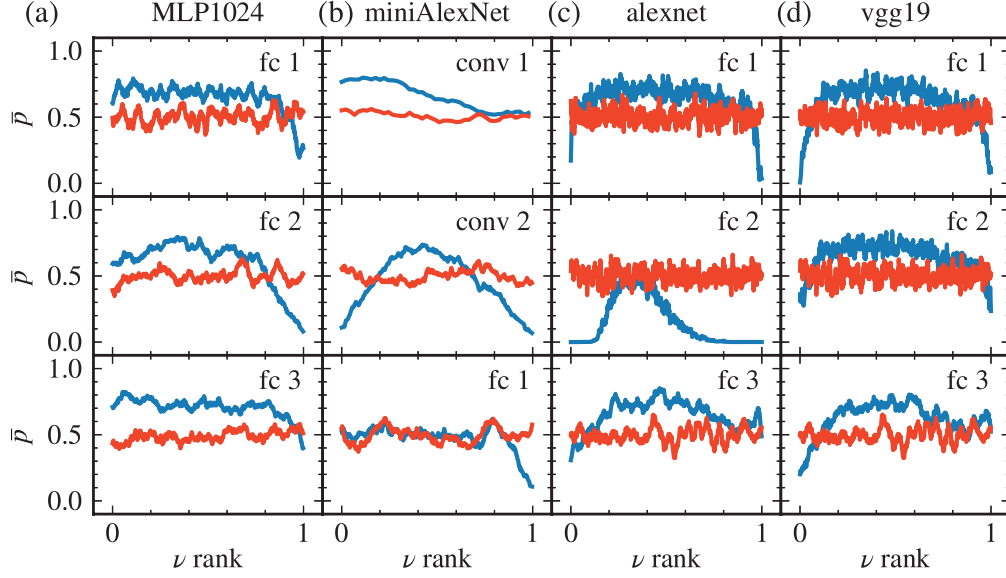


Figure 3.6: Randomness of eigenvector entries as a function of the singular value position in the spectrum. For this, we show p -values of Kolmogorov-Smirnov tests of the singular vector entries against a Porter-Thomas distribution which is the prediction for random matrices. On the x axis, we plot the positions according to the rank ordered singular values, such that 0 corresponds to the smallest and 1 to the largest singular value of each weight matrix. The resulting p -values are averaged over neighboring singular values with a window size of 15 [328]. We show results for random matrices (red) and trained weights (blue) for (a) MLP1024, (b) miniAlexNet, (c) alexnet, and (d) vgg19. For the random vectors that perfectly follow the RMT prediction, the p -values are uniformly distributed in $[0, 1]$ such that the average p -value is 0.5. We find that p -values corresponding to large singular values are reduced compared to the values in the bulk. Additionally, for some layers, there are reduced p -values for vectors corresponding to very small singular values.

Kolmogorov-Smirnov tests for the entries of each singular vector yielding a p -value, $p(\nu_i)$, corresponding to every singular value ν_i . We then average the p -values for surrounding singular values with a window size of $a = 15$:

$$\bar{p}(\nu_i) = \frac{1}{2a} \sum_{j=1}^a [p(\nu_{i-j}) + p(\nu_{i+j})] . \quad (3.47)$$

If the entries of a vector are indeed from a Porter-Thomas distribution, the p -value is uniformly distributed in $[0, 1]$ such that the mean value \bar{p} is close to 0.5, in excellent agreement with the results for initialized, random weights (red lines in Fig. 3.6). For trained weights (blue), we find that p -values in the bulk form a plateau above 0.5 and in the tail, a clear decrease from the plateau is visible. We show later that the increased p -values on the plateaus, compared to the random case, can occur when a low rank matrix W_0 is added to a random matrix W_{random} and the p -values of the singular vectors of $W = W_0 + W_{\text{random}}$ are considered. In some cases, we also observe significantly reduced average p -values for singular vectors corresponding to very small singular values, a phenomenon we will later also observe in the rich learning regime in Sec. 3.2.4.

We would like to emphasize that using the correct Kolmogorov-Smirnov statistics is crucial to obtain average p -values of 0.5 for singular vectors of random matrices. Here it is important

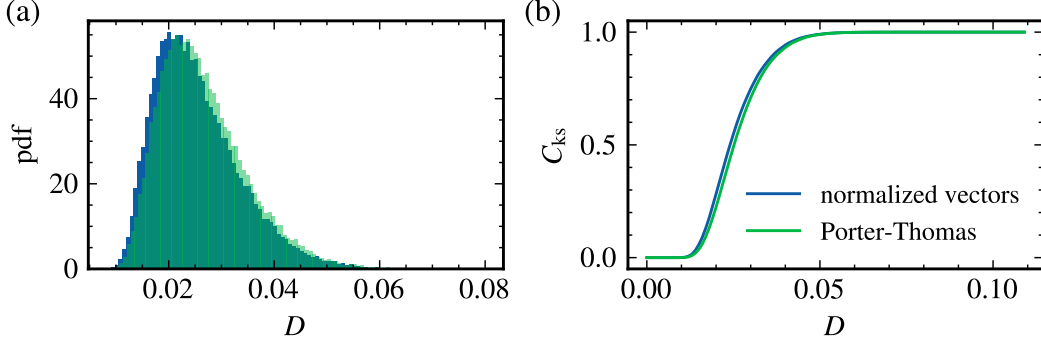


Figure 3.7: Comparison between Kolmogorov-Smirnov test statistics for i.i.d. random vectors of length 512 (green) and for normalized Porter-Thomas vectors (blue). (a) Probability density function of the cdf deviations obtained with Monte-Carlo sampling as described in Sec. (3.2.1). (b) Cumulative distribution functions for the pdfs shown in (a). It becomes apparent that normalizing the vectors significantly changes the Kolmogorov-Smirnov test statistics even though it leaves the cumulative distribution of the vector entries C_{PT} unchanged. (Figure taken from Ref. [329].)

that sampling for computing the statistics as described in Sec. 3.2.1 takes into account the normalization of the singular vectors. If the samples are only drawn from the Porter-Thomas distribution but not normalized in the same way as the singular vectors, there are significant deviations in the p -values (see Fig. 3.7).

As we discussed before, the interpretation of Kolmogorov-Smirnov tests can be tricky and a large p -value alone does not mean that the sample has to be from the test distribution. We therefore additionally consider the inverse participation ratio (IPR) of the entries of a singular vector \mathbf{v} ,

$$\text{IPR}(\mathbf{v}) = \sum_{i=1}^n |v_i|^4. \quad (3.48)$$

This IPR gives some insight into the localization of the vector entries and is approximately given by the inverse of the number of relevant components in \mathbf{v} : for example if a single entry is 1 and all other entries are 0, the IPR is 1. In the other extreme case where all entries are $1/\sqrt{n}$, the IPR is $1/n$. We therefore interpret a small IPR that occurs together with a large p -value as indication of random noise, and a large IPR together with a reduced p -value indicates stored information. In Fig. 3.8, we show the IPRs for the same layers and networks as in Fig. 3.6. The IPR for initialized random weights (red) is consistent with this interpretation and also for trained weights (blue) the IPR is in agreement with the p -values, such that regions with vectors that are more localized (higher IPR) also come with smaller average p -values and large p -values appear together with small IPRs and for random matrices.

Hence, using the RMT predictions as a zero-information hypothesis indicates that the weights consist of a random bulk and information about the rule is stored in the largest singular values and corresponding vectors only.

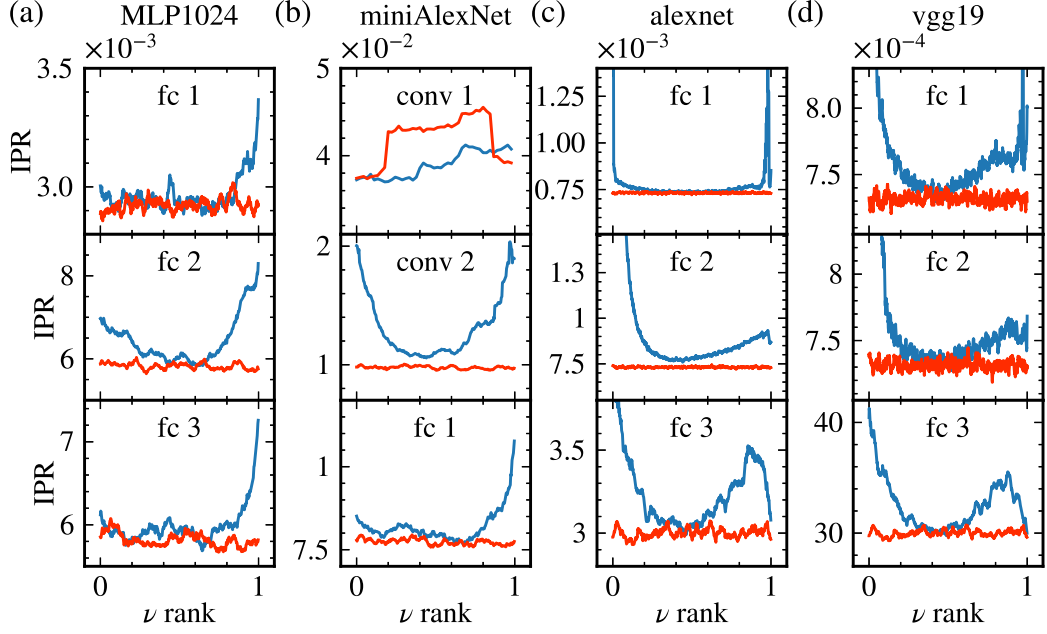


Figure 3.8: Inverse participation ratios (IPRs) of the singular vectors for random weights (red) and trained weights (blue) in (a) MLP1024, (b) miniAlexNet, (c) alexnet, and (d) vgg19. Singular vectors corresponding to large singular values have higher IPRs, i.e. they are more localized, in agreement to the smaller p -values in Fig. 3.6. For some layers, additionally singular vectors for small singular values have large IPRs, however, the bulk agrees with the random control for most of the weights.

3.2.4 Random matrix analysis in different learning regimes

This subsection closely follows the publication: Matthias Thamm, Max Staats, and Bernd Rosenow, *Random matrix analysis of deep neural network weight matrices*, Physical Review E 106, 054124 (2022) [328].

It was shown [128, 224, 452–455] that neural networks can achieve good generalization accuracies even when their weights change only by very small amounts during training. The opposite to this *lazy learning* is denoted as *rich learning*, where the final weights \mathbf{W} after training deviate significantly from the initial ones \mathbf{W}_0 . Based on this, a criterion for estimating the learning regime was proposed by Chizat et al. [456]: For a neural network $f_{\mathbf{W}}$ that maps an input \mathbf{x} to an output, and an accuracy function $\mathcal{A}(f_{\mathbf{W}}, \{\mathbf{x}\}, \{\mathbf{y}\})$, where $\{\mathbf{x}\}$ is a dataset with labels $\{\mathbf{y}\}$, one computes the network's linearization around the initial weights \mathbf{W}_0

$$\tilde{f}_{\mathbf{W}}(\mathbf{x}) = f_{\mathbf{W}_0}(\mathbf{x}) + (\mathbf{W} - \mathbf{W}_0) \cdot \nabla_{\mathbf{W}} f_{\mathbf{W}}|_{\mathbf{W}_0}(\mathbf{x}) . \quad (3.49)$$

In the lazy learning regime, where $\mathbf{W} \approx \mathbf{W}_0$, linearization is a good approximation such that the accuracies are barely different, i.e.

$$\mathcal{A}(f_{\mathbf{W}}, \{\mathbf{x}\}, \{\mathbf{y}\}) \approx \mathcal{A}(\tilde{f}_{\mathbf{W}}, \{\mathbf{x}\}, \{\mathbf{y}\}) . \quad (3.50)$$

On the contrary, in the rich learning regime, one expects significant deviations such that

$$\mathcal{A}(f_{\mathbf{W}}, \{\mathbf{x}\}, \{\mathbf{y}\}) \gg \mathcal{A}(\tilde{f}_{\mathbf{W}}, \{\mathbf{x}\}, \{\mathbf{y}\}) . \quad (3.51)$$

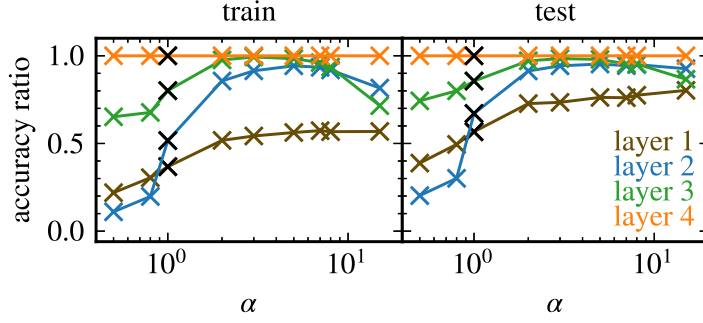


Figure 3.9: Comparison of training and test accuracies for full MLP512 networks and linearized networks around the initial weights of several layers (first layer: brown, second layer: blue, third layer: green, output layer: orange) as a function of the laziness hyperparameter α . Crosses show the ratio between linearized and full network accuracy and lines are a guide to the eye. The black crosses indicate accuracy ratios for typical training with $\alpha = 1$.

This criterion has the advantage that it can also be studied on a layer-wise basis by linearizing around a single weight matrix only, and as accuracies are in the range $[0, 1]$, it gives a scale for laziness comparable between different network architectures. A disadvantage is that it requires to compute the linearization which can be resource intensive for large networks. For obtaining $\tilde{f}_{\mathbf{W}}$, we use the `autodiff` implementation in the `jax` python package together with the `neural_tangents` package [457]. We train several MLP512 networks, where laziness is controlled by introducing a hyperparameter α that modifies the output activations via [456]

$$a_L = \text{softmax}(\alpha(\mathbf{W}_L \mathbf{a}_{L-1} + \mathbf{b}_L)) \quad (3.52)$$

and the cost function as

$$l(\mathbf{W}, \mathbf{b}) = -\frac{1}{N\alpha^2} \sum_{k=1}^N \mathbf{y}^{(k)} \cdot \ln(\mathbf{a}_{\text{out}}^{(k)}) . \quad (3.53)$$

Here, a large $\alpha > 1$ scales down the gradient updates and therefore encourages lazy learning, while small $\alpha < 1$ steers training towards the rich learning regime [456].

In the following, we focus on the MLP512 architecture, for which we show the ratio between accuracies of linearized and full networks as a function of the laziness α in Fig. 3.9. It becomes apparent that layers closer to the input tend more towards the rich learning regime (smaller ratios closer to 0), while layers closer to the output are more lazy (ratios closer to 1). The weights of the output layer, which only have rank 10, are always in the lazy training regime. For our training parameters, we find that controlling the laziness works best for the second hidden layer. We therefore focus on this layer in the following; however, the imbalance in the learning regime between the layers could be lifted by using different learning rates per layer.

Fig. 3.10 depicts the training accuracy (left panel) and the test accuracy (right panel) as a function of the laziness parameter α . As expected, the networks are in the rich regime for $\alpha < 1$, where the full networks (blue crosses) perform significantly better than the linearized networks (green circles), while we observe lazy learning for $\alpha > 1$. The network with $\alpha = 1$ (black symbols), lies about in the middle between the two regimes, where we also find the best test accuracy. We therefore denote $\alpha = 1$ as *typical learning*.

A comparison between the RMT analysis in the three regimes, rich $\alpha = 0.5$ (top panel), typical $\alpha = 1$ (middle panel), and lazy $\alpha = 5$ (bottom panel) in Fig. 3.11 reveals:

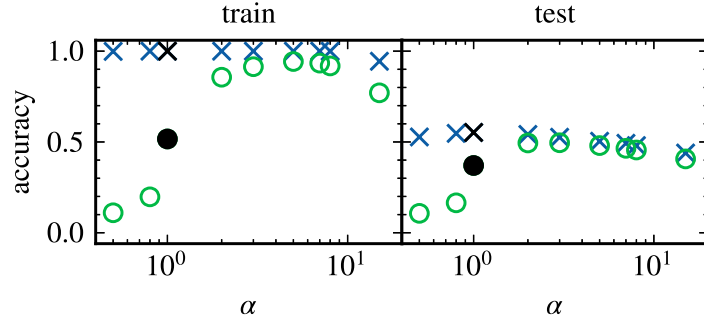


Figure 3.10: Comparison of training and test accuracies for full MLP512 networks (blue crosses) and linearized networks (green circles) around the initial weights of the second layer as a function of the laziness hyperparameter α . The black symbols indicate accuracies for $\alpha = 1$ – similar to the networks trained in the previous sections. For small $\alpha < 1$ accuracies of linearized and full networks deviate significantly which indicates rich learning, while for large $\alpha > 1$ performance differences are small, indicating lazy learning.

- (i) For all networks, the bulks of the spectra are random such that the universal properties, i.e. level spacings (panel b) and level number variance (panel c) agree with RMT predictions. By comparing the level number variance curves for networks trained with various α (not shown), we confirm that the level number variances in panel c display typical deviations from the RMT prediction such that there is no trend of slower growth for networks with larger α .
- (ii) The rich network has more large singular values compared to the typical one, while the lazy network has almost an unchanged Marchenko-Pastur spectrum (panel a). It is therefore surprising that it still achieves a respectable test accuracy of 50.4%, compared to 52.7% for the rich network and 55.2% for the typical network.
- (iii) The p -values for Kolmogorov-Smirnov tests of the eigenvector entries against a Porter-Thomas distribution (panel d) are small only for large singular values in the typical case. In the rich case, we observe small p -values also for singular vectors corresponding to the smallest singular values, and for the lazy network all p -values fluctuate only slightly around 0.5 as expected for random weights.

These findings indicate that networks trained in the lazy regime do not deviate from RMT predictions after training, in striking contrast to rich and typical networks. Thus, an analysis of singular value spectra and singular vector entry distributions can be used to estimate the learning regime of a network on the level of the individual weights, without the need for a potential resource intensive linearization of the networks.

The almost perfect agreement with RMT predictions in the lazy regime raises the question whether the information that allows the network to still generalize quite well is encoded in parts of the spectrum that follow bulk statistics. If this was the case, it would seem impossible to locate this information by means of an RMT analysis. However, we argue in the following that this is in fact not the case for the networks considered here. However, the RMT analysis in the lazy regime faces two major obstacles that make it difficult to detect the information hidden by the dominant random bulk: (i) an individual layer in the lazy network might carry very little information, as it is for example the case for the second hidden layer of the lazy MLP512

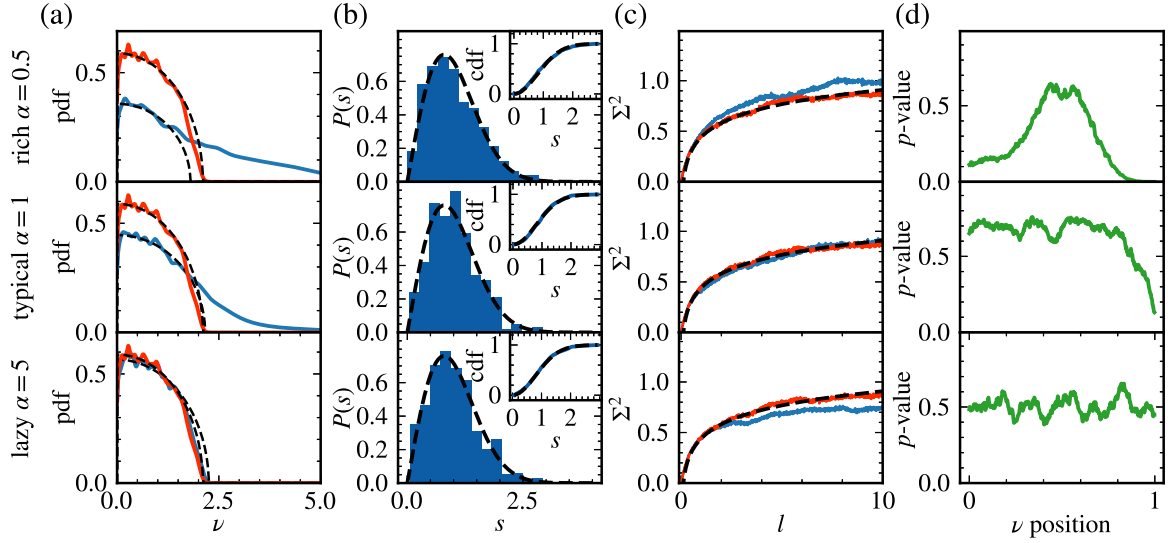


Figure 3.11: Random matrix theory analysis of second layer weights of MLP512 networks trained in different learning regimes: rich learning ($\alpha = 0.5$, top panel), typical learning ($\alpha = 1$, middle panel), and lazy learning ($\alpha = 5$, bottom panel). We show (a) the spectra for trained (blue) and randomly initialized networks (red) together with fits of modified Marcenko-Pastur laws (dashed, black), (b) unfolded level spacing distributions (main panel, blue, window size 15) and corresponding cumulative distributions (insets) together with the Wigner surmise (dashed, black), (c) unfolded level number variance (trained: blue, initialized: red), and (d) p-values for a comparison of singular vector entries to a Porter-Thomas distribution. While trained networks in all cases follow universal RMT predictions (b and c), indicating a random bulk, lazy networks can be distinguished from typical and rich networks by the spectral distributions in (a) and p-values in (d).

network shown in Fig. 3.11, where by replacing the final weight with the initial one for this layer, the generalization accuracy of the network only drops from 50.4% to 42.3%. It is therefore not surprising that no extended region in the spectrum containing information is found by the RMT analysis. (ii) In the lazy case, the difference $\delta W = W - W_0$ between initial weight W_0 and the final weight after training W is small, i.e. $\|\delta W\|/\|W_0\| \ll 1$. In this case, the sensitivity of RMT is not sufficient to detect signatures of δW in W : For instance, the crossover from the Gaussian unitary ensemble to Poisson statistics has been studied in Ref. [458], and it turns out that δW would need to be much larger than in our case in order to have a noticeable effect compared to the bulk statistics from W_0 .

As a solution, we suggest analyzing the statistical properties of the difference matrix δW instead of the full weight matrix. Such an analysis indeed indicates that δW consists of a random bulk and a low-rank contribution that encodes relevant information: In the lazy learning regime the trained weights $W^{(l)}$ remain close to the initial random matrices $W_0^{(l)}$ such that the random bulk dominates any RMT analysis of $W^{(l)}$, and therefore masks small local deviations from the bulk statistics. We thus consider the deviations from the initial weights $\delta W^{(l)} = W^{(l)} - W_0^{(l)}$, which again have random bulk statistics. Their unfolded spectra follow the Wigner surmise (Fig. 3.12a), and one finds a logarithmic increase of the level number variance (panel c). In contrast to the full weight matrix, the difference matrix has a qualitatively different

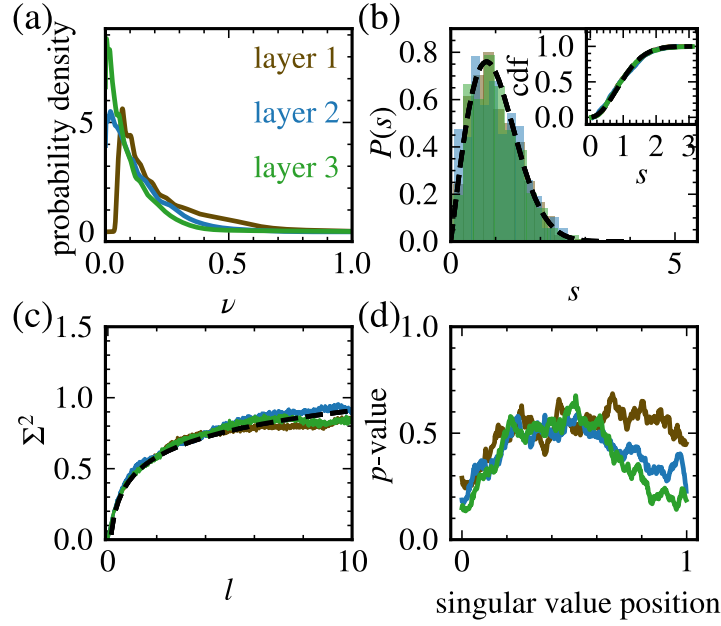


Figure 3.12: Random matrix theory analysis for the difference $\delta W^{(l)} = W^{(l)} - W_0^{(l)}$ between the trained weights $W^{(l)}$ and the initial weights $W_0^{(l)}$ for the lazy MLP512 network trained with $\alpha = 5$. We show results for the first hidden layer (brown), second hidden layer (blue), and third hidden layer (green). Panel (a) depicts the singular value spectra obtained via Gaussian broadening, (b) the level spacing statistics of the unfolded singular value spectra, and (c) the level number variance of the unfolded spectra. Panel (d) shows the averaged p -values for a comparison between singular vector entries and the Porter-Thomas distribution with window size $a = 15$. The spectrum of $\delta W^{(l)}$ agrees well with the Wigner surmise and a logarithmically growing level number variance. However, the p -values in panel (d) indicate that again large parts of the spectrum of $\delta W^{(l)}$ are random and that information is stored in the largest singular values and corresponding vectors of $\delta W^{(l)}$. In addition, the singular value spectrum is not of Marchenko-Pastur type.

distribution function of the singular values (panel a), and the p -values (panel d) are consistent with information being stored in parts of the spectrum corresponding to large singular values, similar to what we found for networks in the typical learning regime.

3.2.5 Tails of weight spectra: Hill estimators and power-law fits

Recently, it has been suggested that the tails of the singular value spectra follow heavy tailed distributions [37, 238], and tail exponents based on power-law fits have been reported [37, 238]. Here, we study the distribution of large singular values using a combination of Hill estimators [459–466] that allow us to visually distinguish power-laws from other heavy tail distributions and power-law fits for which we additionally compute p -values using Kolmogorov-Smirnov tests.

The Hill estimator is obtained by averaging the inverse local slopes of the log-log cumulative distribution and is therefore constant in the case of a power-law $p(x) \propto x^{-\alpha}$. To obtain the Hill estimator, we first rank order the n singular values ν_k such that $\nu_{i_k} \leq \nu_{i_{k+1}}$ and compute the

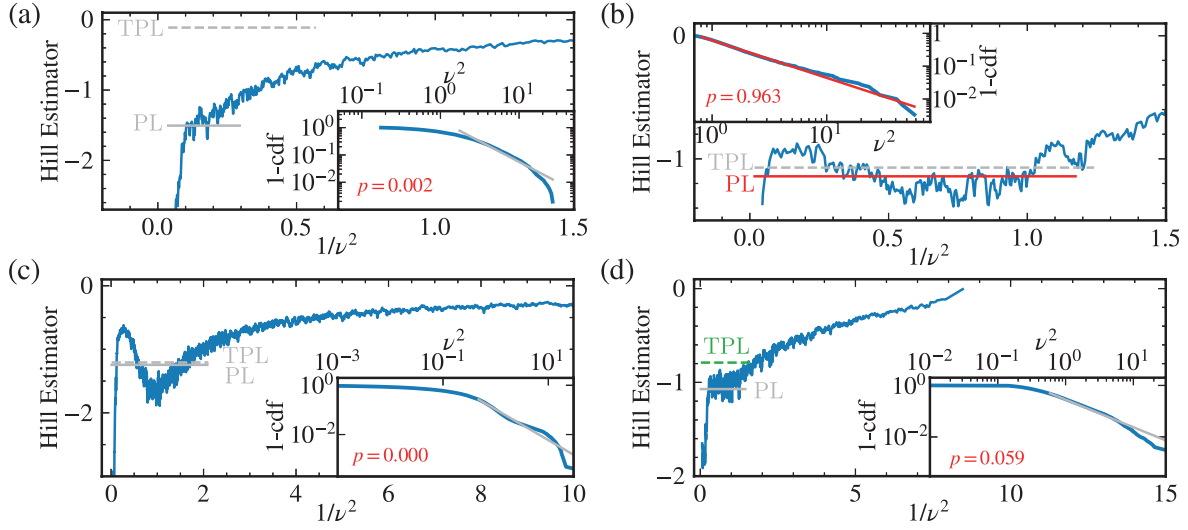


Figure 3.13: Analysis of the cumulative distribution of DNN singular values in the tail region. Hill estimators of the weight matrix spectra for (a) the second layer weight of MLP1024, (b) the second convolutional layer of the CNN miniAlexNet, (c) the second fully connected layer of alexnet, and (d) the third fully connected layer of vgg19. The Hill estimators are obtained using Eq. (3.56) with a window size $a = 20$. The insets depict the corresponding log-log cumulative distributions (blue). In addition, we show results of a power law fit $p(x) \propto x^{-\alpha}$ (solid gray or red) and a truncated power law fit $p(x) \propto x^{-\alpha}e^{-\lambda x}$ (dashed gray or green) to the tails, where a power law tail corresponds to a flat Hill estimator at $-(\alpha - 1)$ starting from ν_{\min} . We show rejected fits in gray and accepted fits colored, based on p-values [445] for the power laws and log-likelihood ratio tests versus truncated power laws (see also Table 3.3). In (b) we cannot reject a power law, however the log-likelihood ratio test tends more towards a truncated power law, but is not statistically reliable ($R = 1.02$ and $p_2 = 0.363$). (Adapted from Ref. [328].)

cumulative distribution of the squared singular values $\lambda_i = \nu_i^2$ as

$$C(\lambda_{i_k}) = \frac{i_k}{n}. \quad (3.54)$$

From this, we get the local inverse slopes

$$\zeta_k = \frac{-\ln[\lambda_{i_{k+1}}/\lambda_{i_k}]}{\ln[C(\lambda_{i_{k+1}})/C(\lambda_{i_k})]} \quad (3.55)$$

that we average over the a surrounding eigenvalues $\frac{1}{a} \sum_{j=-a/2}^{a/2} \lambda_{i_{k+j}}$ to obtain the Hill estimator

$$h(\lambda_k) = \left[\frac{1}{a} \sum_{j=-a/2}^{a/2} \zeta_{k+j} \right]^{-1}. \quad (3.56)$$

In Fig. 3.13, we depict Hill estimators (blue lines in main panels) for various networks. For the dense layers in MLP1024, alexnet, and vgg19 (a, c, d) shown here, the Hill estimators have no extended flat regions that interpolate to $1/\lambda \rightarrow 0$. For the third layer of the `pytorch` vgg19 model (panel d), we find a short flat region followed by a further decay of the Hill estimator

towards $1/\lambda \rightarrow 0$. Hence, we do not find visual evidence for a power-law in the dense layers based on the Hill plots, even though the log-log cdf can look reasonably linear (blue lines in the insets). For the second convolutional layer of miniAlexNet (panel b), there is a longer approximately flat region that persists up to small values of $1/\lambda$, which indicates the presence of a power-law.

To substantiate these findings, we perform power-law fits, $p(\lambda) \propto (\lambda/\lambda_{\min})^{-\alpha}$, to the tails using the `powerlaw` python package [467] to obtain the start of the tail region λ_{\min} and the power-law exponent α . A successful fit however does not mean that there is actually a power-law present in the spectrum, and as the Hill estimators suggest that the tails are not power-law distributed in most cases, we additionally perform Kolmogorov-Smirnov tests to obtain p -values for the confidence of the fits. For computing the Kolmogorov-Smirnov statistics for the fits, we use the method described by Clauset et al. [445]:

- i) We start with a power-law fit to the n eigenvalues to obtain λ_{\min} and α . We denote the number of eigenvalues larger than λ_{\min} as n_{tail} .
- ii) Each sample $\boldsymbol{\lambda}^{(i)}$ for the Kolmogorov-Smirnov statistics is drawn according to
 - a) We draw n random numbers $\mathbf{r}^{(i)}$ uniformly distributed in $[0,1)$ and denote the number of random numbers in $\mathbf{r}^{(i)}$ that satisfy $r_k^{(i)} \leq n_{\text{tail}}/n$ as $\hat{n}_{\text{tail}}^{(i)}$.
 - b) We draw $\hat{n}_{\text{tail}}^{(i)}$ random numbers from the power-law distribution with λ_{\min} and α from step (i) and bootstrap the remaining $n - \hat{n}_{\text{tail}}^{(i)}$ values from the bulk of the original eigenvalues for which $\lambda < \lambda_{\min}$ to construct a sample $\boldsymbol{\lambda}^{(i)}$.
 - c) We then fit a new power-law to this sample $\boldsymbol{\lambda}^{(i)}$ to obtain fit parameters $\lambda_{\min}^{(i)}$ and $\alpha^{(i)}$.
 - d) The relevant Kolmogorov-Smirnov distance is then computed between the empirical cdf of sample $\boldsymbol{\lambda}^{(i)}$ and the cdf of the power-law with $\lambda_{\min}^{(i)}$ and $\alpha^{(i)}$.

This ensures that variations in the fit are also reflected in the Kolmogorov-Smirnov statistics which is crucial for a reliable p -value. Unfortunately, computing the p -values is resource intensive because it requires a power-law fit for each sample and for an accuracy of the p -value of 10^{-2} , about 2500 samples need to be considered [445]. However, obtaining the samples is embarrassingly parallel, i.e., we can obtain the 2500 samples completely independently of each other. This allows us to obtain p -values at least for the weights considered in Fig. 3.13.

But even in the case where a power-law fit is not rejected based on a p -value of $p > 0.05$, there could still be another heavy-tailed distribution that provides a better fit. As suggested in Ref. [37], we also consider truncated power laws (TPL) in the cases where a power-law is not rejected:

To decide which distribution provides a better fit, we perform log-likelihood ratio tests between a truncated power law $p_{\text{TPL}}(x) \propto x^{-\alpha} e^{-\lambda x}$ and a power law $p_{\text{PL}}(x) \propto x^{-\alpha}$. The log-likelihood ratio for n data points in the tail x_i is defined as [445]

$$R = \frac{1}{\sqrt{2n\sigma}} \ln \prod_{i=1}^n \frac{p_{\text{TPL}}(x_i)}{p_{\text{PL}}(x_i)}, \quad (3.57)$$

where σ is the empirical standard deviation of the difference $\ln p_{\text{TPL}}(x_i) - \ln p_{\text{PL}}(x_i)$. A positive sign of R thus indicates that a truncated power law is a better fit, and a negative sign indicates

a better fit for a power law. The p_2 value is then defined as the probability to obtain a ratio with magnitude of at least $|R|$ from a distribution $p(R)$ centered at zero with standard deviation σ , i.e., the probability that the sign is only due to fluctuations. Therefore, small $p_2 < 0.05$ indicate a reliable sign of R , while large p_2 indicate an unreliable sign from fluctuations and hence an inconclusive test.

In Fig. 3.13, we also show the results of power-law fits with solid horizontal lines and the exponent from a truncated power-law fit using dashed lines. We color the lines in gray in cases where we reject the distribution and in red (green) if we cannot reject the power-law (truncated power-law).

Table 3.3: Power law fit results and log-likelihood ratio tests between truncated power law and power law for the dense layers of alexnet, vgg19, and the weights considered in Fig. 3.13. Here, d denotes a dense layer and c a convolutional layer. We reject the power law if $p < 0.05$ or in the case where the two-distribution test favors a truncated power law (positive R , $p_2 < 0.05$) [445]. (Adapted from Ref. [328].)

network	layer	PL fit p	2-distr. R	2-distr. p_2	PL fit α	PL fit x_{\min}	reject PL
MLP1024	d2	0.0019	2.50	0.000	2.51	3.389	yes
miniAlexNet	c2	0.9628	1.02	0.363	2.14	0.852	no
alexnet	d1	0.0004	1.55	0.359	2.29	0.418	yes
	d2	0.0004	1.25	0.211	2.25	0.480	yes
	d3	0.9990	-0.003	0.999	3.02	2.046	no
vgg19	d1	0.0011	2.01	0.142	2.27	0.275	yes
	d2	0.0007	1.98	0.055	2.19	0.291	yes
	d3	0.0590	2.26	0.001	2.07	0.690	yes

In Tab. 3.3, we report the p -values together with fit results for a power-law and log-likelihood tests against a truncated power-law. As already suggested by the Hill plot, the p -values indicate a power-law for the second convolutional layer in miniAlexnet (panel b in Fig. 3.13). For the third dense layer in vgg19 (panel d in Fig. 3.13), a power-law is not rejected as $p > 0.05$; however, a truncated power-law provides a better fit. For the second hidden layer of MLP1024 (panel a) and the second dense layer of alexnet (panel c), we directly reject power-laws based on the p -values. In addition, we consider all dense layers of the `pytorch` models for the large pre-trained networks alexnet and vgg19 (Tab. 3.3). For the third layer, we find agreement with Martin et al. [37] where a power-law for alexnet and a truncated power-law for vgg19 was reported. However, for the other dense layers, which were also reported as power-laws [37], our analysis clearly rejects the power-law fits with p -values much smaller than 0.05.

In summary, we find no evidence that the singular values of DNN weight matrices are generally described by a power law tail distributions, and argue that the exponent resulting from a power law fit to the singular value probability density function can only be viewed as a heuristic tool to characterize different spectra, but not as a genuine property of the tail of the distribution of singular values. In addition, given the absence of power law tails in most of the singular value distributions studied here, it is unclear whether weight matrices of fully trained DNNs indeed belong to an ensemble of heavy tailed random matrices as suggested in Ref. [238].

3.3 Boundary between noise and information applied to filtering neural network weight matrices

Parts of this section closely follow the preprint Max Staats², Matthias Thamm², and Bernd Rosenow, *Boundary between noise and information applied to filtering neural network weight matrices*, arXiv preprint arXiv:2206.03927 [329].

The excellent agreement with RMT predictions suggests that the weights of deep neural networks are predominately random [328]. In addition to the random bulk, a few large singular values and vectors deviate from RMT, which we interpret as an indication for stored information. This also suggests that there is capacity left for the networks to learn random noise in the training data, which has also been reported in Ref. [39]. Inspired by these findings, we train neural networks in the presence of label noise, i.e., we shuffle a portion of the labels in the training data such that the network has to memorize these image label pairs to achieve full training accuracy. We then use RMT to locate how this noise is encoded in the weights, and find that there is a boundary between noise and information in the spectra which we also confirm by setting singular values to zero and monitoring the generalization accuracy. We then propose a filtering algorithm to reduce the influence of the noise in two ways:

- (i) We set singular values to zero starting from the smallest one up to the boundary between noise and information to remove the noise from the spectrum.
- (ii) Due to level repulsion with the noisy bulk, the singular values in the tail are larger than for an ideal low rank weight without noise. We revert this by shifting back large singular values.

We find that filtering can significantly improve the generalization performance for networks trained with label noise. For networks trained with severe overfitting, i.e. for much longer than required to obtain 100% training accuracy, the boundary between noise and information is blurred such that no improvements can be achieved anymore.

3.3.1 Boundary between noise and information in weight spectra

When training with label noise, we observe that the networks still achieve 100% training accuracy, even in the case of 100% label noise where all labels are drawn randomly. However, the generalization performance is significantly deteriorated – proportional to the amount of label noise (Tab. 3.1). To locate where information about the memorized noisy image-label pairs is stored in the weights, we again consider the singular value spectra obtained with Gaussian broadening and the averaged p -values of Kolmogorov-Smirnov tests comparing the singular vector entries to the Porter-Thomas distribution using the methods of Sec. 3.2. In Fig. 3.14, we show the results for the second hidden layer of MLP1024 networks trained with 0% noise (blue), 40% noise (green), and 100% (brown), and a random control (red) where we consider the initialized weights without training (for other layers and miniAlexNet, see Appendix B.1). In the presence of label noise, we observe slightly more intermediate and large singular values in the spectra of MLP1024 networks (lower panel in Fig. 3.14), while the small singular values still follow a Marcenko-Pastur like distribution even in the case of 100% label noise. In addition, the averaged p -values for vectors with singular values in the Marcenko-Pastur region are within

²equal author contribution

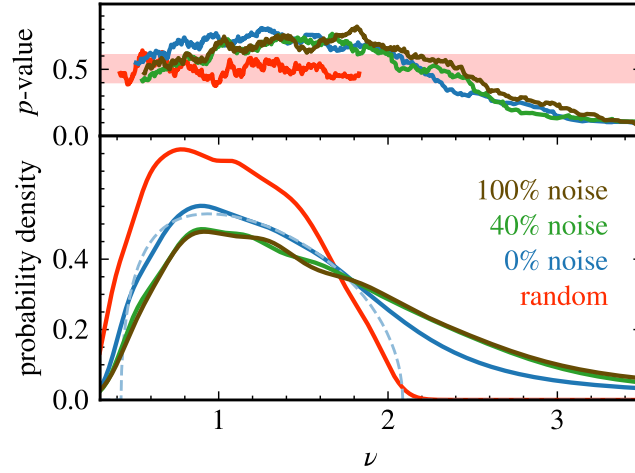


Figure 3.14: Analysis of singular values ν and vectors V of the first hidden layer weight matrix for the MLP1024 network trained with various amounts of label noise: 0% (blue), 40% (green), and 100% (brown). For reference, we show results for randomly initialized weights in red. The upper panel shows the randomness of singular vectors via the p -value of Kolmogorov-Smirnov tests against a Thomas-Porter distribution, averaged over neighboring singular values with a window size of 15; the light red stripe describes the 2σ region around the mean for random vectors. The lower panel depicts the corresponding singular value spectra obtained via Gaussian broadening with a window size of 15 (solid lines). The dashed line shows the fit of a Marchenko-Pastur distribution to the spectrum for 0% label noise. (Taken from Ref. [329].)

or even above the 2σ interval (light red) of the random control (red) regardless of the amount of label noise, and they begin to decline approximately at the same singular value, $\nu \approx 2$. For larger amounts of label noise, the p -values reach the lower bound of the 2σ region at slightly larger singular values; however, the size of the fluctuations in the p -values is comparable with the differences between the curves.

In comparison to the random control, where p -values fluctuate around 0.5, for trained networks the p -values in the random bulk are significantly increased, i.e., they often lie above the 2σ region of the random control. We argue that this is due to the presence of the few non-random singular vectors that store the information. These vectors force the random singular vectors to have a narrower distribution around the most likely part of the Porter-Thomas distribution (normal distribution with zero mean) due to the constraint of orthogonality with the deviating singular vectors with large singular values.

For example, using the same test statistics as described in Sec. 3.2.1, such that random normalized vectors from the Porter-Thomas distribution have on average a p -value of 0.5, the subset of vectors with zero mean have an average p -value of 0.74. We show in Fig. 3.15(a) that the mean values of singular vector entries for small singular values of the weight matrices (0% label noise blue, 40% green, 100% brown) are indeed smaller than the expected values (2σ range in light red stripe) for fully random matrices (red) while the means are much larger for large singular values which store the information.

The increase of p -values can also be shown for a simple model, adding a low-rank matrix δW to a fully random matrix W that would have singular vectors with p -values of 0.5 on average. For this we draw a random 1024×512 matrix W with Gaussian distributed i.i.d. entries with

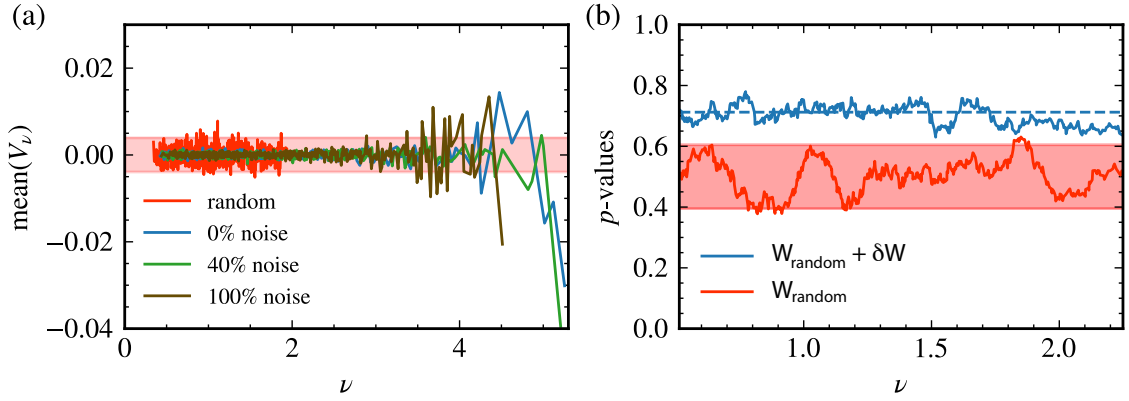


Figure 3.15: (a) Mean values of the singular vectors for the first hidden layer weights of the trained MLP1024 DNNs as a function of the corresponding singular values ν . The red line shows the means for singular vectors of a random Porter-Thomas matrix and the corresponding 2σ region of expected means for such matrices is shown by the transparent red stripe. We observe much smaller means for singular vectors of trained weights in the case of small singular values and significantly larger means for the vectors corresponding to large singular values. (b) Kolmogorov-Smirnov p -values for a random 1024×512 Porter-Thomas matrix W (red) with 2σ region for such matrices in light red, and for the sum $W + \delta W$ (blue), where δW is a matrix of rank ten with entries from a normal distribution with the same standard deviation and mean -0.01 similar to mean values observed for vectors corresponding to the largest singular values in (a). The p -values are averaged over neighboring singular values with a window size of 15. (Taken from Ref. [329].)

mean zero and variance $1/512$, and first show in Fig. 3.15(b) that the p -values fluctuate around 0.5 (red), with most values within the 2σ region (light red region). We then draw a second 1024×512 matrix with i.i.d. Gaussian distributed entries with mean -0.01 and variance $1/512$, compute the singular value decomposition, and reconstruct the matrix only keeping the largest 10 singular values, yielding a rank ten matrix δW . We then analyze the p -values of the singular vectors of $W + \delta W$. We find that the p -values are increased (blue line), with mean 0.71, which shows that in the presence of a few singular vectors with a distribution different from the random bulk, we expect the p -values in the bulk to be increased due to the enforced orthogonality to the singular vectors with a finite mean.

We demonstrated that the results are consistent with the idea that large singular values encode the rule while intermediate singular values store the memorized noise. We further test this hypothesis by setting singular values to zero, starting from the smallest one, and measuring the impact on training and test accuracy. For this, we compute the singular value decomposition $W = U\Sigma V^T$ of a weight W with singular values $\Sigma_{ij} = \delta_{ij}\nu_i$, set the smallest r values of the rank ordered ν_i in Σ_{ij} to zero (we assume $\nu_{i+1} \leq \nu_i$), i.e.

$$\tilde{\Sigma}_{ij} = \begin{cases} \delta_{ij}\nu_i & \text{for } i \leq r \\ 0 & \text{else} \end{cases}, \quad (3.58)$$

and then put the reconstructed weights $\tilde{W} = U\tilde{\Sigma}V^T$ back into the network. In Fig. 3.16, we show the training accuracy as a function of the amount of removed singular values for various layers of MLP1024 (panel a) and miniAlexNet networks (panel b) trained without label noise (blue), with

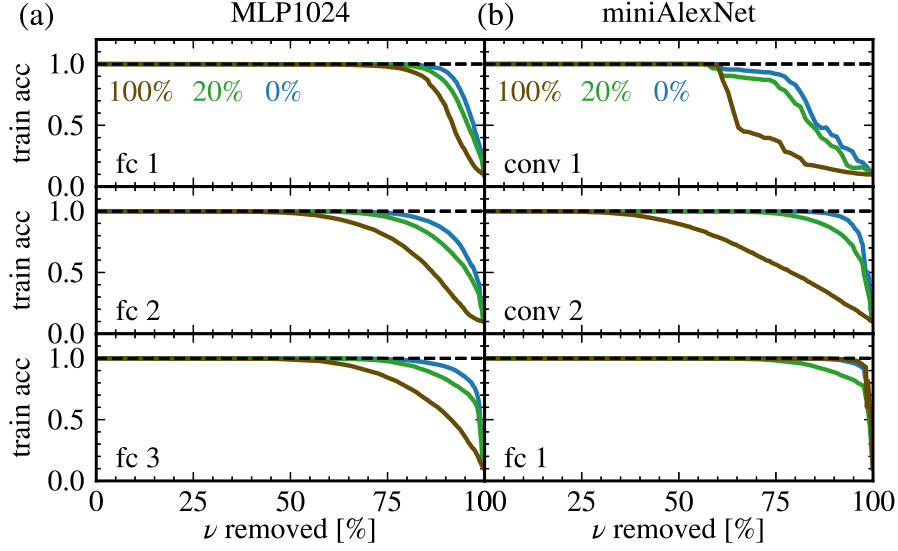


Figure 3.16: Boundary between information and noise demonstrated by setting a given percentage of the singular values to zero. (a) Training accuracy for the MLP1024 network trained with various amounts of label noise (0% blue, 20% green, and 100% brown). (b) Training accuracy for removing singular values from the convolutional network miniAlexNet trained with 0% label noise (blue), 20% (green), and 100% label noise (brown). In all cases, relevant information is stored in the largest singular values and corresponding vectors only. In presence of label noise larger parts of the spectrum are needed to store the noise.

20% label noise (green), and with 100% label noise (brown). We indeed find in all cases that the majority of singular values can be set to zero without any impact on the training accuracy for networks trained without label noise, which demonstrates that these singular values and vectors contain only noise. In addition, we find that more singular values are relevant when training with larger amounts of label noise. This is particularly pronounced in the second convolutional layer of miniAlexNet where the training accuracy drops significantly when removing more than 30% of singular values, compared to a threshold of 90% for clean training data. Here, the convolutional layers seem to encode almost all the memorized training data in the case of 100% label noise, as the first dense layer shows barely any difference to the pristine case when setting singular values to zero.

When considering the test accuracy in Fig. 3.17, it becomes apparent that generalization is due solely to the largest singular values and corresponding vectors even in the case of the large pre-trained networks alexnet (panel d, orange) and vgg19 (panel d, blue). These results suggest that the largest singular values and corresponding vectors encode the underlying rule necessary for generalization, intermediate singular values store memorized image-label pairs when training with label noise, and the bulk of small singular values and corresponding vectors is random noise neither relevant for training nor test accuracy.

We next consider how the generalization performance of networks trained with label noise depends on the removal of singular values. In Fig. 3.18, we consider (a) the hidden layer in MLP1024 networks and (b) the second convolutional layer in miniAlexNet trained with 20% label noise (upper panel) and 40% label noise (lower panel). For the regular training schedule we find that (i) removing small singular values has again no impact on the test accuracy, (ii)

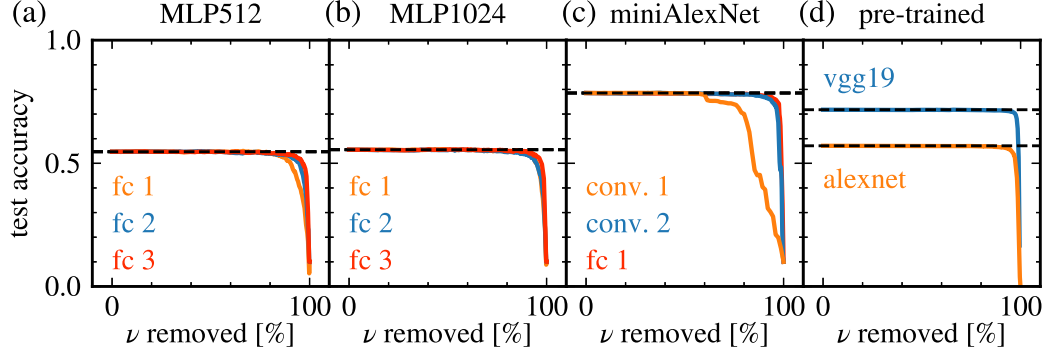


Figure 3.17: Test accuracies for setting a given percentage of the singular values in networks trained without label noise to zero for (a) the hidden layers of an MLP512 network, (b) the hidden layers of an MLP1024 network, (c) the convolutional layers and the first dense layer in miniAlexNet, and (d) the first dense layers in the large pre-trained networks vgg19 (blue) and alexnet (orange). We find that most of the singular values and corresponding vectors do not contribute to the generalization performance of the networks.

setting intermediate singular values to zero even improves the generalization accuracy, and (iii) when removing the largest singular values, the test accuracy sharply drops. Interestingly, the final drop occurs approximately at the same position as for the corresponding test accuracy in the case of 0% noise (Fig. 3.17, blue lines in panel b and c) and the increase is observed when removing singular values in the intermediate region that is relevant for the training accuracy when training with label noise (Fig. 3.16, green). We argue that this indicates the presence of a boundary between information about the rule, stored in the largest singular values and corresponding vectors, and the noise learned to memorize image-label pairs of noisy training data, encodes in the intermediate singular values.

In addition, we train the MLP1024 networks with a learning schedule that promotes overfitting, i.e., we train with a slower learning rate decay for much longer than required to achieve 100% training accuracy (more details in Sec. 3.1.2). This causes an earlier drop of the generalization accuracy (Fig. 3.18, red) without prior improvements. We attribute this behavior to a blurring of the boundary between noise and information, i.e. when training for too long, intermediate singular values that memorize the noise mix with the region of large values that encode the rule.

3.3.2 Filtering of neural network weight matrices

We argued that when training networks with label noise, more intermediate singular values are required for achieving full training accuracy, and that there is a separation between these singular values and the largest ones that encode the underlying rule in case a suitable training schedule is chosen. Based on these findings, we propose a filtering algorithm to reduce the influence of the noise by

- i) setting singular values corresponding to the noise to zero.
- ii) reverting the shift of large singular values due to level repulsion with the random bulk.

We already discussed in the previous section that removing the intermediate singular values can improve the generalization accuracy. In addition, the presence of a random bulk of smaller

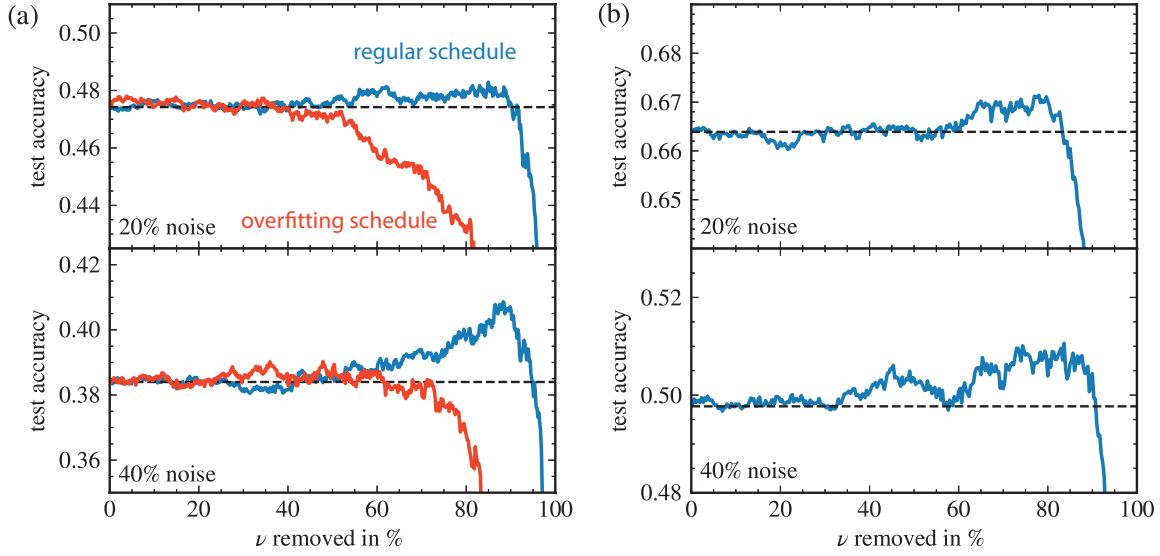


Figure 3.18: Dependence of the test accuracy on the removal of singular values from (a) the second hidden layer weights of MLP1024 networks and (b) the second convolutional layer in miniAlexNet, trained in the presence of label noise (upper panel: 20%, lower panel: 40%). Setting singular values to zero (blue), we observe significant generalization performance improvements. For training with overfitting (red in a), no improvement is observed, indicating that information and noise are mixed in the spectrum. (Adapted from Ref. [329].)

singular values leads to an upward shift of large singular values, which has been discussed for linear networks in Ref. [229] and which is also empirically known from econophysics [30, 31, 468, 469]. We here consider a model for which the weight $W \in \mathcal{M}_{n_i, n_{i+1}}$ consists of a low rank term W_0 that encodes the relevant information about the rule and a full rank random matrix W_{noise} that describes the bulk:

$$W = W_0 + W_{\text{random}} . \quad (3.59)$$

The shift of singular values was explicitly computed [229, 470] in the limit where the matrix dimensions tend to infinity $n_l, n_{l-1} \rightarrow \infty$ while the aspect ratio $A = n_l/n_{l-1} \leq 1$ is fixed. For i.i.d. random entries of W_{noise} with standard deviation σ , the unperturbed singular values ν_0 of W_0 can be obtained from the singular values ν of the full weight $W = W_0 + W_{\text{noise}}$ by

$$\nu_0/\sigma = \frac{1}{\sqrt{2}} \left[(\nu/\sigma)^2 - A - 1 + \sqrt{\left((\nu/\sigma)^2 - A - 1 \right)^2 - 4q} \right]^{1/2} . \quad (3.60)$$

Here, we obtain σ from fitting a Marcenko-Pastur distribution $P_{\text{MP}}(\nu)$, Eq. (3.46), to the singular values ν_i of W using the following algorithm:

- i) As the spectrum contains the large singular values, the Marcenko-Pastur distributed part of the spectrum is not normalized, and the end of the Marcenko-Pastur region ν_+ is not known a priori. We therefore first broaden the spectrum using Gaussian broadening [182, 446], Eq. (3.38), with a window size $a = 15$, and then fit an adjusted two-parameter Marcenko-Pastur distribution where we fix $\nu_- = \min(\nu_i)$ and use ν_+ and the height as independent fit parameters. This yields an estimate for the bounds of the Marcenko-Pastur region ν_- and ν_+ .

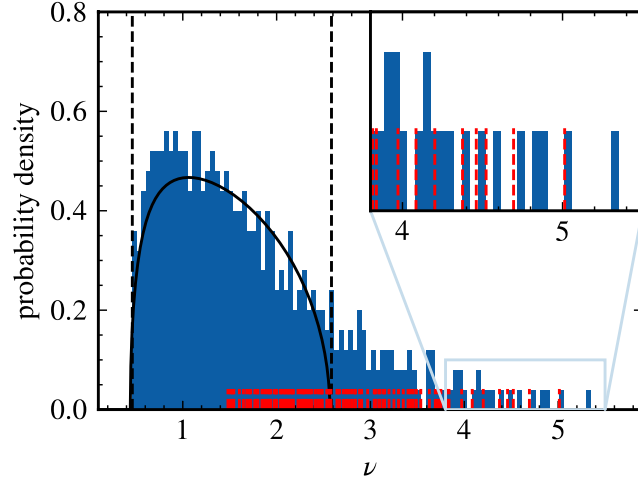


Figure 3.19: Shifting of singular values: histogram of singular values for the first hidden layer weight matrix of the MLP1024 network (blue) trained with 40% label noise and Marcenko-Pastur fit (solid, black) with boundaries of the MP region (dashed black lines). The dashed red lines show the locations of shifted singular values according to Eq. (3.60), and the inset zooms into the tail region. (Taken from Ref. [329].)

- ii) We then fit a proper Marcenko-Pastur fit, Eq. (3.46), with single fit parameter σ to the normalized histogram of the ν_i between ν_- and ν_+ .

The shifting formula can be applied for singular values above the Marcenko-Pastur region $\nu_i > \nu_+$, and we keep the other values unchanged. An example for this shifting in the second hidden layer of an MLP1024 network trained with 40% label noise is shown in Fig. 3.19. Here, large singular values are shifted by a comparably small amount while smaller singular values are strongly shifted such that many values close to ν_+ are pushed into the Marcenko-Pastur region (marked by the dashed, vertical lines).

In Fig. 3.20, we apply filtering to MLP1024 and miniAlexNet networks trained with label noise (upper panels: 20%, lower panels: 40%). We find that shifting singular values in addition to removing (green) gives a further significant improvement compared to only removing singular values (blue). Here, we first set singular values from the Marcenko-Pastur region to zero, before removing shifted values even for values that are shifted into the Marcenko-Pastur region.

To systematically study the improvements due to noise filtering, we train many MLP1024 networks on the CIFAR-10 dataset for various amounts of label noise each with different seeds of the random number generator for initialization. We consider networks trained with 9 different amounts of label noise from 0% to 80%, and ten different seeds for each amount of noise. For noise filtering of a full network, we use the following algorithm:

- i) Start with the last hidden layer $l = L - 1$.
- ii) Compute the singular value decomposition $W^{(l)} = U^{(l)} \text{diag}(\nu_i^{(l)}) [V^{(l)}]^T$.
- iii) Rank-order the singular values and corresponding vectors such that $\nu_i^{(l)} \leq \nu_{i+1}^{(l)}$.
- iv) Perform a Marcenko-Pastur fit to obtain the Marcenko-Pastur region $\nu_{\pm}^{(l)}$ and the width $\sigma^{(l)}$.

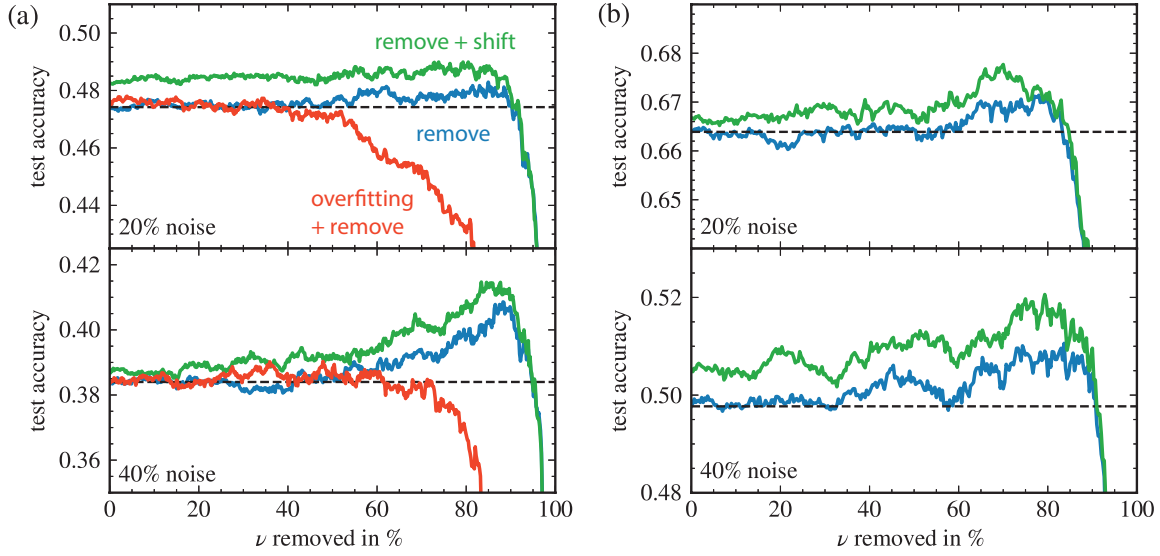


Figure 3.20: Dependence of the test accuracy on the removal and shifting of singular values from (a) the second hidden layer weights of MLP1024 networks and (b) the second convolutional layer in miniAlexNet, trained in the presence of label noise (upper panel: 20%, lower panel: 40%). For setting singular values to zero (blue) and when additionally shifting them according to Eq. (3.60) (green), we observe a significant improvement in performance. For training with overfitting (red in a) no improvement is observed, indicating that information and noise are mixed in the spectrum. (Adapted from Ref. [329].)

- v) Shift the singular values above the Marcenko-Pastur region, $\nu_i^{(l)} > \nu_+^{(l)}$, according to Eq. (3.60) but keeping the rank in the spectrum

$$\tilde{\nu}_i^{(l)} = \begin{cases} \nu_i^{(l)} & \text{for } \nu_i^{(l)} \leq \nu_+^{(l)} \\ \nu_0(\nu_i^{(l)}) & \text{for } \nu_i^{(l)} > \nu_+^{(l)} \end{cases} . \quad (3.61)$$

- vi) Set singular values $\tilde{\nu}_i^{(l)}$ to zero starting from index $i = 1$, reconstruct the weights $\tilde{W}^{(l)} = U^{(l)} \text{diag}(\tilde{\nu}_i^{(l)}) V^{(l)T}$, and measure the accuracy on a validation set. In case there is no improvement from shifting, we take the unshifted values for this layer.
- vii) Determine the amount of removed singular values for which the best validation accuracy is obtained, yielding the optimal singular values $\bar{\nu}_i$, and replace the original weight of the network by $\bar{W}^{(l)} = U^{(l)} \text{diag}(\bar{\nu}_i) [V^{(l)}]^T$.
- viii) Proceed with the previous layer $l \rightarrow l - 1$ from step ii) until one arrives at the input layer.

We apply the algorithm to all 90 networks, using a validation set of 2000 images from CIFAR-10, and compute the accuracy improvement on the test set of the remaining 8000 CIFAR-10 images. The average improvements over the different initial realizations for each amount of the noise together with the error of the mean are depicted in Fig. 3.21. For the regular learning schedule (blue and green), we find significant improvements of up to 6% already by only removing

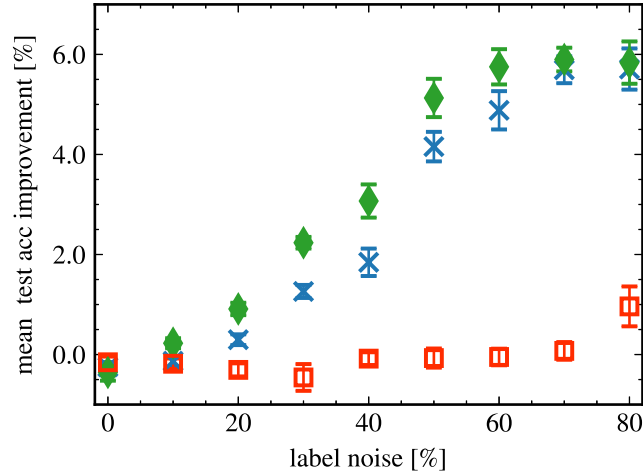


Figure 3.21: Average improvement of the test accuracy when removing singular values (blue, red) from all layers and when additionally shifting singular values (green) of the first two layers in MLP1024 networks, with results for both the standard learning rate schedule (blue crosses, green diamonds) and an overfitting schedule (red squares). We observe that the average improvements increase with increasing amount of label noise, with an enhanced improvement for additionally shifting singular values. There are no improvements for networks trained with overfitting. (Taken from Ref. [329].)

singular values (blue), i.e. by skipping step iv) in the filtering algorithm. When additionally shifting singular values in the first and second hidden layer (green), i.e. performing all steps of the algorithm, we find additional improvements such that already for 20% label noise, an average improvement of the generalization accuracy of about 1% is achieved.

We additionally train another 90 MLP1024 networks using the learning rate schedule that promotes severe overfitting (red squares). In this case, no improvements are found which we again attribute to mixing between the parts of the spectrum that memorizes noisy data and the region that encodes the underlying rule.

In principle the way how label noise is implemented can affect the generalization accuracy in two ways: (i) there is overall less valuable information about the rule in the dataset when replacing a portion of the labels by random labels, and (ii) the network may get distracted by the wrong labels. To test how strong the former effect is, we consider a smaller control training set, for each training set with label noise, by removing all noisy image-label pairs. The performance on the control dataset defines an upper bound to the test accuracy that the network can learn when trained with the corresponding noisy set. In Fig. 3.22, we show average generalization accuracies of MLP1024 networks trained on the control set (brown circles), reference accuracies for training on datasets with label noise without noise filtering (black squares), and for using our noise filtering algorithm (green diamonds). It becomes apparent that label noise severely reduces the generalization performance, while just training with fewer data has relatively little influence on the test accuracy. Even though we observe significant improvements by using noisy filtering, they cannot fully make up for the reduction of the generalization performance due to label noise.

In the proposed algorithm, we do not optimize the amount of singular values that are shifted, but instead we determine ν_+ and σ from Marcenko-Pastur fits and shift the values larger than ν_+ .

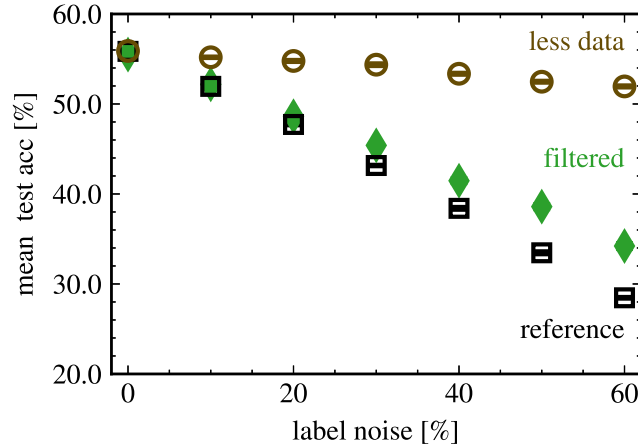


Figure 3.22: Average test accuracy as a function of the label noise for MLP1024 networks trained with the typical learning schedule. For each noise, we trained ten networks initialized using different seeds for the pseudo random number generator. We show mean accuracies and error of the mean without filtering (black squares), and for removing singular values from all layers and shifting singular values of the first two layers (green diamonds). For comparison, we also show accuracies for networks where we removed a percentage of images and labels from the training data set instead of assigning random labels (brown circles). For a fair comparison, we removed exactly those image-label pairs for which random labels are drawn in the corresponding network trained with label noise. It becomes apparent that label noise significantly reduced the training accuracy by perturbing learning of the underlying rule, and while weight filtering can significantly improve the generalization performance, it cannot make up for the influence of the noise.

To estimate to what extent these values optimize the accuracy improvements, we measure the generalization performance for a MLP1024 network trained with 40% label noise as a function of both σ and the portion of removed singular values. For the data presented in Fig. 3.23, we split the CIFAR-10 test dataset into 5 batches of 2000 images to simulate different validation data sets and average the accuracy improvements over all 5 batches. We find that optimal improvements on all batches are found for filtering parameters that lie in the same region but do not agree perfectly (black circles). For the first two hidden layers, where our filtering algorithm decides to shift singular values, the parameters from the Marcenko-Pastur fits (intersection point of the dashed lines) are within the region of the best improvements. For the third hidden layer, the fit results do not yield an improvement such that our algorithm skips shifting singular values in this layer. However, allowing for different σ reveals an optimum close to the boundary value

$$\sigma_{\max} = \begin{cases} \nu_{\max}(1 - \sqrt{A}) & \text{for } A < 1 \\ \nu_{\max}/2 & \text{for } A = 1 \end{cases}, \quad (3.62)$$

for which only the largest singular value ν_{\max} is shifted.

3.4 Conclusions

We studied weights of trained neural networks using RMT as a zero information hypothesis and demonstrated that large parts of the weights remain random during training. In particular,

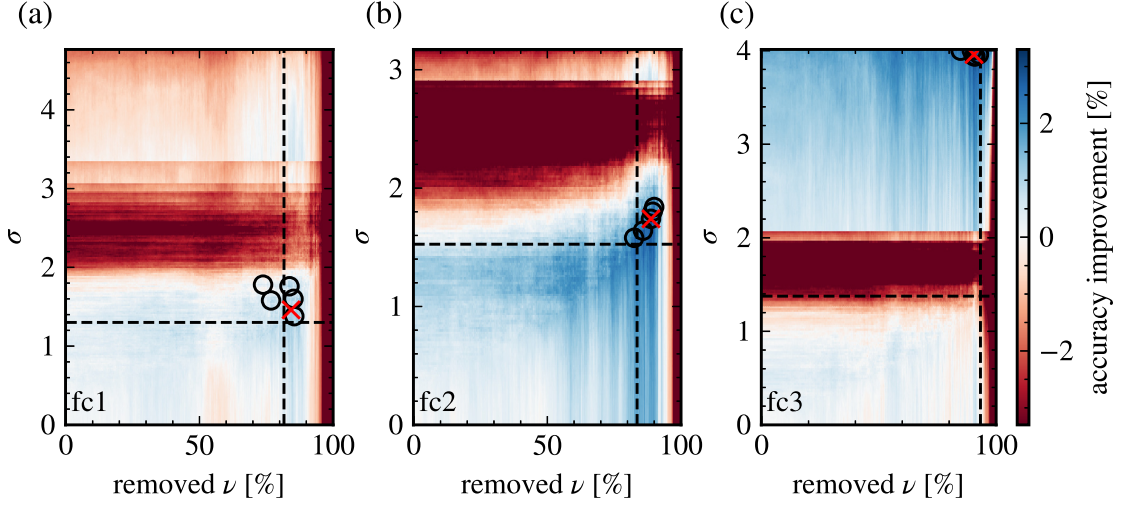


Figure 3.23: Accuracy improvements as a function of both the ratio of removed singular values and the value of σ used for shifting singular values for (a) the first hidden layer, (b) the second hidden layer, and (c) the third hidden layer of MLP1024 networks trained with 40% label noise. We split the 10 000 CIFAR-10 test images into 5 batches of 2000 images each and averaged accuracies obtained on these batches. The black circles show the location of the maximal improvement for each of these batches, and the red crosses mark the maximum of the averaged accuracy. The dashed, horizontal lines show the value of σ obtained from a Marcenko-Pastur fit to the spectrum, and the dashed, vertical lines show the percentage of singular values within the Marcenko-Pastur region. In the first two hidden layers, the optimal improvement is obtained for parameters close to the ones obtained from the Marcenko-Pastur fit. In the third hidden layer, a Marcenko-Pastur fit does not find a σ that yields an improvement, and the optimum lies close to σ_{\max} where only a few of the largest singular values are shifted and most of the singular values are set to zero.

the agreement of the level spacing distribution of singular values with the Wigner surmise and logarithmic growth of the level number variance reveals that the singular value spectra of trained weights are predominately random. In addition, a comparison of singular vectors with the Porter-Thomas distribution shows that singular vectors are random as well, with exceptions only for the vectors corresponding to large singular values. This is also supported by directly comparing Marcenko-Pastur distributions to the spectra, indicating that relevant information is located in the tail of the spectra and the corresponding singular vectors. Furthermore, comparing to RMT predictions is also fruitful for distinguishing between the learning regimes.

By training networks with label noise, systematically setting singular values to zero, and measuring training, and generalization performance, we find that (i) the bulk of small singular values is random and carries no information, (ii) intermediate singular values can be important for the training accuracy by memorizing noise, and (iii) large singular values carry the relevant information about the underlying rule that allows the networks to generalize. If networks are trained in a suitable way, we observe a separation between regions (ii) and (iii), and when training for too long after reaching 100% training accuracy, severe overfitting sets in and the boundary gets blurred. In the former case, we proposed an algorithm for filtering weight matrices to mitigate the influence of the noise by (a) setting singular values to zero to remove the parts of the

spectrum that do not carry relevant information and (b) by reverting the level repulsion between the random bulk of small singular values and the large relevant values. We find significant improvements when applying the filtering algorithm to networks training with label noise.

So far, for reverting the level repulsion, we focused on a simple model for the weights trained with label noise, by considering a random bulk W_{random} added to a low rank weight W_0 , encoding the rule. In the future it would be interesting to study the level repulsion more rigorously to aim for a better shifting method to potentially bring the generalization performance of the networks trained with label noise even closer to the values for networks trained with pristine data. Furthermore, it may be interesting to study filtering techniques that also operate on the directions of singular vectors.

Setting small singular values to zero and reducing the size of large singular values is in fact not too dissimilar from what L_2 regularization tries to achieve during training. It might therefore also be interesting to consider implementing some filtering methods already during training. For instance, one could think about a training algorithm that early in training – after the relevant directions are sorted out, and the largest singular values separated from the bulk – systematically shifts small singular values to zero such that training can focus on the relevant directions without the problem of level repulsion.

4. Entanglement entropy in one dimensional quantum chains of interacting fermions

Parts of this chapter closely follow the publication: Matthias Thamm, Harini Radhakrishnan, Hatem Barghathi, Bernd Rosenow, and Adrian Del Maestro, *One-particle entanglement for one dimensional spinless fermions after an interaction quantum quench*, Physical Review B 106, 165116 (2022) [327].

In this chapter, we consider one dimensional systems of spinless, interacting fermions. In particular, we assume that the system is in some pure state $|\Psi\rangle$ such that the density matrix is given by $\rho = |\Psi\rangle\langle\Psi|$. We then split the system into two subregions A and B and study entanglement between the subspaces by tracing out region B , yielding the reduced density matrix $\rho_A = \text{Tr}_B \rho$ that can be used to obtain the von Neumann entanglement entropy

$$S_1 = -\text{Tr} [\rho_A \ln \rho_A] . \quad (4.1)$$

An alternative measure for the entanglement entropy is the Rényi entropy of power $\alpha \in \mathbb{N}$

$$S_\alpha = \frac{1}{1-\alpha} \ln \text{Tr} [\rho_A^\alpha] , \quad (4.2)$$

which for $\alpha \rightarrow 1$ reduces to the von Neumann entropy, and for $\alpha \geq 2$ does not require computation of the logarithm of a matrix. As discussed in the introduction, there are different ways to divide the one-dimensional system with N fermions into two partitions:

- (i) a spatial bipartition, where A is a connected region of length ℓ and the remaining system has length $L - \ell$.
- (ii) a particle bipartition by fixing n particles and tracing out the position of the remaining $N - n$ particles.

To study the role of entanglement between the bipartitions, we consider free fermions for times $t < 0$ and at $t = 0$ we suddenly switch on interactions – a so-called interaction quantum quench. Initially after the quench, the spatial entanglement entropy increases linearly until the system equilibrates, leading to a saturation of the entanglement entropy. Here, spatial entanglement entropy takes the role of a thermodynamic entropy after thermalization [471, 472] and the initial increase describes the spread of information through the system after the quench.

For most parts of this chapter, we consider particle entanglement entropy which initially grows much faster after a quench [308]. It has been shown [308] that in the thermodynamic limit, spatial and particle entanglement entropy are equivalent asymptotically after the quench. In addition, particle entanglement is encoded in the spectrum of the n -body reduced density matrix which gives access to all n -body observables. Here, particle entanglement provides information about quantum correlations between indistinguishable particles. In the following, we mainly focus on the one-body reduced density matrix and one-particle entanglement entropy in a one-dimensional lattice of interacting, spinless fermions with nearest neighbor hopping and short range interactions – the J - V model. In the case of small interactions – the Luttinger liquid phase where low energy excitations are described by density fluctuations on top of an average density background – we obtain the one-body density from an analytic bosonization calculation.

However, to account for the short range nature of the interactions in the J - V model, we need to introduce an interaction cutoff ε that cannot be fixed from the Luttinger liquid calculation alone. We therefore perform large scale numerical computations using exact diagonalization (ED) for large systems up to $N = 19$ fermions on $L = 38$ sites. We consider even larger systems up to $L = 102$ sites at half filling using density matrix renormalization group calculations (DMRG), which is an approximate method that we control and stabilize using ED results and analytical considerations about the ground state. Being able to consider large systems allows reliable finite size scaling to the thermodynamic limit and to determine the interaction cutoff of the Luttinger liquid model from the numerical results.

In summary, we (i) provide a definite picture of the one-body reduced density matrix and entanglement in a one-dimensional integrable model both in equilibrium and after a quantum quench, and (ii) make a fruitful connection between state-of-the-art numerical techniques and an analytical bosonization calculation, which allows the field theory approach to make accurate predictions even beyond universal scaling relations [327].

This chapter is structured as follows: In Sec. 4.2, we introduce the J - V model, which describes a one-dimensional lattice of interacting fermions, its phase diagram, and an instructive example calculation for two fermions on four sites – the smallest non-trivial system that illustrates how entanglement entropy is computed in practice. In Sec. 4.3, we perform the Luttinger liquid bosonization calculation in equilibrium and after a quantum quench, followed by a detailed description of the methods and tricks used for numerical computation with ED (Sec. 4.4) and DMRG (Sec. 4.5). Finally, in Sec. 4.6, we compare numerical results with analytical Luttinger liquid results to unambiguously determine the interaction cutoff of the bosonization calculation.

The results of this chapter have been obtained during my research stay with the Del Maestro group at the University of Tennessee in Knoxville (UTK). The bosonization calculation was performed in close collaboration with Harini Radhakrishnan (UTK), and large parts of the ED code have already been written by Adrian Del Maestro and Hatem Barghathi (UTK) for a previous project [473]. We added improvements to the ED code for this project to allow for larger system sizes and wrote new code for the DMRG calculation [474] based on `ITensor.jl` [475].

4.1 Schmidt decomposition

In this section, we describe a technique to efficiently compute the entanglement entropy based on a singular value decomposition for a given bipartition of the system. For a general partitioning into A and B , we can write the state $|\Psi\rangle$ as

$$|\Psi\rangle = \sum_{ab} C_{ab} |\theta_a\rangle_A \otimes |\chi_b\rangle_B , \quad (4.3)$$

where $|\theta_a\rangle_A$ is a basis in A and the $|\chi_b\rangle_B$ form a basis in partition B . We perform a singular value decomposition of the coefficient matrix C such that we can write

$$C_{ab} = \sum_k U_{ak} \lambda_k V_{kb}^\dagger \quad (4.4)$$

$$|\Psi\rangle = \sum_{abk} U_{ak} \lambda_k V_{kb}^\dagger |\theta_a\rangle_A \otimes |\chi_b\rangle_B . \quad (4.5)$$

By defining new basis states

$$|\Theta_k\rangle_A = \sum_a U_{ak} |\theta_a\rangle_A \quad (4.6)$$

$$|X_k\rangle_B = \sum_b V_{kb}^\dagger |\chi_b\rangle_B, \quad (4.7)$$

we can write the state in the Schmidt decomposed representation using a single sum

$$|\Psi\rangle = \sum_k \lambda_k |\Theta_k\rangle_A \otimes |X_k\rangle_B. \quad (4.8)$$

Due to the properties of the singular value decomposition, U and V are unitary matrices such that $|\Theta_k\rangle_A$ and $|X_k\rangle_B$ both form a basis in their respective subspace and can therefore be used to compute the trace for the reduced density matrix

$$\begin{aligned} \rho_A = \text{Tr}_B [|\Psi\rangle\langle\Psi|] &= \sum_{kk'k''} \lambda_k \lambda_{k''}^* \langle X_k | X_{k'} \rangle_B |\Theta_{k'}\rangle_A \langle \Theta_{k''}|_A \langle X_{k''} | X_k \rangle_B \\ &= \sum_k |\lambda_k|^2 |\Theta_k\rangle_A \langle \Theta_k|_A \end{aligned} \quad (4.9)$$

$$\rho_B = \text{Tr}_A [|\Psi\rangle\langle\Psi|] = \sum_k |\lambda_k|^2 |X_k\rangle_B \langle X_k|_B. \quad (4.10)$$

Once, the Schmidt decomposition has been performed, the von Neumann and Rényi entropies can be directly computed as

$$S_1 = - \sum_k |\lambda_k|^2 \ln (|\lambda_k|^2) \quad (4.11)$$

$$S_\alpha = \frac{1}{1-\alpha} \ln \left(\sum_k |\lambda_k|^{2\alpha} \right). \quad (4.12)$$

It is therefore sufficient to construct the matrix C and compute its singular values λ_k to obtain the entropies without explicitly performing the partial trace.

4.2 The J-V model

We study a system of N spinless fermions on a one-dimensional lattice with L sites at half-filling $L = 2N$ described by the J - V Hamiltonian

$$H = -J \sum_{i=1}^L (c_{i+1}^\dagger c_i + c_i^\dagger c_{i+1}) + V \sum_{i=1}^L n_i n_{i+1}. \quad (4.13)$$

Here J is the hopping amplitude, V is the nearest neighbor interaction, c_i^\dagger creates a fermion at site i , and $n_i = c_i^\dagger c_i$ is the occupation number operator for site i . In the case of even number of particles N we use anti-periodic boundary conditions and for odd N we use periodic boundary conditions, which ensures that the ground state is always non-degenerate. We can express states using an occupation number basis

$$|n_1, n_2, \dots, n_L\rangle = \prod_{i=1}^L c_i^\dagger |0\rangle, \quad (4.14)$$

where $|0\rangle$ is the empty lattice, and we use the convention that creation operators c_i^\dagger are rank ordered according to the site index i in the product above, such that operators with smaller i act before those with larger site index.

In the following, we first consider the phase diagram of the J - V model, and then discuss the simple example for $N = 2$ fermions on $L = 4$ lattice sites, where we apply the important concepts that are generalized to large systems for the numerical computations later.

4.2.1 Phases from mapping onto a spinful XXZ-chain

To understand the phase diagram of the J - V model, it is advantageous to start from a spin-1/2, one-dimensional XXZ model [476, Eq. (6.2)] on a lattice of L sites

$$H_{\text{XXZ}} = \sum_i J_{xy} \left(S_{i+1}^x S_i^x + S_{i+1}^y S_i^y \right) + J_z \sum_i S_{i+1}^z S_i^z , \quad (4.15)$$

where the spin operators $S^k = \sigma_k/2$ can be represented by Pauli matrices and have the commutation relations [476]

$$[S^k, S^{k'}] = i\varepsilon_{kk'k''} S^{k''} . \quad (4.16)$$

This model can be solved using a Bethe ansatz [323, 324, 476, 477], and it is well known that the system has phase transitions at $J_z/J_{xy} = \pm 1$ [476]. For $J_z/J_{xy} < -1$ the ground state is ferromagnetic with aligned spins, and for $J_z/J_{xy} > 1$ the system is an anti-ferromagnet which prefers alternating spins. We consider $S_j^\pm = S_j^x \pm iS_j^y$ and the Wigner transformation [476]

$$S_j^+ \longrightarrow (-1)^j c_j^\dagger e^{i\pi \sum_{k=1}^{j-1} n_k} \quad (4.17)$$

$$S_j^- \longrightarrow (-1)^j e^{-i\pi \sum_{k=1}^{j-1} n_k} c_j \quad (4.18)$$

$$S_j^z \longrightarrow n_j - \frac{1}{2} , \quad (4.19)$$

which ensures the correct commutation relations [476]. Using that

$$S_{i+1}^x S_i^x + S_{i+1}^y S_i^y = \frac{1}{2} \left(S_{i+1}^+ S_i^- + S_{i+1}^- S_i^+ \right) \quad (4.20)$$

$$S_{j+1}^+ S_j^- \longrightarrow -c_{j+1}^\dagger e^{i\pi n_j} c_j = -c_{j+1}^\dagger c_j \quad (4.21)$$

$$S_{j+1}^- S_j^+ \longrightarrow -c_{j+1} e^{-i\pi n_j} c_j^\dagger = c_{j+1} c_j^\dagger = -c_j^\dagger c_{j+1} , \quad (4.22)$$

the transformed Hamiltonian becomes

$$H_{\text{XXZ}} \longrightarrow -\frac{J_{xy}}{2} \sum_i \left(c_{j+1}^\dagger c_j + c_j^\dagger c_{j+1} \right) + J_z \sum_i n_{j+1} n_j + J_z (L/4 - N) , \quad (4.23)$$

where a fixed particle number N in the transformed model corresponds to fixed magnetization of the XXZ model. By using the replacement [476]

$$\frac{J_{xy}}{2} \longrightarrow J \quad (4.24)$$

$$J_z \longrightarrow V \quad (4.25)$$

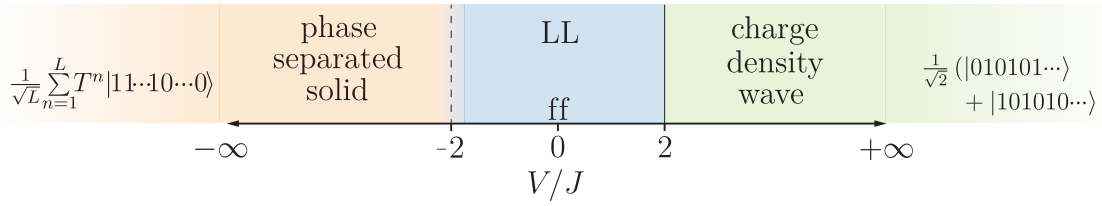


Figure 4.1: Phases of the J - V model. For small interactions, $|V/J| \leq 2$, the system is in the *Luttinger liquid* phase (LL), where $V/J = 0$ corresponds to free fermions (ff). At $V/J = 2$, there is a first order phase transition into the *charge density wave* phase where the ground state favors separation between the fermions. At $V/J = -2$ there is a second order phase transition to the *phase separated solid* phase, where the ground state prefers clustering the fermions together.

one can therefore map the XXZ model onto the J - V model

$$H_{\text{XXZ}} \longrightarrow H_{JV} + V(L/4 - N) . \quad (4.26)$$

Here, the constraint of half-filling of the lattice, $L = 2N$, maps onto a magnetization of zero in the XXZ model. For $V/J > 0$, in the absence of a magnetic field in the XXZ model (as it is the case in Eq. 4.15), the average of the spin operator S_z vanishes in the ground state, i.e. $\langle S^z \rangle = 0$, and the ground state has zero magnetization [145]. The situation is more complicated for negative interactions $V/J < 0$ and especially the limit of infinite negative interactions where the unrestricted XXZ model has a ferromagnetic ground state with maximal magnetization. However, it has recently been shown in Ref. [325], that the phase boundaries are not altered by restriction to zero magnetization (half-filling) and the ground state in the clustered solid phase favors clustering of spins of the same direction with an equal number of \uparrow and \downarrow spins.

We therefore conclude that the J - V model has phase transitions at

$$\frac{J_z}{J_{xy}} = \pm 1 \longrightarrow \frac{V}{J} = \pm 2 , \quad (4.27)$$

such that the J - V system can be in one of three phases (Fig. 4.1):

- (i) For $V/J < -2$ the system is a phase separated solid, where the strong attractive interactions favor clustering of fermions such that the ground state for $V/J \rightarrow -\infty$ becomes

$$|\Psi_{V/J \rightarrow -\infty}\rangle = \frac{1}{\sqrt{L}} \sum_{n=1}^L T^n |11 \dots 1100 \dots 0\rangle , \quad (4.28)$$

where $|11 \dots 1100 \dots 0\rangle$ is the state for which the first N sites are occupied by a fermion and the remaining N sites are empty. Here, T is the translation operator that shifts each fermion one site to the right, *e.g.* $T|011001\rangle = |101100\rangle$.

For the limit $V/J \rightarrow -\infty$ at half filling $L = 2N$, the first row of the one-body reduced density matrix

$$\rho_1^{ij} = \frac{1}{N} \langle \Psi_{V/J \rightarrow -\infty} | c_i^\dagger c_j | \Psi_{V/J \rightarrow -\infty} \rangle \quad (4.29)$$

is given by

$$\rho_1^{1,k} = \frac{1}{NL} \begin{cases} N & \text{for } k = 1 \\ 2 & \text{for } k = N + 1 \\ 0 & \text{else} \end{cases} . \quad (4.30)$$

Due to the translational symmetry (see also Sec. 4.4.1), the one-body reduced density matrix can be diagonalized by a Fourier transform which yields the spectrum

$$\lambda_n = \sum_{k=1}^L \rho_1^{1,k} \cos[q_n(k-1)] \quad (4.31)$$

$$= \frac{1}{2N} \left[1 + \frac{2}{N} \cos(q_n N) \right] \quad (4.32)$$

$$= \frac{1}{2N} + \mathcal{O}(N^{-2}) . \quad (4.33)$$

From this, one finds the entropy in the thermodynamic limit as

$$S_1(V/J \rightarrow -\infty) - \ln(N) = - \sum_{k=1}^L \lambda_k \ln(\lambda_k) - \ln(N) = \ln(2) \quad (4.34)$$

$$S_\alpha(V/J \rightarrow -\infty) - \ln(N) = \frac{1}{1-\alpha} \ln \left(\sum_{k=1}^L \lambda_k^\alpha \right) - \ln(N) = \ln(2) . \quad (4.35)$$

- (ii) For the other strongly interacting case $V/J > 2$, the system is in the charge density wave phase where strong repulsion results in a ground state with maximal separations between the fermions. For a lattice at half filling, the ground state in the limit $V/J \rightarrow \infty$ becomes

$$|\Psi_{V/J \rightarrow +\infty}\rangle = \frac{1}{\sqrt{2}} (|10101 \dots\rangle + |01010 \dots\rangle) . \quad (4.36)$$

In this limit, the one-body reduced density matrix is diagonal $\rho_1^{ij} = \delta_{ij}/(2N)$, such that the eigenvalues are given by $\lambda_n = 1/(2N)$. Therefore, the entropy is again given by $S_1(V/J \rightarrow \infty) - \ln(N) = \ln(2)$ and $S_\alpha(V/J \rightarrow \infty) - \ln(N) = \ln(2)$.

- (iii) In the intermediate region, $-2 < V/J < 2$, the system is in the Tomonaga-Luttinger liquid (LL) phase, where the relatively weak interactions allow the description with an effective low-energy theory.

These phases are also visible from the one-particle von Neumann entanglement entropy S_1 : In Fig. 4.2, we show $S_1 - S_{\text{ff}}$ for $N = 51$ particles obtained numerically from the one-body reduced density matrix, where $S_{\text{ff}} = \ln(N)$ is the entropy for free fermions. At the second order transition $V/J = 2$ (dotted, vertical line), the slope changes and the entropy asymptotically approaches the theory value $\ln(2)$. At the first order phase transition $V/J = -2$ (solid, vertical line), the entropy rapidly approaches the theory value found for an infinite lattice, $\ln(2)$ [145] (dotted, horizontal line). Here, reaching an entropy of $\ln(2)$ is related to a flat spectrum of the one-body reduced density matrix (1-RDM). For $V/J < -2$, the diagonal elements of the 1-RDM scale like $\sim N^{-1}$ while the off diagonal elements are of the order $\sim N^{-2}$ such that the spectrum rapidly flattens when N becomes large – as it is the case for $N = 51$ in Fig. 4.2.

In Fig. 4.3, we additionally show the one-particle entanglement entropy for various N as a function of the interaction V/J close to the phase transition to the clustered solid phase obtained with exact diagonalization ($N < 20$) and DMRG ($N \geq 20$) (details in the following sections). From the upper left panel it becomes apparent how the value of the entropy for infinite negative interaction strength is approached: For even N , where $\cos(q_n N) = 0$ in Eq. (4.32), the entropy for $V/J \rightarrow -\infty$ is given by $\ln(2) + \ln(N)$, while for odd N , where $\cos(q_n N) = (-1)^n$, there is

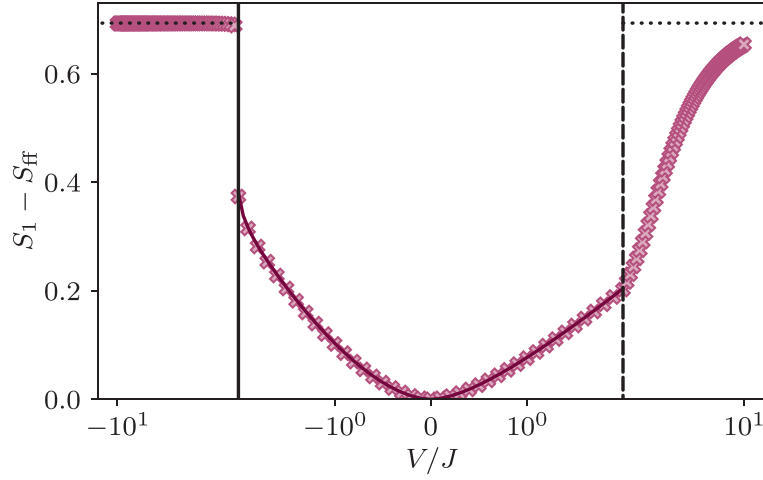


Figure 4.2: Interaction dependence of the one-particle von Neumann entanglement entropy S_1 obtained numerically from the J - V model. Here, S_{ff} is the entropy for free fermions. The equilibrium ground state entropy is obtained numerically from DMRG for a system of $L = 102$ lattice sites at half filling (crosses). The solid line represents finite size scaling of numerical data to the thermodynamic limit. The excellent agreement with the finite size DMRG data shows that the system with $N = 51$ fermions is large enough to describe the thermodynamic limit accurately in the whole LL phase.

a finite size correction $\mathcal{O}(N^{-2})$ (see App. C.2). The lower left panel shows that the increase after the phase transition becomes steeper with larger system sizes such that at $N = 51$, the finite step size in V/J values makes it look like a discrete jump, which in reality will always be a smooth but very steep increase as long as the system size is finite. The drastic increase is also a special property of the von Neumann entropy: The right panel shows the Rényi entropy for a large order $\alpha = 10$, where we find a much smoother increase even for the system with $N = 51$ fermions.

4.2.2 A simple example: two fermions on four sites

For the case of $N = 2$ fermions on a lattice with $L = 4$ sites and anti-periodic boundary condition, i.e. $c_5 = -c_1$, we can compute the entanglement entropy analytically (for a similar calculation with $N = 3$, see Ref. [308]). We therefore use this simple example to motivate and introduce the important concepts that are used in the following sections to perform the numerical computations for larger systems.

Naive basis

To obtain the ground state of the Hamiltonian, we diagonalize the matrix $\mathcal{H}_{ij} = \langle \phi_i | H | \phi_j \rangle$ of the J - V model in the basis $|\phi_i\rangle$ with the $\binom{L}{N} = 6$ elements

$$\{|0011\rangle, |0101\rangle, |0110\rangle, |1001\rangle, |1010\rangle, |1100\rangle\}. \quad (4.37)$$

We rank order the basis states according to the integer that is obtained when reading the occupation numbers as bits in a binary representation, i.e. $0011 = 3$, $0101 = 5$, $0110 = 6$, $1001 = 9$, $1010 = 10$, $1100 = 12$. We also use the convention that the basis states are created by

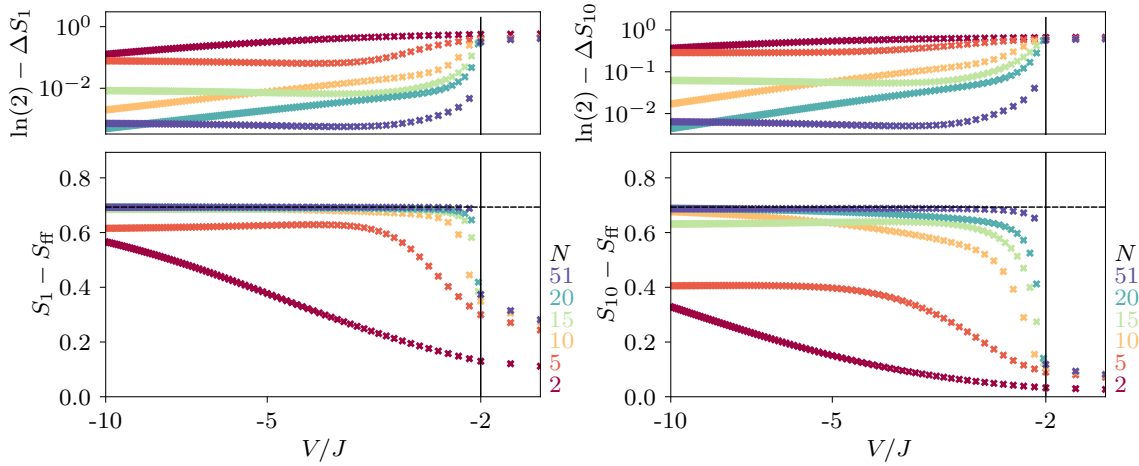


Figure 4.3: Interaction dependence of one particle von Neumann entanglement entropy S_1 (left panel) and the Rényi entropy S_{10} obtained numerically from the J - V model. Here, $S_{\text{ff}} = \ln(N)$ is the entropy for free fermions. The equilibrium ground state entropy is obtained numerically from ED for $N < 20$ and with DMRG for systems with $N > 20$ fermions at half filling. The black, vertical line marks the phase transition at $V/J = -2$ and the dashed, horizontal line shows the theory prediction $\ln(2)$ for $V/J \rightarrow -\infty$ in the thermodynamic limit. Upper panels show the difference of the entropy to this asymptotic value. For comparison, the dark blue crosses show the entropy for $L = 102$ lattice sites that is shown in Fig. 4.2. It becomes apparent that the transition is smooth for all considered finite systems and becomes steeper for larger systems and smaller Rényi power α .

acting with creation operators, starting with the smallest site index first on the empty lattice $|0\rangle$:

$$|1001\rangle = c_4^\dagger c_1^\dagger |0\rangle = -c_1^\dagger c_4^\dagger |0\rangle \quad (4.38)$$

$$|0110\rangle = c_3^\dagger c_2^\dagger |0\rangle. \quad (4.39)$$

Thus, the J - V Hamiltonian for $N = 2$, $L = 4$ has the matrix representation

$$\mathcal{H} = \begin{pmatrix} V(t) & -J & 0 & 0 & -J & 0 \\ -J & 0 & -J & -J & 0 & -J \\ 0 & -J & V(t) & 0 & -J & 0 \\ 0 & -J & 0 & V(t) & -J & 0 \\ -J & 0 & -J & -J & 0 & -J \\ 0 & -J & 0 & 0 & -J & V(t) \end{pmatrix}. \quad (4.40)$$

Diagonalizing the Hamiltonian, one finds the eigenstates and energy eigenvalues

$$\begin{aligned} \lambda_0 &= \frac{1}{2} \left(V - \sqrt{32J^2 + V^2} \right) & \mathbf{u}_0 &= \frac{1}{\sqrt{4 + 2\zeta_-^2}} (1, \zeta_-, 1, 1, \zeta_-, 1) \\ \lambda_1 &= 0 & \mathbf{u}_1 &= \frac{1}{\sqrt{2}} (0, -1, 0, 0, 1, 0) \\ \lambda_2 &= V & \mathbf{u}_2 &= \frac{1}{\sqrt{2}} (-1, 0, 0, 0, 0, 1) \end{aligned}$$

$$\begin{aligned}
\lambda_3 &= V & \mathbf{u}_3 &= \frac{1}{\sqrt{2}} (-1, 0, 0, 1, 0, 0) \\
\lambda_4 &= V & \mathbf{u}_4 &= \frac{1}{\sqrt{2}} (-1, 0, 1, 0, 0, 0) \\
\lambda_5 &= \frac{1}{2} (V + \sqrt{32J^2 + V^2}) & \mathbf{u}_5 &= \frac{1}{\sqrt{4 + 2\zeta_+^2}} (1, \zeta_+, 1, 1, \zeta_+, 1) ,
\end{aligned} \tag{4.41}$$

where we used the abbreviation

$$\zeta_{\pm}(V/J) = \frac{V \mp \sqrt{32J^2 + V^2}}{4J} . \tag{4.42}$$

Here, we find the ground state with energy λ_0 as

$$\begin{aligned}
|\Psi_0(V/J)\rangle &= \frac{1}{\sqrt{4 + 2\zeta_-^2(V/J)}} \Big[|0011\rangle + \zeta_-(V/J)|0101\rangle + |0110\rangle \\
&\quad + |1001\rangle + \zeta_-(V/J)|1010\rangle + |1100\rangle \Big] .
\end{aligned} \tag{4.43}$$

For the special case of free fermions $V/J = 0$, the coefficient ζ_- becomes $\zeta_-(0) = \sqrt{2}$.

Translational symmetry cycles

As the number of basis states is given by $\binom{L}{N}$, this calculation very quickly becomes computationally infeasible as N, L increase. However, using translational symmetry T of the Hamiltonian, i.e. $[T, H] = 0$, one can block diagonalize \mathcal{H} and focus only on a single block to obtain the ground state. Here, the unitary operator T moves all particles one place to the right, $Tn_i = n_{i+1}T$ where $n_i = c_i^\dagger c_i$ is the particle number operator at site i . Using anti-periodic (periodic) boundary conditions for odd (even) number of fermions N , ensures that the states pick up no negative signs when a fermion hops long the boundary, i.e. $Tc_L^\dagger = c_1^\dagger e^{i\pi\hat{N}}T$. When applying the translation operator to the six basis states,

$$\begin{aligned}
|\varphi_{1,1}\rangle &\equiv |0011\rangle & |\varphi_{2,1}\rangle &\equiv |0101\rangle \\
T|\varphi_{1,1}\rangle &= |1001\rangle \equiv |\varphi_{1,2}\rangle & T|\varphi_{2,1}\rangle &= |1010\rangle \equiv |\varphi_{2,2}\rangle \\
T^2|\varphi_{1,1}\rangle &= |1100\rangle \equiv |\varphi_{1,3}\rangle & T^2|\varphi_{2,1}\rangle &= |\varphi_{2,1}\rangle \\
T^3|\varphi_{1,1}\rangle &= |0110\rangle \equiv |\varphi_{1,4}\rangle & & \\
T^4|\varphi_{1,1}\rangle &= |\varphi_{1,1}\rangle , & &
\end{aligned} \tag{4.44}$$

one observes that there are two translational cycles with length $M_1 = 4$ and $M_2 = 2$ that are not mixed among each other by T . Within these cycles, the unitary operator T satisfies $T^{M_\nu} = 1$ such that $|\varphi_{\nu, M_\nu+1}\rangle = |\varphi_{\nu,1}\rangle$. We can therefore define new basis states

$$|\gamma_{\nu,q}\rangle = \frac{1}{\sqrt{M_\nu}} \sum_{m=1}^{M_\nu} e^{i\frac{2\pi q}{M_\nu}(m-1)} |\varphi_{\nu,m}\rangle = \frac{1}{\sqrt{M_\nu}} \sum_{m=1}^{M_\nu} e^{i\frac{2\pi q}{M_\nu}(m-1)} T^{m-1} |\varphi_{\nu,1}\rangle , \tag{4.45}$$

which are eigenstates of T with

$$T|\gamma_{\nu,q}\rangle = e^{-i\frac{2\pi q}{M_\nu}} |\gamma_{\nu,q}\rangle . \tag{4.46}$$

Acting with H on these states and using that $\sum_{m=1}^{M_\mu} e^{-i\frac{2\pi q}{M_\nu}(m-1)} = M_\nu \delta_{q,0}$ and $\langle \varphi_{\nu,q} | \varphi_{\nu',q'} \rangle = \delta_{qq'} \delta_{\nu\nu'}$, we find

$$H|\gamma_{1,q}\rangle = -2\sqrt{2}J\delta_{q,0}|\gamma_{2,q=0}\rangle + V|\gamma_{1,q}\rangle \quad (4.47)$$

$$H|\gamma_{2,q}\rangle = -2\sqrt{2}J\delta_{q,0}|\gamma_{1,q=0}\rangle . \quad (4.48)$$

We therefore directly found the three eigenstates $|\gamma_{1,q=1}\rangle$, $|\gamma_{1,q=2}\rangle$, $|\gamma_{1,q=3}\rangle$ with eigenvalue V , one eigenstate $|\gamma_{2,q=1}\rangle$ with eigenvalue 0, and the Hamiltonian in the basis $\{|\gamma_{2,0}\rangle, |\gamma_{1,0}\rangle, |\gamma_{2,1}\rangle, |\gamma_{1,1}\rangle, |\gamma_{1,2}\rangle, |\gamma_{1,3}\rangle\}$, sorted by the value of q , takes the block diagonal form

$$\mathcal{H} = \begin{pmatrix} 0 & -2\sqrt{2}J & 0 & 0 & 0 & 0 \\ -2\sqrt{2}J & V & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & V & 0 & 0 \\ 0 & 0 & 0 & 0 & V & 0 \\ 0 & 0 & 0 & 0 & 0 & V \end{pmatrix} . \quad (4.49)$$

Importantly, if we are only interested in the ground state, we only need to construct and diagonalize the $q = 0$ block, which has the size $n_{\text{cycles}} \times n_{\text{cycles}}$, where n_{cycles} is the number of cycles – about a factor of L smaller than the size of the Hilbert space. From this, we again recover the ground state

$$|\Psi_0\rangle = \frac{1}{\sqrt{1 + \zeta_-^2(V/J)/2}} \left[\frac{\zeta_-(V/J)}{\sqrt{2}} |\gamma_{2,0}\rangle + |\gamma_{1,0}\rangle \right] . \quad (4.50)$$

Entanglement entropy under spatial bipartition

We next split the system into two spatial regions A and B of length $\ell = 2$ and compute the entanglement entropy between the two regions. We assume region A contains sites 1 and 2 and region B the remaining sites 3 and 4, such that we can write the basis states as tensor products between the basis states of the subregions

$$|n_1 n_2 n_3 n_4\rangle = |n_1 n_2\rangle_A \otimes |n_3 n_4\rangle_B . \quad (4.51)$$

A basis in the subregions is given by

$$\begin{aligned} |\theta_m\rangle_A &\in \{|00\rangle_A, |01\rangle_A, |10\rangle_A, |11\rangle_A\} \\ |\chi_m\rangle_B &\in \{|00\rangle_B, |01\rangle_B, |10\rangle_B, |11\rangle_B\} . \end{aligned} \quad (4.52)$$

For the ground state (Eq. (4.43))

$$\begin{aligned} |\Psi_0\rangle = \frac{1}{\sqrt{4 + 2\zeta_-^2}} &\left[|00\rangle_A |11\rangle_B + \zeta_- |01\rangle_A |01\rangle_B + |01\rangle_A |10\rangle_B \right. \\ &\left. + |10\rangle_A |01\rangle_B + \zeta_- |10\rangle_A |10\rangle_B + |11\rangle_A |00\rangle_B \right] , \end{aligned} \quad (4.53)$$

the full density matrix is given by $\rho = |\Psi_0\rangle\langle\Psi_0|$. We obtain the reduced density matrix of subregion B by tracing out region A :

$$\rho_B = \text{Tr}_A \rho = \sum_{m=1}^4 \langle \theta_m |_A \rho | \theta_m \rangle_A \quad (4.54)$$

$$= (|00\rangle_B, |11\rangle_B, |01\rangle_B, |10\rangle_B) \frac{1}{4+2\zeta_-^2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \zeta_-^2 + 1 & 2\zeta_- \\ 0 & 0 & 2\zeta_- & \zeta_-^2 + 1 \end{pmatrix} \begin{pmatrix} \langle 00 |_B \\ \langle 11 |_B \\ \langle 01 |_B \\ \langle 10 |_B \end{pmatrix}. \quad (4.55)$$

Clearly, ρ_A has the same matrix representation and $\text{Tr}_B(\rho_A) = \text{Tr}_A(\rho_B) = 1$. For computing the von Neumann entanglement entropy, we need the logarithm of this matrix, which can be obtained by diagonalizing it. As the one-body reduced density matrix is already block diagonal, we only need to diagonalize the 2×2 block yielding

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}, \quad D = \frac{1}{2(2+\zeta_-^2)} \begin{pmatrix} (\zeta_- - 1)^2 & 0 \\ 0 & (\zeta_- + 1)^2 \end{pmatrix} \quad (4.56)$$

such that

$$\ln \rho_B = U \ln(D) U^\dagger = \begin{pmatrix} -\ln(4+2\zeta_-^2) & 0 & 0 & 0 \\ 0 & -\ln(4+2\zeta_-^2) & 0 & 0 \\ 0 & 0 & \ln \left| \frac{\zeta_-^2 - 1}{4+2\zeta_-^2} \right| & \ln \left| \frac{\zeta_- + 1}{\zeta_- - 1} \right| \\ 0 & 0 & \ln \left| \frac{\zeta_- + 1}{\zeta_- - 1} \right| & \ln \left| \frac{\zeta_-^2 - 1}{4+2\zeta_-^2} \right| \end{pmatrix}. \quad (4.57)$$

Thus, the von Neumann entanglement entropy is given by

$$S(\ell=2) = -\text{Tr}(\rho_B \ln \rho_B) = \frac{-1}{4+2\zeta_-^2} \left\{ 4\zeta_- \ln \left| \frac{\zeta_- + 1}{\zeta_- - 1} \right| + 2(1+\zeta_-^2) \ln \left| \frac{\zeta_- - 1}{4+2\zeta_-^2} \right| - 2 \ln [4+2\zeta_-^2] \right\}. \quad (4.58)$$

As an alternative approach for computing the entanglement entropy, we consider the Schmidt decomposition of the ground state (Sec. 4.1):

$$|\Psi_0\rangle = \sum_{ab} C_{ab} |\theta_a\rangle_A |\chi_b\rangle_B \quad (4.59)$$

$$C = \frac{1}{\sqrt{4+2\zeta_-^2}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \zeta_- & 1 \\ 0 & 0 & 1 & \zeta_- \end{pmatrix}. \quad (4.60)$$

Because C is symmetric here, we can directly diagonalize the matrix, and find

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}, \quad D = \frac{1}{\sqrt{4+2\zeta_-^2}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \zeta_- - 1 & 0 \\ 0 & 0 & 0 & \zeta_- + 1 \end{pmatrix}. \quad (4.61)$$

Hence, using the new basis

$$(|\Theta_1\rangle_A, |\Theta_2\rangle_A, |\Theta_3\rangle_A, |\Theta_4\rangle_A)^T = \left(|00\rangle_A, |11\rangle_A, \frac{|10\rangle_A - |01\rangle_A}{\sqrt{2}}, \frac{|10\rangle_A + |01\rangle_A}{\sqrt{2}} \right)^T \quad (4.62)$$

diagonalizes the reduced density matrix

$$\rho_B = \frac{1}{4 + 2\zeta_-^2} \left[|\Theta_1\rangle_A \langle \Theta_1|_A + |\Theta_2\rangle_A \langle \Theta_2|_A + (\zeta_- - 1)^2 |\Theta_3\rangle_A \langle \Theta_3|_A + (\zeta_- + 1)^2 |\Theta_4\rangle_A \langle \Theta_4|_A \right] . \quad (4.63)$$

From the matrix D , we can directly read off the von Neumann entropy

$$S_1(\ell = 2) = \frac{-1}{4 + 2\zeta_-^2} \left[2 \ln \frac{1}{4 + 2\zeta_-^2} + (\zeta_- - 1)^2 \ln \frac{(\zeta_- - 1)^2}{4 + 2\zeta_-^2} + (\zeta_- + 1)^2 \ln \frac{(\zeta_- + 1)^2}{4 + 2\zeta_-^2} \right] , \quad (4.64)$$

and the Rényi entropies

$$S_\alpha(\ell = 2) = \frac{1}{1 - \alpha} \ln \left[\frac{1}{(4 + 2\zeta_-^2)^2} \left(2^\alpha + (\zeta_- - 1)^{2\alpha} + (\zeta_- + 1)^{2\alpha} \right) \right] . \quad (4.65)$$

After rearranging terms, one indeed finds that Eq. (4.64) is equivalent to Eq. (4.58) found before. It becomes apparent that using the Schmidt decomposition makes computing the entanglement entropy much easier.

Entanglement entropy under particle bipartition

Instead of a spatial bipartition, we next consider a particle bipartition by fixing $n = 1$ particles and tracing out the second one. To define the particle bipartition [308], we need to artificially distinguish individual fermions and therefore write the basis states in anti-symmetrized first quantization notation, where subscripts indicate the particle label, e.g. $|1010\rangle = (|1_1 0 1_2 0\rangle - |1_2 0 1_1 0\rangle)/\sqrt{2}$. We introduce a new notation $|i_1 i_2\rangle$ where i_1 and i_2 are the positions of the first and second fermion, respectively. We use a sign convention $(-1)^{N_p}$ where N_p is the number of nearest neighbor permutations required to rank order the positions i_1, \dots, i_n . In this notation, we do not explicitly write down minus signs, except when computing matrix elements, e.g. $|1010\rangle = (|13\rangle + |31\rangle)/\sqrt{2}$. Therefore, the ground state, Eq. (4.43), can be written as

$$|\Psi_0\rangle = \frac{1}{\sqrt{4 + 2\zeta_-^2}} \frac{1}{\sqrt{2}} \left[(|34\rangle + |43\rangle) + \zeta_- (|24\rangle + |42\rangle) + (|23\rangle + |32\rangle) + (|14\rangle + |41\rangle) \right. \\ \left. + \zeta_- (|13\rangle + |31\rangle) + (|12\rangle + |21\rangle) \right] . \quad (4.66)$$

We could obtain the reduced density matrix from $\Psi_0(j_1, j_2) = \langle j_1, j_2 | \Psi_0 \rangle$ by tracing out the second particle position

$$\rho_A^{i_1, j_1} = \sum_{i_2=1}^4 \Psi_0^*(i_1, i_2) \Psi_0(j_1, i_2) . \quad (4.67)$$

However, it is easier to consider the Schmidt decomposition for which we define the bases in A and B as

$$\begin{aligned} |\theta_a\rangle_A &\in \{|1\rangle_A, |2\rangle_A, |3\rangle_A, |4\rangle_A\} \\ |\chi_b\rangle_B &\in \{|1\rangle_B, |2\rangle_B, |3\rangle_B, |4\rangle_B\} \end{aligned} \quad (4.68)$$

such that the ground state can be written as

$$|\Psi_0\rangle = \sum_{ab} C_{ab} |\theta_a\rangle_A |\chi_b\rangle_B \quad (4.69)$$

$$C = \frac{1}{\sqrt{4+2\zeta_-^2}} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -1 & -\zeta_- & -1 \\ 1 & 0 & -1 & -\zeta_- \\ \zeta_- & 1 & 0 & -1 \\ 1 & \zeta_- & 1 & 0 \end{pmatrix}. \quad (4.70)$$

Here, we reinserted the signs when constructing the coefficient matrix C . Computing the singular value decomposition $C = U\Sigma V^T$, we obtain the diagonal form of the reduced density matrix (see Sec. 4.1) with eigenvalues $\text{diag}(\Sigma^2) = \frac{1}{2(4+2\zeta_-^2)} \left((\sqrt{2}-\zeta_-)^2, (\sqrt{2}-\zeta_-)^2, (\sqrt{2}+\zeta_-)^2, (\sqrt{2}+\zeta_-)^2 \right)$. We can therefore directly write down the von Neumann one particle entanglement entropy

$$S_1(n=1) = \frac{-1}{4+2\zeta_-^2} \left[(\sqrt{2}-\zeta_-)^2 \ln \left(\frac{(\sqrt{2}-\zeta_-)^2}{2(4+2\zeta_-^2)} \right) + (\sqrt{2}+\zeta_-)^2 \ln \left(\frac{(\sqrt{2}+\zeta_-)^2}{2(4+2\zeta_-^2)} \right) \right] \quad (4.71)$$

and the Rényi entropy of order α

$$S_\alpha(n=1) = \frac{1}{1-\alpha} \ln \left[\frac{2}{2^\alpha(4+2\zeta_-^2)^\alpha} \left((\sqrt{2}-\zeta_-)^{2\alpha} + (\sqrt{2}+\zeta_-)^{2\alpha} \right) \right]. \quad (4.72)$$

Particle entanglement entropy from the correlation matrix

There is yet another way to compute the one-body reduced density matrix $\rho_A(n=1)$ for N fermions that is based on the correlation function

$$\mathcal{C}_{ij} = \langle \Psi_0 | c_i^\dagger c_j | \Psi_0 \rangle \quad (4.73)$$

$$\rho_A(n=1) = \frac{1}{N} \mathcal{C}. \quad (4.74)$$

For the ground state Eq. (4.43), this yields

$$\rho_A(n=1) = \frac{1}{2(4+2\zeta_-^2)} \begin{pmatrix} 2+\zeta_-^2 & 2\zeta_- & 0 & -2\zeta_- \\ 2\zeta_- & 2+\zeta_-^2 & 2\zeta_- & 0 \\ 0 & 2\zeta_- & 2+\zeta_-^2 & 2\zeta_- \\ -2\zeta_- & 0 & 2\zeta_- & 2+\zeta_-^2 \end{pmatrix}. \quad (4.75)$$

Due to the translational symmetry, it is even sufficient to only compute the first row, and diagonalization reduces to a Fourier transform. Determining the correct signs, however, is a non-trivial problem. For a fermion operator c_i or c_i^\dagger acting on the state $|n_1 n_2 n_3 n_4\rangle$, the sign can be obtained from the number of occupied sites before site i

$$\text{sgn}(c_i, |n_1 n_2 n_3 n_4\rangle) = (-1)^{\sum_{k<i} n_k}. \quad (4.76)$$

The sign of a term in \mathcal{C}_{ij} is therefore given by $(-1)^{\sum_{k<i} (1-\delta_{kj}) n_k} (-1)^{\sum_{k<j} n_k}$. The eigenvalues of $\rho_A(n=1)$, Eq. (4.75), are given by

$$\frac{1}{2(4+2\zeta_-^2)} \left((\sqrt{2}-\zeta_-)^2, (\sqrt{2}-\zeta_-)^2, (\sqrt{2}+\zeta_-)^2, (\sqrt{2}+\zeta_-)^2 \right), \quad (4.77)$$

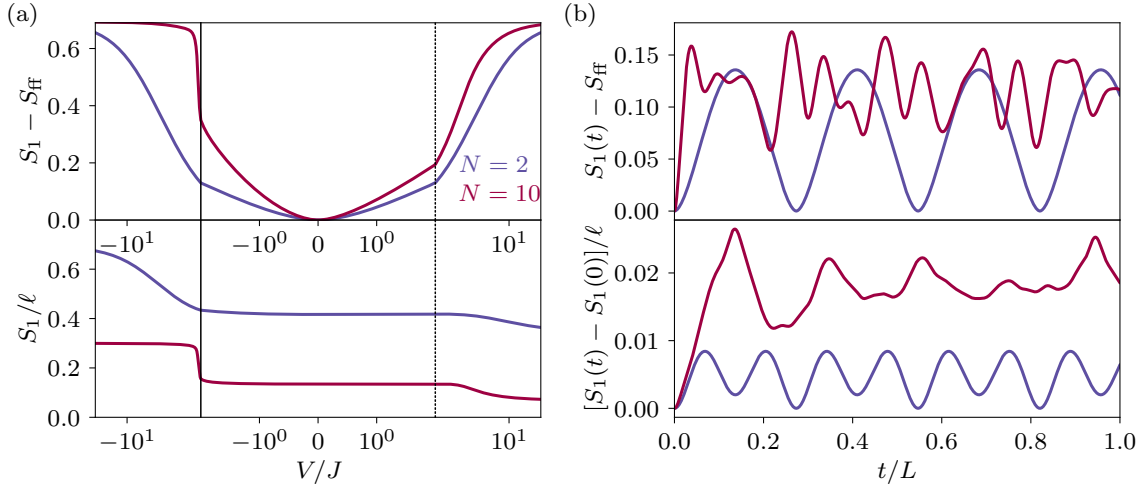


Figure 4.4: Entanglement entropy for the simple example of $N = 2$ fermions on $L = 4$ lattice sites (blue). (a) Ground state von Neumann entropies, Eqs. (4.71) and (4.64), under particle bipartition with $n = 1$ (upper panel) and under spatial bipartition with $\ell = 2$ (lower panel). (b) Waiting time t dependence of the growth of entanglement entropy after a quantum quench from free fermions to $V/J = 1.0$, Eqs. (4.94) and (4.89). For particle entanglement, we subtract the entropy for free fermions $S_{\text{ff}} = \ln(N)$. We note that finite size effects dominate for such a small system. For comparison, we show the numerical results obtained with exact diagonalization for $N = 10, L = 20, n = 1, \ell = 10$ in red.

as obtained before, such that we find the same entanglement entropy Eq. (4.71) and Eq. (4.72).

In Fig. 4.4a, we show the von Neumann entanglement entropy for $N = 2$ fermions on $L = 4$ sites as a function of the interaction strength V/J . For the one-particle entanglement entropy, we additionally subtract the entropy for free fermions $S_{\text{ff}} = S_1(n = 1, V/J = 0) = \ln(N)$. At the phase transitions $V/J = \pm 2$ (vertical black lines), we observe qualitative changes in the slopes of the entropy (blue). However, a comparison with numerical results for a larger system (red) shows, that some qualitative features are not visible due to finite size effects in the small system.

Time evolutions after an interaction quantum quench

We next consider an interaction quantum quench, i.e., we start from $N = 2$ free fermions on $L = 4$ lattice sites in the ground state $|\Psi_0(0)\rangle$ and at time $t = 0$, we suddenly switch on interactions with strength V/J such that $V(t)/J = \theta(t)V/J$. For $t < 0$ the ground state with energy $-2\sqrt{2}J$ is given by

$$|\Psi(t \leq 0; V/J = 0)\rangle = \frac{1}{2\sqrt{2}} \left(|0011\rangle + \sqrt{2}|0101\rangle + |0110\rangle + |1001\rangle + \sqrt{2}|1010\rangle + |1100\rangle \right), \quad (4.78)$$

according to Eq. (4.43) where $V/J \rightarrow 0$ corresponds to $\zeta_- \rightarrow \sqrt{2}$. After the quench, $t > 0$, we compute the time evolution of $|\Psi_0(t)\rangle$ using the full eigensystem $|\Psi_\alpha\rangle, E_\alpha$, Eq. (4.41), of the post quench Hamiltonian as

$$|\Psi(t)\rangle = \sum_{\alpha} e^{-iE_\alpha t} \langle \Psi_\alpha | \Psi_0(0) \rangle |\Psi_\alpha\rangle. \quad (4.79)$$

For this, we need the full eigenstates $|\Psi_\alpha\rangle$ – at least of the $q = 0$ block of the Hamiltonian as $|\Psi_0(0)\rangle$ is still orthogonal to the post quench $q > 0$ block for all V/J . We find the overlaps

$$\langle\Psi_i|\Psi_0(0)\rangle = 0 \quad \text{for } 1 \leq i \leq 4 \quad (4.80)$$

$$\langle\Psi_0|\Psi_0(0)\rangle = \frac{1}{2\sqrt{2}} \frac{1}{\sqrt{4+2\zeta_-^2}} (4+2\sqrt{2}\zeta_-) \quad (4.81)$$

$$\langle\Psi_5|\Psi_0(0)\rangle = \frac{1}{2\sqrt{2}} \frac{1}{\sqrt{4+2\zeta_+^2}} (4+2\sqrt{2}\zeta_+) \quad (4.82)$$

and energy eigenvalues $E_0 = (V - \sqrt{32J^2 + V^2})/2$, $E_5 = (V + \sqrt{32J^2 + V^2})/2$. Hence, the time evolution of the initial ground state at waiting time t after the quench is given by

$$\begin{aligned} |\Psi(t)\rangle &= [\xi_-(t) + \xi_+(t)](|0011\rangle + |0110\rangle + |1001\rangle + |1100\rangle) \\ &\quad + [\zeta_-\xi_-(t) + \zeta_+\xi_+(t)](|0101\rangle + |1010\rangle), \end{aligned} \quad (4.83)$$

where we defined the abbreviations

$$\xi_\pm(t) = \frac{1}{2\sqrt{2}} e^{-it(V \pm \sqrt{32J^2 + V^2})/2} \frac{4 + 2\sqrt{2}\zeta_\pm}{4 + 2\zeta_\pm^2} \quad (4.84)$$

$$\zeta_\pm = \frac{V \mp \sqrt{32J^2 + V^2}}{4J}. \quad (4.85)$$

Waiting time dependence of spatial entanglement entropy

Equipped with the time evolution of the state after the quench, we can compute the waiting time dependence of the spatial entanglement entropy by first constructing the Schmidt decomposition (with basis Eq. (4.52))

$$|\Psi(t)\rangle = \sum_{ab} C_{ab} |\theta_a\rangle_A |\chi_b\rangle_B \quad (4.86)$$

$$C = \begin{pmatrix} 0 & \xi_- + \xi_+ & 0 & 0 \\ \xi_- + \xi_+ & 0 & 0 & 0 \\ 0 & 0 & \zeta_-\xi_- + \zeta_+\xi_+ & \xi_- + \xi_+ \\ 0 & 0 & \xi_- + \xi_+ & \zeta_-\xi_- + \zeta_+\xi_+ \end{pmatrix}. \quad (4.87)$$

The magnitudes of squared singular values of the coefficient matrix are given by

$$\text{diag}\Sigma^2 = \left(|\xi_-(\zeta_- - 1) + \xi_+(\zeta_+ - 1)|^2, |\xi_+ + \xi_-|^2, |\xi_+ + \xi_-|^2, |\xi_-(\zeta_- + 1) + \xi_+(\zeta_+ + 1)|^2 \right), \quad (4.88)$$

such that we find the entropies

$$\begin{aligned} S_1(t; \ell = 2) &= - \left[4 |\xi_+ + \xi_-|^2 \ln |\xi_+ + \xi_-| \right. \\ &\quad + 2 |\xi_-(\zeta_- - 1) + \xi_+(\zeta_+ - 1)|^2 \ln |\xi_-(\zeta_- - 1) + \xi_+(\zeta_+ - 1)| \\ &\quad \left. + 2 |\xi_-(\zeta_- + 1) + \xi_+(\zeta_+ + 1)|^2 \ln |\xi_-(\zeta_- + 1) + \xi_+(\zeta_+ + 1)| \right], \end{aligned} \quad (4.89)$$

$$\begin{aligned} S_\alpha(t, \ell = 2) &= \frac{1}{1-\alpha} \ln \left[2 |\xi_+ + \xi_-|^{2\alpha} \right. \\ &\quad + |\xi_-(\zeta_- - 1) + \xi_+(\zeta_+ - 1)|^{2\alpha} \\ &\quad \left. + |\xi_-(\zeta_- + 1) + \xi_+(\zeta_+ + 1)|^{2\alpha} \right]. \end{aligned} \quad (4.90)$$

Waiting time dependence of particle entanglement entropy

For the particle bipartition, the Schmidt decomposition takes the form (basis: Eq. (4.68))

$$|\Psi(t)\rangle = \sum_{ab} C_{ab} |\theta_a\rangle_A |\chi_b\rangle_B \quad (4.91)$$

$$C = \begin{pmatrix} 0 & -(\xi_+ + \xi_-) & -(\zeta_- \xi_- + \zeta_+ \xi_+) & -(\xi_+ + \xi_-) \\ \xi_+ + \xi_- & 0 & -(\xi_+ + \xi_-) & -(\zeta_- \xi_- + \zeta_+ \xi_+) \\ \zeta_- \xi_- + \zeta_+ \xi_+ & \xi_+ + \xi_- & 0 & -(\xi_+ + \xi_-) \\ \xi_+ + \xi_- & \zeta_- \xi_- + \zeta_+ \xi_+ & \xi_+ + \xi_- & 0 \end{pmatrix}. \quad (4.92)$$

The coefficient matrix has singular values with squared magnitudes given by

$$\text{diag} \Sigma^2 = \frac{1}{2} \left(\left| \xi_- (\zeta_- - \sqrt{2}) + \xi_+ (\zeta_+ - \sqrt{2}) \right|^2, \left| \xi_- (\zeta_- - \sqrt{2}) + \xi_+ (\zeta_+ - \sqrt{2}) \right|^2, \right. \\ \left. \left| \xi_- (\zeta_- + \sqrt{2}) + \xi_+ (\zeta_+ + \sqrt{2}) \right|^2, \left| \xi_- (\zeta_- + \sqrt{2}) + \xi_+ (\zeta_+ + \sqrt{2}) \right|^2 \right), \quad (4.93)$$

such that the one-particle entanglement entropy is

$$S_1(t; n=1) = - \left[\left| \xi_- (\zeta_- - \sqrt{2}) + \xi_+ (\zeta_+ - \sqrt{2}) \right|^2 \ln \left(\frac{1}{2} \left| \xi_- (\zeta_- - \sqrt{2}) + \xi_+ (\zeta_+ - \sqrt{2}) \right|^2 \right) \right. \\ \left. + \left| \xi_- (\zeta_- + \sqrt{2}) + \xi_+ (\zeta_+ + \sqrt{2}) \right|^2 \ln \left(\frac{1}{2} \left| \xi_- (\zeta_- + \sqrt{2}) + \xi_+ (\zeta_+ + \sqrt{2}) \right|^2 \right) \right] \quad (4.94)$$

$$S_\alpha(t; n=1) = \frac{1}{1-\alpha} \ln \left[\frac{1}{2^{\alpha-1}} \left| \xi_- (\zeta_- - \sqrt{2}) + \xi_+ (\zeta_+ - \sqrt{2}) \right|^{2\alpha} \right. \\ \left. + \frac{1}{2^{\alpha-1}} \left| \xi_- (\zeta_- + \sqrt{2}) + \xi_+ (\zeta_+ + \sqrt{2}) \right|^{2\alpha} \right] \quad (4.95)$$

In Fig. 4.4b, we show the waiting time dependence of the increase in spatial (Eq. (4.94)) and particle entanglement entropy (Eq. (4.89)) after the quench.

Unfortunately, finite size effects dominate for such small systems such that many important features are not visible in the results for $N = 2$. We therefore additionally plot exact diagonalization results for a much larger system with $N = 10$ particles on $L = 20$ lattice sites (red). From this it becomes apparent that the initial increase for the spatial entanglement entropy is much slower, $\propto t$, compared to the particle entanglement [308]. Nevertheless, the calculations shown in this section can be straightforwardly expanded to larger systems and therefore allows getting insight into the methods implemented in the code [474] that is used for the numerical simulations.

4.3 Mapping onto an effective low-energy Luttinger liquid model

This section closely follows the publication: Matthias Thamm, Harini Radhakrishnan, Hatem Barghathi, Bernd Rosenow, and Adrian Del Maestro, *One-particle entanglement for one dimensional spinless fermions after an interaction quantum quench*, Physical Review B 106, 165116 (2022) [327].

4.3.1 One-particle entanglement entropy in the equilibrium case

We start with deriving an analytical result for the one body reduced density matrix (1-RDM) $\rho_1(x, 0)$ for the corresponding LL model of length L . Here, we measure lengths in units of the lattice constant. From the LL 1-RDM, we compute the one-particle entanglement entropy, and then compare with numerical results obtained for the J - V model. In this phase with intermediate interaction strength, observables are dominated by low energy excitations in the form of density fluctuations around a static average density background. Such fluctuations of the density are bosonic in nature, which allows us to describe the low energy physics with an effective Hamiltonian [163] in bosonization notation after linearizing the dispersion around the Fermi points

$$H = \sum_{q \neq 0} [\omega_0(q) + m(q)] b_q^\dagger b_q + \frac{1}{2} \sum_{q \neq 0} g_2(q) (b_q b_{-q} + b_q^\dagger b_{-q}^\dagger), \quad (4.96)$$

where b_q^\dagger (b_q) are bosonic creation (annihilation) operators with $[b_q, b_{q'}^\dagger] = \delta_{q, q'}$, $\omega_0(q) = v_F |q|$. We work in units where $\hbar = 1$ and measure lengths in units of the lattice constant. The sum is taken over discrete momenta $q_n = n2\pi/L$ with $n \in \mathbb{Z} \setminus \{0\}$. The nearest neighbor coupling in the lattice model has a finite interaction range, which we take into account by assuming that $g_2(q)$ and $m(q)$ vanish for momenta $q\varepsilon \gg 1$ larger than an interaction cutoff ε (for the detailed implementation of the cutoff procedure see Eq. (4.118)). For small $|q|$, the parameters have a linear q dependence, $m(q) = g_4 |q|$, $g_2(q) = g_2 |q|$, where g_2 and g_4 can be related to the parameters of the J - V model as discussed below. As $\omega_0(q)$ is the dispersion for free fermions, the terms g_2 and g_4 are zero in this case. By comparing the final bosonization results with numerical simulations of the J - V chain the interaction cutoff can be unambiguously determined. The Hamiltonian Eq. (4.96) is quadratic in the boson operators and can therefore be solved analytically. In order to compute the one body density matrix, we use refermionization to express the fermionic field operators $\psi_\alpha(x)$ in terms of bosonic fields as

$$\psi_\alpha(x) = \frac{\chi_\alpha}{\sqrt{2\pi\eta}} e^{i(\varphi_{0,\alpha} + \alpha \frac{2\pi x}{L} N_\alpha)} e^{-i\alpha\phi_\alpha(x)} \quad (4.97)$$

$$\phi_\alpha(x) = - \sum_{q>0} \sqrt{\frac{2\pi}{qL}} e^{-q\eta/2} \left[e^{i\alpha q x} b_{\alpha q} + e^{-i\alpha q x} b_{\alpha q}^\dagger \right], \quad (4.98)$$

where $\alpha = (-)1$ indicates right (left) moving fermions, $\chi_\alpha = e^{\alpha i \frac{\pi}{2} N_\alpha}$ are Klein factors with $\chi_\alpha^\dagger \chi_\alpha = 1$, η is a short distance cutoff measured in units of the lattice spacing (not to be confused with the interaction cutoff ε), and $\phi_\alpha(x)$ are Hermitian operators [162, 326]. Here, N_α is the particle number operator, and $\varphi_{0,\alpha}$, N_α are zero mode operators satisfying the commutation relation $[N_\alpha, \varphi_{0,\alpha}] = i$. The one-body density matrix can be obtained from the one point correlation functions for left and right movers in terms of the fermion operators Eq. (4.97) via

$$\rho_1(x, 0) = \frac{1}{N} \left[e^{-ik_F x} C_+(x, 0) + e^{ik_F x} C_-(x, 0) \right] \quad (4.99)$$

$$C_\alpha(x, 0) = \langle \psi_\alpha^\dagger(x) \psi_\alpha(0) \rangle \quad (4.100)$$

with Fermi momentum $k_F = \pi N/L$. To relate the results for the effective LL model to numerical results of the J - V model at half filling, we use Bethe ansatz results obtained via a mapping to

the spin-1/2 XXZ chain [326]

$$K \equiv \sqrt{\frac{v_F + g_4 + g_2}{v_F + g_4 - g_2}} = \frac{\pi}{2 \cos^{-1} \left(\frac{-V}{2J} \right)} \quad (4.101)$$

$$\frac{v}{J} = \frac{1}{1 - (2K)^{-1}} \sin \left[\pi(1 - (2K)^{-1}) \right] , \quad (4.102)$$

where K is the LL interaction parameter, and $v|q|$ is the dispersion relation for low energy excitations. We use the above expressions for v and K in the diagonalized version of the LL Hamiltonian Eq. (4.96) to parametrize the interaction strength and velocity.

The Hamiltonian Eq. (4.96) can be diagonalized using a Bogoliubov transformation

$$\begin{aligned} a_q &= \cosh(\theta_q) b_q + \sinh(\theta_q) b_{-q}^\dagger \\ a_{-q}^\dagger &= \sinh(\theta_q) b_q + \cosh(\theta_q) b_{-q}^\dagger , \end{aligned} \quad (4.103)$$

in contrast to simply diagonalizing the Hamiltonian in a basis (b_q^\dagger, b_{-q}) (as one would do for a fermionic BCS Hamiltonian), since this would not preserve the bosonic commutation relations [478, 479]. The choice of coefficients in Eq. (4.103) guarantees bosonic commutation relations $[a_q, a_{q'}^\dagger] = \delta_{q,q'}$, $[a_q, a_{q'}] = 0$, $[a_q^\dagger, a_{q'}^\dagger] = 0$, and one finds that

$$\begin{aligned} & \sum_q f(|q|) \frac{a_q^\dagger a_q}{\cosh^2(\theta_q) + \sinh^2(\theta_q)} \\ &= \sum_q f(|q|) b_q^\dagger b_q + f(|q|) \frac{\sinh(\theta_q) \cosh(\theta_q)}{\sinh^2(\theta_q) + \cosh^2(\theta_q)} (b_q b_{-q} + b_q^\dagger b_{-q}^\dagger) . \end{aligned} \quad (4.104)$$

Choosing $f(|q|) = \omega_0(q) + m(q)$ and $\tanh(2\theta_q) = g_2(q)/f(|q|)$, which in the limit $q \rightarrow 0$ is given by $g_2/(v_F + g_4)$, the Hamiltonian becomes diagonal

$$H = \sum_q \omega(q) a_q^\dagger a_q \quad (4.105)$$

$$\omega(q) = \sqrt{(\omega_0(q) + m(q))^2 - g_2(q)^2} \equiv v|q| . \quad (4.106)$$

This allows us to evaluate the ground state expectation values

$$\langle a_q^\dagger a_{q'} \rangle = \delta_{qq'} f_b(q) = \delta_{qq'} - \langle a_q a_{q'}^\dagger \rangle \quad (4.107)$$

$$\langle a_q a_{q'} \rangle = 0 = \langle a_q^\dagger a_{q'}^\dagger \rangle . \quad (4.108)$$

where $f_b(q)$ is the Bose-Einstein distribution function with energies $\omega(q)$.

Using Eq. (4.97) in Eq. (4.99) together with the Baker-Hausdorff formula $e^A e^B = e^{A+B} e^{[A,B]/2}$, the one point correlation function becomes

$$C_\alpha(x, 0) = \frac{1}{2\pi\eta} e^{\alpha \frac{\pi x}{L} [N_\alpha, \varphi_{0,\alpha}]} e^{\frac{1}{2} [\phi_\alpha(x), \phi_\alpha(0)]} \langle e^{i\alpha(\phi_\alpha(x) - \phi_\alpha(0))} \rangle . \quad (4.109)$$

Here, we use the boson cumulant formula $\langle e^{i\alpha(\phi_\alpha(x) - \phi_\alpha(0))} \rangle = e^{-\frac{1}{2} \langle (\phi_\alpha(x) - \phi_\alpha(0))^2 \rangle}$, which is valid in equilibrium for a quadratic Hamiltonian, for any linear combination of bosons $\sum_n A_n b_n + B_n b_n^\dagger$. In addition,

$$\frac{1}{2} [\phi_\alpha(x), \phi_\alpha(0)] = \frac{1}{2} \sum_{q>0} \frac{2\pi}{qL} \left[e^{-qi(-i\eta - \alpha x)} - e^{-iq(-i\eta + \alpha x)} \right] . \quad (4.110)$$

Due to the regularization η , we can perform the q sum, where we use that for any complex number z with $\text{Im}[z] > 0$ holds [162]

$$\sum_{q>0} \frac{2\pi}{qL} e^{-iqz} = -\ln \left[1 - e^{-i\frac{2\pi}{L}z} \right] = -\ln \left[2ie^{-i\frac{\pi}{L}z} \sin \left(\frac{\pi}{L}z \right) \right]. \quad (4.111)$$

One needs to be careful when using logarithm laws with complex numbers, as we need to stay on the main branch of the logarithm: $\ln(z) = \ln(|z|) + i \arg(z)$. With this in mind, we find for Eq. (4.110)

$$\begin{aligned} \frac{1}{2}[\phi_\alpha(x), \phi_\alpha(0)] &= \frac{1}{2} \left\{ -\ln \left[2ie^{-i\frac{\pi}{L}(-\alpha x - i\eta)} \sin \left(\frac{\pi}{L}(\alpha x + i\eta) \right) \right] \right. \\ &\quad \left. + \ln \left[-2ie^{i\frac{\pi}{L}(-\alpha x + i\eta)} \sin \left(\frac{\pi}{L}(\alpha x - i\eta) \right) \right] \right\} \end{aligned} \quad (4.112)$$

$$= -\frac{1}{2} \ln \left| \frac{\sin \left(\frac{\pi}{L}(\alpha x + i\eta) \right)}{\sin \left(\frac{\pi}{L}(\alpha x - i\eta) \right)} \right| - i \arg \left[\sin \left(\frac{\pi}{L}(\alpha x + i\eta) \right) \right] - i\frac{\pi}{2} - i\frac{\pi}{L}\alpha x. \quad (4.113)$$

In the limit $\eta/x \rightarrow 0$, the \arg term is 0 for $\alpha x > 0$ and $\pm\pi$ if $\alpha x < 0$ such that

$$\lim_{\eta/x \rightarrow 0} e^{\frac{1}{2}[\phi_\alpha(x), \phi_\alpha(0)]} = -\text{sgn}(\alpha x) i e^{-i\frac{\pi}{L}\alpha x}. \quad (4.114)$$

In order to evaluate the expectation value in Eq. (4.109) by using the boson cumulant formula, we need the expectation values $\langle \phi_\alpha(x) \phi_\alpha(x') \rangle$. To compute them by utilizing the expectation values Eq. (4.107), we insert the inverse of the transformation Eq. (4.103) into the expression for $\phi_\alpha(x)$, Eq. (4.98), such that

$$\begin{aligned} \phi_\alpha(x) &= -\sum_{q>0} \sqrt{\frac{2\pi}{qL}} e^{-q\eta/2} \left[e^{i\alpha qx} \left(\cosh(\theta_q) a_q - \sinh(\theta_q) a_{-q}^\dagger \right) \right. \\ &\quad \left. + e^{-i\alpha qx} \left(\cosh(\theta_q) a_q^\dagger - \sinh(\theta_q) a_{-q} \right) \right]. \end{aligned} \quad (4.115)$$

Using the expectation values of pairs for a_q operators with $\langle a_q^\dagger a_{q'} \rangle = \delta_{qq'} f_b(q)$, we find

$$\begin{aligned} \langle \phi_\alpha(x) \phi_\alpha(x') \rangle &= \sum_{q>0} \frac{2\pi}{qL} e^{-q\eta} \left\{ e^{i\alpha(x-x')} \left[(1 - f_b(q)) \cosh^2(\theta_q) + f_b(q) \sinh^2(\theta_q) \right] \right. \\ &\quad \left. + e^{-i\alpha(x-x')} \left[f_b(q) \cosh^2(\theta_q) + (1 - f_b(q)) \sinh^2(\theta_q) \right] \right\}. \end{aligned} \quad (4.116)$$

At zero temperature, the Bose-Einstein distribution becomes $f_b(q > 0) = 0$, such that we find for the exponent appearing in the correlation function

$$\begin{aligned} -\frac{1}{2} \langle (\phi_\alpha(x) - \phi_\alpha(0))^2 \rangle &= \frac{1}{2} \sum_{q>0} \frac{2\pi}{qL} e^{-q\eta} \left(\cosh^2(\theta_q) + \sinh^2(\theta_q) - 1 + 1 \right) \\ &\quad \times \left[-2 + e^{i\alpha qx} + e^{-i\alpha qx} \right]. \end{aligned} \quad (4.117)$$

Here, we added a zero $(-1 + 1)$ to separate the free term Eq. (4.111) from the interaction term of the correlation function. Including the -1 in the interaction term ensures that it vanishes in

the non-interacting case where $\theta_q \rightarrow 0$. We now precisely define the interaction cutoff ε by using it to describe the q dependence of the interaction term as [163]

$$\cosh^2(\theta_q) + \sinh^2(\theta_q) - 1 \approx \frac{K + K^{-1} - 2}{2} e^{-\varepsilon|q|} \equiv \gamma_{\text{eq}}^2 e^{-\varepsilon|q|} , \quad (4.118)$$

where now $K = \lim_{q \rightarrow 0} e^{2\theta_q}$ and γ_{eq} are independent of q . While ε appears to be a free parameter of the model, we will show later that for not too strong interactions, a fixed value can be chosen such that the analytic calculation reproduces numerical results from exact diagonalization and DMRG for a range of interaction strengths. In addition, ε regularizes the interaction part of the correlation function and therefore allows us to take the limit $\eta q \rightarrow 0$ when keeping εq finite.

At zero temperature, we use Eq. (4.111) to perform the q sums in Eq. (4.117), and find

$$F_\alpha^0(x; \eta) \equiv -\frac{1}{2} \langle (\phi_\alpha(x) - \phi_\alpha(0))^2 \rangle_0 \quad (4.119)$$

$$= \frac{1}{2} \sum_{q>0} \frac{2\pi}{qL} e^{-iq(-i\eta)} \left[-2 + e^{i\alpha qx} + e^{-i\alpha qx} \right] \quad (4.120)$$

$$\begin{aligned} &= \ln \left[-2\mathbf{i} e^{-\frac{\pi}{L}\eta} \sin\left(\frac{\pi}{L}\mathbf{i}\eta\right) \right] \\ &\quad - \frac{1}{2} \ln \left[2\mathbf{i} e^{-i\frac{\pi}{L}(\alpha x - i\eta)} \sin\left(\frac{\pi}{L}(\alpha x - \mathbf{i}\eta)\right) \right] \\ &\quad - \frac{1}{2} \ln \left[2\mathbf{i} e^{-i\frac{\pi}{L}(\alpha x + i\eta)} \sin\left(\frac{\pi}{L}(\alpha x + \mathbf{i}\eta)\right) \right] , \end{aligned} \quad (4.121)$$

such that the interaction term can be obtained from the free one by multiplying with a factor γ_{eq}^2 while changing the regularization to include the interaction cutoff, i.e. $\eta \rightarrow \eta + \varepsilon$,

$$-\frac{1}{2} \left[\langle (\phi_\alpha(x) - \phi_\alpha(0))^2 \rangle - \langle (\phi_\alpha(x) - \phi_\alpha(0))^2 \rangle_0 \right] = \gamma_{\text{eq}}^2 F_\alpha^0(x; \eta + \varepsilon) . \quad (4.122)$$

We use this and the translational invariance of the expectation value, i.e. $\langle \phi_\alpha(x) \phi_\alpha(x') \rangle = \langle \phi_\alpha(x - x') \phi_\alpha(0) \rangle$, to obtain the expectation value that appears in the one point correlation function

$$e^{-\frac{1}{2} \langle (\phi_\alpha(x) - \phi_\alpha(0))^2 \rangle} = \frac{-\mathbf{i} \sin\left(\frac{\pi}{L}\mathbf{i}\eta\right)}{\left| \sin\left(\frac{\pi}{L}(\alpha x + \mathbf{i}\eta)\right) \right|} \left[\frac{\mathbf{i} \sin\left(\frac{\pi}{L}\mathbf{i}(\eta + \varepsilon)\right)}{\left| \sin\left(\frac{\pi}{L}(\alpha x + \mathbf{i}(\eta + \varepsilon))\right) \right|} \right]^{\gamma_{\text{eq}}^2} . \quad (4.123)$$

Using Eq. (4.114) and Eq. (4.123) in the expression for the correlation function Eq. (4.109), taking the limit $\eta/x, \eta/L \rightarrow 0$, where $\sin(\pi\mathbf{i}\eta/L)/\eta \rightarrow \pi\mathbf{i}/L$, $\mathbf{i} \sin(\mathbf{i}b) = -|\sin(\mathbf{i}b)|$, and $\sqrt{\sin(b + \mathbf{i}c) \sin(-b + \mathbf{i}c)} = \mathbf{i} |\sin(b + \mathbf{i}c)|$, we find

$$C_\alpha(x, 0) = \frac{\mathbf{i}\pi}{2\pi L} \frac{\text{sgn}(\alpha x)}{\left| \sin\left(\frac{\pi}{L}(\alpha x)\right) \right|} \left| \frac{\sin\left(\frac{\pi}{L}\mathbf{i}\varepsilon\right)}{\sin\left(\frac{\pi}{L}(\alpha x + \mathbf{i}\varepsilon)\right)} \right|^{\gamma_{\text{eq}}^2} \quad (4.124)$$

$$= \frac{\alpha\mathbf{i}}{2\sin(\pi x/L)} \left| \frac{\sin(\pi\mathbf{i}\varepsilon/L)}{\sin\left(\frac{\pi}{L}(x + \mathbf{i}\varepsilon)\right)} \right|^{\gamma_{\text{eq}}^2} . \quad (4.125)$$

Using that $\alpha = -1$ for left movers and $\alpha = +1$ for right movers, we obtain the full one body

density matrix Eq. (4.99) [163]

$$\rho_1(x, 0) = \rho_1^0(x, 0) \left| \frac{\sin(\pi \mathbf{i} \varepsilon / L)}{\sin(\frac{\pi}{L}(x + \mathbf{i} \varepsilon))} \right|^{\gamma_{\text{eq}}^2} \quad (4.126)$$

$$\rho_1^0(x, 0) = \frac{1}{N} \frac{\sin(k_F x)}{L \sin(\pi x / L)} . \quad (4.127)$$

Here, Eq. (4.127) is equivalent to the exact one body density matrix for non-interacting lattice fermions [327].

Because x is a relative coordinate, and we are interested in the short distance behavior that dominates the Fourier transform and the entropy, we consider the limit of large L with $x/L \ll 1$ and neglect terms of order $\mathcal{O}(x/L)$. For the leading term $L \sin(\pi x/L) \rightarrow \pi x$ we then arrive at the following expression for the one body density matrix:

$$\rho_1(x, 0) = \frac{\sin(k_F x)}{N \pi x} \left(\frac{\varepsilon^2}{x^2 + \varepsilon^2} \right)^{\gamma_{\text{eq}}^2/2} + \mathcal{O}\left(\frac{x}{L}\right) , \quad (4.128)$$

which is normalized such that $L\rho_1(0, 0) = 1$ where $k_F = \pi N/L$. Because the particle number N appears explicitly in the normalization, we cannot directly take the thermodynamic limit. Therefore, we first compute the entropy density, and only then take the thermodynamic limit $1/N \rightarrow 0$. To diagonalize $\rho_1(x, 0)$, we compute the Fourier transform which yields

$$\begin{aligned} \rho_1(q) &= \int_{-\infty}^{\infty} dx \rho_1(x, 0) e^{-iqx} \\ &= \frac{\Gamma[\frac{1}{2}(-1 + \gamma_{\text{eq}}^2)]\sqrt{\pi}}{2\pi N \Gamma(\gamma_{\text{eq}}^2/2)} [f_1(\tilde{q}) + f_1(-\tilde{q})] \\ &\quad - \frac{2\Gamma(-\gamma_{\text{eq}}^2) \sin(\pi \gamma_{\text{eq}}^2/2)}{2\pi N} [f_2(\tilde{q}) + f_2(-\tilde{q})] \end{aligned} \quad (4.129)$$

where $\tilde{q} = \varepsilon q$, $\tilde{k}_F = \varepsilon k_F$, $L/(2\pi) \int dq \rho_1(q) = 1$, and

$$\begin{aligned} f_1(\tilde{q}) &= (\tilde{k}_F + \tilde{q}) {}_1F_2 \left[\left\{ \frac{1}{2} \right\}, \left\{ \frac{3}{2}, \frac{3 - \gamma_{\text{eq}}^2}{2} \right\}, \frac{1}{4}(\tilde{k}_F + \tilde{q})^2 \right] \\ f_2(\tilde{q}) &= (\tilde{k}_F + \tilde{q}) |\tilde{k}_F + \tilde{q}|^{\gamma_{\text{eq}}^2 - 1} {}_1F_2 \left[\left\{ \frac{\gamma_{\text{eq}}^2}{2} \right\}, \left\{ \frac{1 + \gamma_{\text{eq}}^2}{2}, \frac{2 + \gamma_{\text{eq}}^2}{2} \right\}, \frac{1}{4}(\tilde{k}_F + \tilde{q})^2 \right] . \end{aligned}$$

Here, ${}_pF_q$ are the generalized hypergeometric functions. From the Fourier transformed one body density matrix $\rho_1(q)$ we obtain the fermionic distribution function as $N\rho_1(q)$, which in the absence of interactions $\gamma_{\text{eq}} = 0$ reduces to a step function $\theta(|q| - k_F)$, and in presence of interactions decays like a power law (see Fig. 4.5).

Using that the 1-RDM is diagonal in Fourier space, we can directly compute the one-particle Rényi entanglement entropy

$$S_\alpha = \frac{1}{1 - \alpha} \ln \left(\frac{N}{2k_F} \int dq \rho_1(q)^\alpha \right) \quad (4.130)$$

$$S_1 = -\frac{N}{2k_F} \int dq \rho_1(q) \ln \rho_1(q) , \quad (4.131)$$

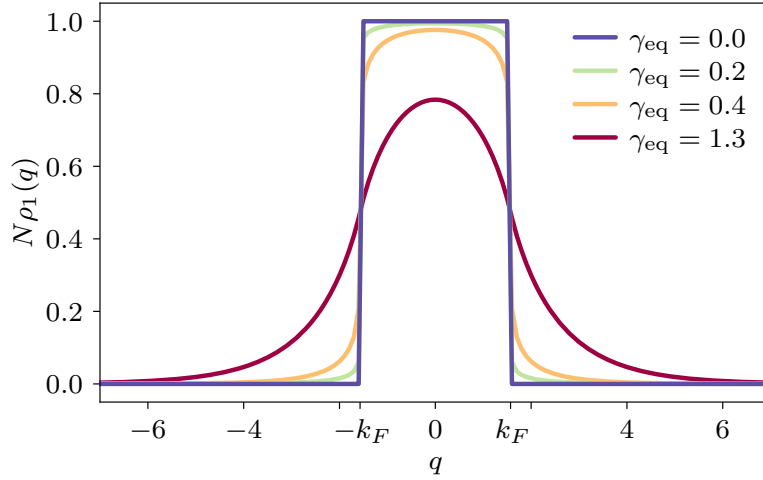


Figure 4.5: Distribution function $N\rho_1(q)$, Eq. (4.129), obtained from the $x/L \ll 1$ limit of the one body reduced density matrix in the Luttinger liquid model for a fixed interaction cutoff $\varepsilon = 0.84$ and various interaction strengths V/J . Without interactions, $\gamma_{\text{eq}} = 0$, the distribution function is a step function up at the Fermi momenta $k_F = \pm\pi/2$.

where $\alpha = 1$ is the von Neumann entropy, and the factor $N/(2k_F) = (2\pi/L)^{-1}$ originates from turning the sums into integrals in the limit of large L . When comparing to numerical results, we additionally subtract the entropy for free fermions S_{ff} .

In the absence of interactions, $\gamma_{\text{eq}} = 0$, the one-body density is given by $\varrho_1^0(x, 0)$ in Eq. (4.127). Performing the Fourier transform, we recover the expected zero temperature distribution function

$$N\varrho_1^0(q) = \theta(|q| - k_F) . \quad (4.132)$$

Using this expression in Eq. (4.131), one finds the free fermion von Neumann entropy [311, 480–485]

$$S_{\text{ff}} = -\frac{N}{2k_F} \int_{-k_F}^{k_F} dq \frac{1}{N} \ln \frac{1}{N} = \ln(N) . \quad (4.133)$$

The same one-particle entanglement entropy is obtained for any other Rényi power $\alpha > 1$ [483], which can be seen from Eq. (4.130)

$$\begin{aligned} S_{\alpha, \text{ff}} &= \frac{1}{1-\alpha} \ln \left(\frac{N}{2k_F} \int_{-k_F}^{k_F} dq \frac{1}{N^\alpha} \right) \\ &= \frac{1}{1-\alpha} \ln \frac{1}{N^{\alpha-1}} \\ &= \ln(N) = S_{\text{ff}} . \end{aligned} \quad (4.134)$$

4.3.2 Time evolution of the one-body density matrix after a quantum quench

We next consider free fermions for times $t < 0$ and suddenly turn on the V/J interaction at $t = 0$, so that the J - V Hamiltonian for this quench is given by

$$H = -J \sum_{i=1}^L (c_{i+1}^\dagger c_i + c_i^\dagger c_{i+1}) + V(t) \sum_{i=1}^L n_i n_{i+1} , \quad (4.135)$$

with $V(t) = \theta(t)V$. This allows us to study the growth and spread of entanglement entropy after the quench by considering the difference $S_\alpha - \ln(N)$, in which the entropy of free fermions is subtracted. We again start by computing the one body density

$$\rho_1(x, 0; t) = \frac{1}{N} \left[e^{-ik_F x} C_+(x, 0; t) + e^{ik_F x} C_-(x, 0; t) \right] \quad (4.136)$$

$$C_\alpha(x, 0; t) = \langle \psi_\alpha^\dagger(x, t) \psi_\alpha(0, t) \rangle \quad (4.137)$$

for the quench in the LL model

$$H = \sum_q [\omega_0(q) + m(q, t)] b_q^\dagger(t) b_q(t) + \frac{1}{2} \sum_q g_2(q, t) \left(b_q(t) b_{-q}(t) + b_q^\dagger(t) b_{-q}^\dagger(t) \right), \quad (4.138)$$

where in this case $g_2(q, t) = \theta(t)g_2(q) = \theta(t)g_2|q|$, $m(q, t) = \theta(t)g_4|q|$, and again, $\omega_0(q) = v_F|q|$. Analogous to the equilibrium case, we can diagonalize the Hamiltonian for any fixed time $t > 0$ using the same Bogoliubov transformation Eq. (4.103) with $\tanh(2\theta_q) = g_2(q)/(\omega_0(q) + g_4(q))$, but now the operators $a_q(t)$ are time dependent. For $t > 0$ this yields the diagonal Hamiltonian

$$H = \sum_q v|q| a_q^\dagger(t) a_q(t) \quad (4.139)$$

$$v = \sqrt{(v_F + g_4)^2 - g_2^2}. \quad (4.140)$$

Since the Hamiltonian for $t > 0$ is diagonal in the a_q operators, we can use the trivial time evolution

$$a_q(t) = e^{-iv|q|t} a_q. \quad (4.141)$$

Substituting this time evolution into the inverse of the transformation Eq. (4.103), we obtain the time evolution of the b_q operators as [163]

$$b_q(t) = w_q(t) b_q + u_q(t) b_{-q}^\dagger \quad (4.142)$$

$$w_q(t) = \cos(v|q|t) - i \sin(v|q|t) \cosh(2\theta_q) \quad (4.143)$$

$$u_q(t) = -i \sin(v|q|t) \sinh(2\theta_q).$$

A very important conceptual difference to the equilibrium case is that the Hamiltonian is not diagonal in the a_q operators for $t \rightarrow 0^-$, and therefore we cannot easily write down expectation values of the a_q operators. However, since H is diagonal in the b_q operators for $t \rightarrow 0^-$, we can use $b_q(t=0) \equiv b_q$ and

$$\begin{aligned} \langle b_q^\dagger b_{q'} \rangle &= f_b(q) \delta_{q,q'} \\ \langle b_q b_{q'} \rangle &= 0 = \langle b_q^\dagger b_{q'}^\dagger \rangle. \end{aligned} \quad (4.144)$$

This together with the more complicated time evolution of the b_q operators Eq. (4.142) gives rise to a different exponent $\gamma \geq \gamma_{\text{eq}}$ as we show in the following. Up to Eq. (4.114) the calculation for the correlation function is analogous to the equilibrium case such that

$$C_\alpha(x, t) = \frac{e^{\alpha \frac{i\pi x}{L}}}{2\pi\eta} e^{\frac{1}{2}[\phi_\alpha(x, t), \phi_\alpha(0, t)]} e^{-\frac{1}{2}\langle (\phi_\alpha(x, t) - \phi_\alpha(0, t))^2 \rangle}. \quad (4.145)$$

The exponential $e^{\frac{1}{2}[\phi_\alpha(x,t), \phi_\alpha(0,t)]}$ is unchanged by the time dependence and is still given by Eq. (4.114). In order to evaluate $\langle \phi_\alpha(x, t) \phi_\alpha(0, t) \rangle$, we use the time evolution of the b_q operators from Eq. (4.142) in the definition of the bosonic fields

$$\begin{aligned} \phi(x, t) = & - \sum_{q>0} \sqrt{\frac{2\pi}{qL}} e^{-\frac{q\eta}{2}} [e^{i\alpha qx} (w_q(t) b_{\alpha, q} + u_q(t) b_{\alpha, -q}^\dagger) \\ & + e^{-i\alpha qx} (w_q^*(t) b_{\alpha, q}^\dagger + u_q^*(t) b_{\alpha, -q})] . \end{aligned} \quad (4.146)$$

Analogous to the equilibrium case, we use $\langle b_q^\dagger b_q \rangle = f_b(q)$, to obtain

$$\begin{aligned} \langle \phi(x, t) \phi(x', t) \rangle = & - \sum_{q>0} \frac{2\pi}{qL} e^{-q\eta} \left\{ e^{i\alpha q(x-x')} [(1 - f_b(q)) |w_q(t)|^2 + f_b(q) |u_q(t)|^2] \right. \\ & \left. + e^{-i\alpha q(x-x')} [(1 - f_b(q)) |u_q(t)|^2 + f_b(q) |w_q(t)|^2] \right\} . \end{aligned} \quad (4.147)$$

We again consider the zero temperature case with $f_b(q > 0) = 0$. This allows us to rewrite the desired exponential term from Eq. (4.145) as follows

$$-\frac{1}{2} \langle (\phi_\alpha(x, t) - \phi_\alpha(0, t))^2 \rangle = \sum_{q>0} \frac{2\pi}{qL} e^{-q\eta} (|w_q(t)|^2 + |u_q(t)|^2) \left[-1 + \frac{1}{2} e^{i\alpha qx} + \frac{1}{2} e^{-i\alpha qx} \right] . \quad (4.148)$$

Using $|w_q(t)|^2 + |u_q(t)|^2 = \cosh^2(2\theta_q) - \cos(2v|q|t) \sinh^2(2\theta_q)$, the above becomes

$$\begin{aligned} -\frac{1}{2} \langle (\phi_\alpha(x, t) - \phi_\alpha(0, t))^2 \rangle = & \sum_{q>0} \frac{2\pi}{qL} e^{-q\eta} (2 \sin^2(v|q|t) \sinh^2(2\theta_q) + 1) \\ & \times \left[-1 + \frac{1}{2} e^{i\alpha qx} + \frac{1}{2} e^{-i\alpha qx} \right] . \end{aligned} \quad (4.149)$$

We define the momentum dependence of interaction parameter in the quench case as

$$\sinh^2(2\theta_q) \approx \left(\frac{K - K^{-1}}{2} \right)^2 e^{-\varepsilon|q|} \equiv \gamma^2 e^{-\varepsilon|q|} . \quad (4.150)$$

The free term $\langle (\phi_\alpha(x, t) - \phi_\alpha(0, t))^2 \rangle_0$ is equivalent to that in Eq. (4.121). To compute the interaction term, we need to compute the q -sum, where we can use Eq. (4.111) such that

$$\exp \left[\kappa \sum_{q>0} \frac{2\pi}{L} e^{-iqz(x)} \sin^2(v|q|t) \right] = \frac{[\sin(\frac{\pi}{L}(z(x) - 2v|q|t)) \sin(\frac{\pi}{L}(z + 2v|q|t))]^{\kappa/4}}{\sin(\frac{\pi}{L}z(x))^{\kappa/2}} . \quad (4.151)$$

The interaction term is then found to be

$$\begin{aligned} & e^{-\frac{1}{2}[\langle (\phi_\alpha(x, t) - \phi_\alpha(0, t))^2 \rangle - \langle (\phi_\alpha(x, t) - \phi_\alpha(0, t))^2 \rangle_0]} \\ & = \left| \frac{\sin(\frac{\pi}{L}(\mathbf{i}(\eta + \epsilon)))}{\sin(\frac{\pi}{L}(\alpha x + \mathbf{i}(\eta + \epsilon)))} \right|^{\gamma^2} \\ & \times \left| \frac{\sin(\frac{\pi}{L}(\alpha x + 2vt + \mathbf{i}(\eta + \epsilon)))}{\sin(\frac{\pi}{L}(2vt + \mathbf{i}(\eta + \epsilon)))} \right|^{\gamma^2/2} \\ & \times \left| \frac{\sin(\frac{\pi}{L}(\alpha x - 2vt + \mathbf{i}(\eta + \epsilon)))}{\sin(\frac{\pi}{L}(-2vt + \mathbf{i}(\eta + \epsilon)))} \right|^{\gamma^2/2} . \end{aligned} \quad (4.153)$$

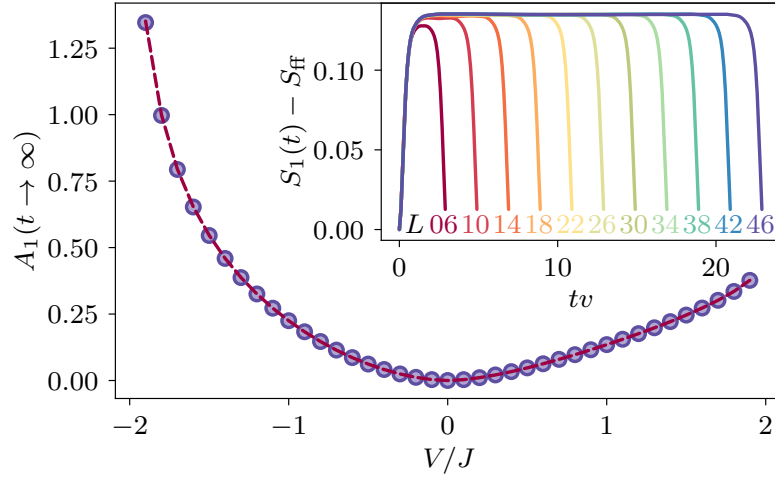


Figure 4.6: Thermodynamic limit growth of the 1-particle von Neumann entanglement entropy as a function of the interaction strength V/J after the quantum quench at $t = 0$ obtained from the Luttinger liquid steady state limit (red, dashed line) in Eq. (4.157). Blue circles show the result of finite size scaling of the plateau values shown in the inset. The inset depicts the first plateau of the entanglement entropy obtained by numerical integration from the time dependent one body density matrix in Eq. (4.154) for several system sizes $L = 2N$. We observe that the plateau size increases linearly with L while the region where the entropy increases to the plateau and the timescale where it drops from the plateau is independent of the system length. In the thermodynamic limit, the average of the entropy and the plateau values agree with each other. The main panel demonstrates that the plateau value of the entropy coincides with the entropy obtained from the steady state result for the 1-RDM in the thermodynamic limit.

Inserting the above and Eq. (4.113) into Eq. (4.145), and taking the limit $\eta/L \rightarrow 0$, yields the correlation function for α -movers. Adding together the right and left movers as in Eq. (4.126) gives the final expression [163]

$$\rho(x, t) = \rho_1^0(x, 0) \left| \frac{\sin(\pi \mathbf{i} \varepsilon / L)}{\sin(\pi(x + \mathbf{i} \varepsilon) / L)} \right|^{\gamma^2} \left| \frac{\sin(\frac{\pi}{L}(x - 2vt + \mathbf{i} \varepsilon)) \sin(\frac{\pi}{L}(x + 2vt + \mathbf{i} \varepsilon))}{\sin(\frac{\pi}{L}(-2vt + \mathbf{i} \varepsilon)) \sin(\frac{\pi}{L}(2vt + \mathbf{i} \varepsilon))} \right|^{\gamma^2/2}. \quad (4.154)$$

4.3.3 Steady state after quench

The 1-RDM consists of the free part $\rho_1^0(x, 0)$ Eq. (4.127), the interaction factor with exponent $\gamma^2 \neq \gamma_{\text{eq}}^2$, and a time dependent oscillatory term with exponent $\gamma^2/2$. To obtain the one-particle entanglement entropy with Eq. (4.130), we need to numerically compute the Fourier transform of Eq. (4.154). However, we can already extract information about the time dependence from the real space correlation function. We find that the entropy obtained from the LL correlation function is strictly periodic with period $\Delta t = L/(2v)$ and plateaus centered around $t_{n, \text{plateau}} = L/(2v)(n + 1/2), n \in \mathbb{N}_0$ (see inset of Fig. 4.6) corresponding to times when all sine functions turn into cosine functions in Eq. (4.154). Because the size of the plateaus is proportional to L and the timescale for increase and decrease from the plateaus is independent of L (inset Fig. 4.6), the average converges to the plateau value in the thermodynamic limit. We compute the plateau

values for many system lengths by numerically Fourier transforming Eq. (4.154), evaluating the entanglement entropy at $t_{0,\text{plateau}}$, and performing finite size scaling to show the thermodynamic limit averaged entropies with blue circles in Fig. 4.6.

The steady state estimate of the entropy can also be analyzed by generalizing the scaling form

$$S_\alpha(N, V/J) = \ln(N) + A_\alpha(V/J) + \mathcal{O}(N^{-\lambda}) , \quad (4.155)$$

to include the post-quench waiting time:

$$A_\alpha(V/J, t) = \lim_{N \rightarrow \infty} S_\alpha(N, V/J, t) - \ln(N) . \quad (4.156)$$

Its steady state value can be obtained from the 1-RDM in the $x/L \ll 1$ limit (dashed, red line in Fig. 4.6) obtained from Eq. (4.154)

$$\rho_{t \rightarrow \infty}(x) = \frac{\sin(k_F x)}{N\pi x} \left(\frac{\varepsilon^2}{x^2 + \varepsilon^2} \right)^{\gamma^2/2} + \mathcal{O}\left(\frac{x}{L}\right) , \quad (4.157)$$

similar to the equilibrium case Eq. (4.126) and Eq. (4.128) with γ_{eq} replaced by γ .

4.4 Numerical computations with exact diagonalization

To study how well the low energy field theory approach describes the J - V model in Eq. (4.13) in the LL phase and to fix the interaction cutoff, we perform a series of numerical calculations on finite sized systems. The software needed to reproduce all results is open source and has been made available online [474].

We first utilize exact diagonalization (ED), where we construct all $\binom{2N}{N}$ basis states for a lattice with $L = 2N$ sites and N fermions to determine the corresponding matrix elements of Eq. (4.13) and construct the Hamiltonian as a sparse matrix. We then use the Lanczos algorithm [486] to determine the ground state $|\Psi_0\rangle$, from which the full density matrix can be determined as $\rho = |\Psi_0\rangle\langle\Psi_0|$. The reduced one-body density matrix is obtained by fixing one coordinate in the anti-symmetrized many particle wave function $\Psi_0(i_1, \dots, i_N) = \langle i_1, \dots, i_N | \Psi_0 \rangle$ and tracing out the other $N - 1$ particle positions [161]

$$\rho_1^{i_1, j_1} = \sum_{i_2, \dots, i_N} \Psi_0^*(i_1, i_2, \dots, i_N) \Psi_0(j_1, i_2, \dots, i_N) . \quad (4.158)$$

From the reduced density matrix, one can then obtain the von Neumann and Rényi entropies according to Eq. (4.1), $S_1 = -\text{Tr}[\rho_A \ln \rho_A]$, and Eq. (4.2), $S_\alpha = \frac{1}{1-\alpha} \ln \text{Tr}[\rho_A^\alpha]$, respectively. For the time evolution after the quench, we first obtain the ground state in the pre-quench case $t < 0$ for free fermions $|\Psi_0(0)\rangle$ and then evolve the state using unitary time evolution

$$|\Psi(t)\rangle = e^{-itH} |\Psi_0(0)\rangle = \sum_{\alpha} e^{-iE_\alpha t} \langle \Psi_\alpha | \Psi_0(0) \rangle |\Psi_\alpha\rangle , \quad (4.159)$$

where $|\Psi_\alpha\rangle$ and E_α are the full set of eigenstates and eigenvalues of the post-quench Hamiltonian, respectively. From the density matrix $|\Psi(t)\rangle\langle\Psi(t)|$, we can then first obtain the reduced density matrix and then the entanglement entropy as before.

The advantage of ED is that by constructing a full basis of the Hilbert space, we obtain exact results without any approximations and can thus also use ED to benchmark other approximate

methods. Here, we would like to consider systems as large as possible; however, unfortunately the number of basis states is proportional to $\binom{L}{N}$ and thus grows incredibly fast with the size of the system at half-filling $L = 2N$. A naive implementation as described so far would only allow us to obtain one particle entanglement entropies for the ground state of systems up to $N = 12$ fermions on $L = 24$ lattice sites within any reasonable time (on the order of weeks) and 1 TB of system memory. However, we can use a series of tricks to reduce the computational resource requirements that ultimately allows us to consider the one-particle entanglement entropy for the ground state of up to $N = 19$ fermions on $L = 38$ lattice sites and even the time evolution after an interaction quench for up to $L = 26$ sites using ED. In the following, we describe the most important technical aspects: On a periodic lattice at half-filling, we (i) use three symmetries to reduce the size of the matrix from which we obtain the ground state by about a factor $1/(4L)$, (ii) efficiently implement the basis as an integer fermion basis, and (iii) for obtaining the spectrum of the 1-RDM, construct the Schmidt decomposition to find the singular values of the 1-RDM.

4.4.1 Symmetries

Translational symmetry

The most important symmetry is the translational symmetry T which commutes with the Hamiltonian of the J - V model Eq. (4.13), i.e. $[T, H] = 0$, due to the (anti-) periodic boundary conditions. Here, the unitary operator T moves each fermion one site to the right, e.g. $T|01100101\rangle = |10110010\rangle$. For any occupation number basis state $|\phi_j\rangle$ at half-filling, the resulting state $T|\phi_j\rangle$ is another state that is also in the set of basis states. Applying T multiple times to $|\phi_j\rangle$, the states $T^k|\phi_j\rangle$ for $k = 0, 1, \dots, M_\nu - 1$ with $2 \leq M_\nu \leq L$ define a symmetry cycle ν of M_ν elements. This allows us to group the whole basis into translation symmetry cycles $|\varphi_{\nu,m}\rangle$ ($m = 1, \dots, M_\nu$), where we denote the state $|\varphi_{\nu,1}\rangle$ as the leader of cycle ν , and $|\varphi_{\nu,m+1}\rangle = T|\varphi_{\nu,m}\rangle$ with $|\varphi_{\nu,M_\nu+1}\rangle = T|\varphi_{\nu,1}\rangle$. Due to the construction, the cycles cannot be mixed among each other by application of T , and for each cycle ν of M_ν basis states, we can define a number of M_ν new orthonormal basis states as linear combinations

$$|\gamma_{\nu,q}\rangle = \frac{1}{\sqrt{M_\nu}} \sum_{m=1}^{M_\nu} e^{i2\pi q(m-1)/M_\nu} T^{m-1} |\varphi_{\nu,1}\rangle \text{ with } q = 1, 2, \dots, M_\nu. \quad (4.160)$$

These states are eigenstates of T with $T|\gamma_{\nu,q}\rangle = e^{-i2\pi q/M_\nu} |\gamma_{\nu,q}\rangle$. As H and T commute, the Hamiltonian cannot mix between $|\gamma_{\nu,q}\rangle$ with different q such that it becomes block diagonal in the representation of the basis $\{|\gamma_{\nu,q}\rangle\}$ sorted by the value of q . The main advantage comes from the fact that the ground state lies in the $q = 0$ block [487] and that the eigenstates of blocks with different q are orthogonal to each other. We therefore only need to construct and store this single block of the Hamiltonian with $q = 0$. The ground state is then obtained from the $q = 0$ block stored in a sparse matrix format using the Lanczos algorithm, and for the time evolution after the quench, it is sufficient to find all eigenstates $|\Psi_{\alpha,q=0}\rangle$ and eigenvalues $E_{\alpha,q=0}$ of the dense $q = 0$ block of the Hamiltonian (as $\langle\Psi_0(0)|\Psi_{\alpha,q>0}\rangle = 0$). This therefore reduces the size of the problem by about a factor of $1/L$, as L is the typical size of a cycle.

Particle-hole symmetry

At half filling, the particle-hole operator P , which flips all occupation numbers $P|101001\rangle = |010110\rangle$ is another symmetry of the Hamiltonian which also commutes with the translation operator, $[T, P] = 0$. Because $P^2 = 1$, the particle-hole operator has eigenvalues $n_P = \pm 1$. If

ket	binary representation of Int64	integer
$ 1010\rangle =$	<div style="display: inline-block; border: 1px solid black; padding: 2px 5px; margin: 0 2px;"> ⁶⁴0 </div> <div style="display: inline-block; border: 1px solid black; padding: 2px 5px; margin: 0 2px;"> ⁶³0 </div> <div style="display: inline-block; border: 1px solid black; padding: 2px 5px; margin: 0 2px;"> ⁶²0 </div> <div style="display: inline-block; border: 1px solid black; padding: 2px 5px; margin: 0 2px;"> ⁶¹0 </div> <div style="display: inline-block; border: 1px solid black; padding: 2px 5px; margin: 0 2px;"> ⁶⁰0 </div> <div style="display: inline-block; border: 1px solid black; padding: 2px 5px; margin: 0 2px;"> ⁵⁹0 </div> ... <div style="display: inline-block; border: 1px solid black; padding: 2px 5px; margin: 0 2px;"> ⁶0 </div> <div style="display: inline-block; border: 1px solid black; padding: 2px 5px; margin: 0 2px;"> ⁵0 </div> <div style="display: inline-block; border: 1px solid black; padding: 2px 5px; margin: 0 2px;"> ⁴1 </div> <div style="display: inline-block; border: 1px solid black; padding: 2px 5px; margin: 0 2px;"> ³0 </div> <div style="display: inline-block; border: 1px solid black; padding: 2px 5px; margin: 0 2px;"> ²1 </div> <div style="display: inline-block; border: 1px solid black; padding: 2px 5px; margin: 0 2px;"> ¹0 </div>	= 10

Figure 4.7: Sketch of the occupation number representation of basis states as integers. The state $b = 10$ ($|1010\rangle$) is an example for a basis state for $N = 2$ fermions on $L = 4$ sites. We treat the occupation numbers for each site as the bits of the binary representation of an integer. This allows effective operations on these integers – much faster than acting on arrays of occupation numbers.

$P|\gamma_{\tilde{\nu},q}\rangle$ lies in a different cycle than $|\gamma_{\tilde{\nu},q}\rangle$, we can use P to further subdivide the $q = 0$ block by using the projection $(1 \pm P)/\sqrt{2}$ onto its eigenstates

$$|\theta_{\tilde{\nu},q,n_P=\pm 1}\rangle = \frac{1}{\sqrt{2}} (|\gamma_{\tilde{\nu},q}\rangle \pm P|\gamma_{\tilde{\nu},q}\rangle) . \quad (4.161)$$

Reflection symmetry

The third symmetry we exploit is spatial inversion R , which reflects the occupation numbers about a site, e.g. $R|011011\rangle = |110110\rangle$, and commutes with the Hamiltonian $[R, H] = 0$. However, in general, R does not commute with T , but fortunately in the $q = 0$ block translation and spatial inversion do commute. Since $R^2 = 1$, the eigenvalues are also given by $n_R = \pm 1$ and the corresponding projection operator is given by $(1 \pm R)/\sqrt{2}$. If R maps either $|\gamma_{\nu,q}\rangle$ or $|\theta_{\tilde{\nu},q,n_P}\rangle$ into another cycle, projecting onto eigenstates of R further subdivides the $q = 0$ block of the Hamiltonian in analogy to Eq. (4.161).

We therefore only need to construct the $q = 0$, $n_R = +1$, $n_P = +1$ block of the Hamiltonian which is a major reduction in memory and time complexity for obtaining the ground state.

4.4.2 Integer fermion basis

A straightforward way of implementing a state in the occupation number basis, is to store the occupation numbers in an array \mathbf{n} of length L . However, to more efficiently store and operate on the basis states, we implement them as 64-bit integers such that the binary representation of the integer has the occupation numbers as bits, e.g. $|1010\rangle \rightarrow 1010_2 = 10$ (Fig. 4.7). In cases, where we need more than 64 lattice sites, we can even use 128-bit integers to allow for a total of up to 128 sites. One direct advantage of the integer basis is that rank ordering the integers provides a natural order for the basis states such that a state can be looked up in the set of basis states by searching in a sorted list which only scales logarithmically with the length of the list and is thus very efficient.

In the following, we discuss the efficient application of operators on the integer basis states for which we use several logical bit operations: (i) right `bitshift`, $b \gg i$, that shifts every bit of b by i places to the right, filling up the leftmost bits with zeros and losing information about the first i bits, e.g. $0\dots 0101 \gg 1 = 0\dots 0010$, (ii) left `bitshift`, $b \ll i$, shifting i places to the left, e.g. $0\dots 00101 \ll 1 = 0\dots 01010$, (iii) logical and operations, $\&$, e.g. $1010 \& 1001 = 1000$, (iv) local or operations, $|$, e.g. $1010 | 1001 = 1011$, and the not operation \sim , e.g. $\sim 1010 = 0101$.

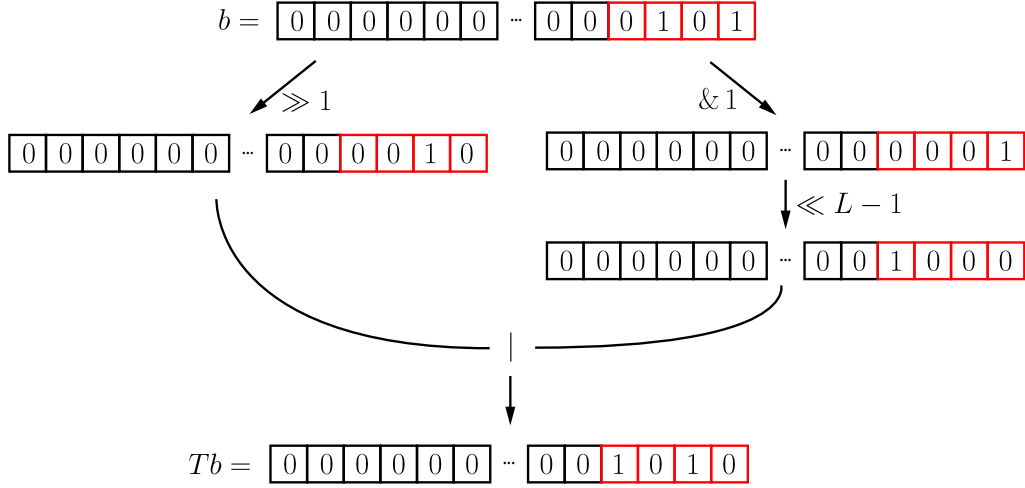


Figure 4.8: Example of translation by one site using bitshift operations. We show application of T on the state $b = 5$ ($|0101\rangle$) – an example of a state from the $N = 2$, $L = 4$ basis.

Creation and annihilation operators

We can represent an operator c_i^\dagger or c_i as an integer with a 1 at position i and zeros in all other bits, i.e. $c_i \rightarrow 1 \ll i$. However, this integer only stores the position of the site the operator acts on, not the fermion character. To act on a basis state $|\phi\rangle \rightarrow b$, we first need to check if site i is empty, that is $(1 \ll i) \& b$ is zero. If this is the case, action of c_i destroys the state, and if it is not the case, then c_i^\dagger destroys the state. For the case that this check is successfully passed, and the state is not destroyed, the result of applying the operator is obtained via

$$c_i^\dagger |\phi\rangle \rightarrow (-1)^{\sum_{j=1}^{i-1} b_j} (1 \ll i) | b \quad (4.162)$$

$$c_i |\phi\rangle \rightarrow (-1)^{\sum_{j=1}^{i-1} b_j} [(1 \ll i) | (\sim b)] . \quad (4.163)$$

Here, the sign $(-1)^{\sum_{j=1}^{i-1} b_j}$ accounts for the minus signs when moving the operator across the creation operators that build up the state, $|\phi\rangle = |n_1, n_2, \dots, n_L\rangle = \prod_{n_i=1}^i c_i^\dagger |0\rangle$, where we sort the creation operators according to their site index i , e.g. $c_2^\dagger |0101\rangle = c_2^\dagger (c_1^\dagger c_3^\dagger |0\rangle) = -c_1^\dagger c_2^\dagger c_3^\dagger |0\rangle = -|0111\rangle$.

This can be straightforwardly extended to action of more operators of the same kind which allows for efficient computations of matrix elements of the type $\langle \phi_i | c_{i_1}^\dagger \dots c_{i_n}^\dagger c_{j_1} \dots c_{j_n} | \phi_j \rangle$ with just two integers from integer bases with n occupied bits on L sites.

Translation operator

The translation operator shifts every fermion one site to the right which can be achieved by a bitshift, $b \gg 1$. However, due to the boundary condition, we need to keep information about the first bit, $b \& 1$, which needs to be moved to position L via a shift $\ll L - 1$. We can then combine the right shifted integer with the former first bit at position L by using an *or* operation. An example of this process is shown in Fig. 4.8, and the full operation is given by

$$T|\phi\rangle \rightarrow (b \gg 1) | ((b \& 1) \ll L - 1) . \quad (4.164)$$

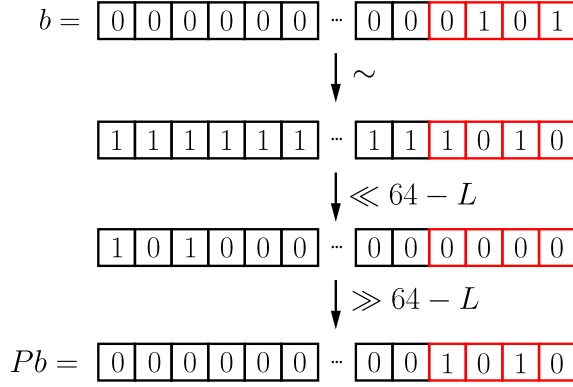


Figure 4.9: Example of particle-hole operation on an integer fermion basis state. We show application of P on the state $b = 5$ ($|0101\rangle$) – an example state from the $N = 2$, $L = 4$ basis.

An efficient implementation of the translation operator T even allows us to only store the translation cycle leaders of the basis and generate the other elements of the cycle on demand by action of T , thus saving about a factor $1/L$ of system memory for the basis.

Particle-hole operator

To act on a state with the particle-hole operator P , we need to flip all occupation numbers, which is achieved by a not operation, $\sim b$. However, this also flips the bits other than the first L ones. To remove these bits outside the lattice, we shift $64 - L$ bits to the left and again to the right, thus truncating the $64 - L$ last bits. An example is shown in Fig. 4.9, and the full operation is given by

$$P|\phi\rangle \rightarrow [(\sim b) \ll (64 - L)] \gg (64 - L) . \quad (4.165)$$

Reflection operator

The reflection operator can be implemented using a `bitreverse` that inverts the order of all 64 bits. We then only need to move the last L bits back to the first L sites by a bitshift of $64 - L$ sites. An example is shown in Fig. 4.10 and the full operation is given by

$$R|\phi\rangle \rightarrow \text{bitreverse}(b) \gg (64 - L) . \quad (4.166)$$

Computational resource requirements

The full basis for N fermions on $L = 2N$ lattice sites contains $\binom{2N}{N}$ elements such that the memory needed to store the integer basis is given by $\binom{2N}{N} \frac{8}{(1024)^3}$ GB. For example for $N = 19$, this is about 263 GB for storing the basis alone, and for $N = 20$ fermions, the size already exceeds 1 TB. Using the integer fermion basis and only saving the cycle leaders allows us to push the full exact diagonalization computations of the one-particle entanglement entropy to a maximum of $N = 19$ fermions for 1.5 TB of system memory (which also needs to accommodate the $q = 0$ block of the Hamiltonian, the ground state, and the Schmidt decomposition).

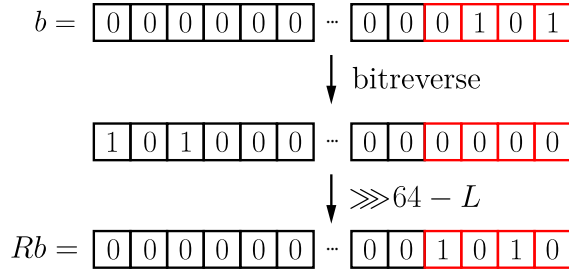


Figure 4.10: Example of reflection operation on an integer fermion basis state. We show application of R on the state $b = 5$ ($|0101\rangle$) – an example state from the $N = 2$, $L = 4$ basis.

4.4.3 Structure matrix

Parts of this subsection closely follow the publication: Harini Radhakrishnan, Matthias Thamm, Hatem Barghathi, Bernd Rosenow, and Adrian Del Maestro, *A Scaling Function for the Particle Entanglement Entropy of Fermions*, preprint arXiv:2302.09093 [488].

For efficiently computing the n -particle entanglement entropy, we construct the coefficient matrix C used for the Schmidt decomposition, i.e. we write a state $|\Psi\rangle$ in terms of the basis states $|\theta_a\rangle_A$ and $|\chi_b\rangle_B$ in the partitions A and B according to

$$|\Psi\rangle = \sum_{ab} C_{ab} |\theta_a\rangle_A \otimes |\chi_b\rangle_B . \quad (4.167)$$

The first trick is to constrict the sub-bases states in terms of translational cycles μ (ν), i.e. $|\theta_{\mu,i}\rangle$ ($|\chi_{\nu,j}\rangle$), where $i = 1$ corresponds to the cycle leader, and $T|\theta_{\mu,i}\rangle_A = |\theta_{\mu,i+1}\rangle_A$. Due to the translational symmetry, it is sufficient to only consider the $|\theta_{\mu,i}\rangle_A$ and the cycle leaders in the B partition, $|\chi_{\nu,1}\rangle_B$, as the other contributions are redundant by application of T . The challenge is then to decompose the full basis states into the Kronecker product of the sub-bases with the correct signs and to extract the coefficients from the state $|\Psi\rangle$. For this, we precompute the so-called structure matrix $\mathcal{A} \in \mathcal{M}_{n_{c,A} \times n_{c,B} \times L}$, where $n_{c,A}$ ($n_{c,B}$) is the number of symmetry cycles in partition A (B). We define the magnitude $|\mathcal{A}_{\mu\nu i}|$ as the index of the symmetry cycle of the state $|\theta_{\mu,i}\rangle_A \otimes |\chi_{\nu,1}\rangle_B$ in the full basis $|\gamma_{|\mathcal{A}_{\mu,\nu,i}|,k}\rangle$, which is also the index for the corresponding coefficient in $|\Psi\rangle$. The sign of $\mathcal{A}_{\mu\nu i}$ is defined as the sign of the state $|\theta_{\mu,i}\rangle_A \otimes |\chi_{\nu,1}\rangle_B$ in the anti-symmetrized first quantization basis. In the integer fermion basis language, the integers representing the state of the full basis is obtained from the integers b_μ^A , b_ν^B of the sub-bases via

$$|\theta_{\mu,i}\rangle_A \otimes |\chi_{\nu,1}\rangle_B \rightarrow T^{i-1} b_\nu^A | b_\nu^B , \quad (4.168)$$

and the sign has two contributions:

- (i) $T^{i-1} b_\nu^A$ may contribute a minus sign due to the boundary conditions if n is even and a fermion is moved across the boundary by T^{i-1} (as the corresponding creation operator has to be moved to the first position through the $n - 1$ other creation operators).
- (ii) When combining $T^{i-1} b_\nu^A$ with b_ν^B , we order the creation operators to obtain the state in the integer basis and the corresponding contribution to the sign, e.g.

$$|0001\rangle_A \otimes |1000\rangle_B \rightarrow c_1^\dagger c_4^\dagger |0\rangle \longrightarrow |1_2 001_1\rangle \quad (4.169)$$

$$|1000\rangle_A \otimes |0001\rangle_B \rightarrow c_4^\dagger c_1^\dagger |0\rangle = -c_1^\dagger c_4^\dagger |0\rangle \longrightarrow -|1_1 001_2\rangle . \quad (4.170)$$

Here, the subscript labels the particles. Writing the states in the fermion position basis $|\theta_{\mu,i}\rangle_A \rightarrow |i_1, \dots, i_n\rangle$ and $|\chi_{\mu,1}\rangle_B \rightarrow |j_1, \dots, j_n\rangle$ (for example $|0110\rangle \rightarrow |2, 3\rangle$), the sign is negative (positive) if $(i_1, \dots, i_n, j_1, \dots, j_n)$ is an odd (even) permutation of $(1, 2, \dots, 2n)$. The permutation is odd if it needs an odd number of nearest neighbor permutations to sort the list $(i_1, \dots, i_n, j_1, \dots, j_n)$.

With this information at hand, we can directly construct the coefficient matrix C by combining the coefficients from the state $|\Psi\rangle$ state – according to the magnitude of the entry in the structure matrix – with the corresponding phase – according to the sign of the entry in the structure matrix. In addition, exploiting translational symmetry, the matrix C can be diagonalized using a Fourier transform to efficiently obtaining the spectrum $|\lambda_i|^2$ and thus the particle entanglement entropy.

4.5 Numerical computations with DMRG and ITensors.jl

As discussed in the previous section, the major disadvantage of exact diagonalization is the exponential size of the Hilbert space that severely limits the size of systems that can be considered – even with 1.5 TB of memory. However, for accurate finite size scaling to the thermodynamic limit, it is crucial to consider systems as large as possible. For this reason, we describe approximate methods in this section which allow for the truncation of the Hilbert space to a much smaller subspace of relevant states.

In order to study systems with $L \geq 40$ and thus improve finite size scaling to the thermodynamic limit, we use the implementation of the density matrix renormalization group algorithm (DMRG) in the `ITensors.jl` software package [475] for the Julia programming language. As an approximate method, DMRG does not need to explore the entire Hilbert space and therefore requires fewer resources, but at the price of inaccuracies with magnitudes that are difficult to estimate a priori. We therefore also use ED as a benchmark to estimate the reliability of DMRG results, where a direct naive DMRG application to the J - V model with periodic boundary conditions leads to significant errors already for systems of size $N > 17$. We thus use a number of checks and detailed knowledge of the physical system to stabilize the DMRG calculation as we describe in the following.

After we computed the ground state $|\Psi_0\rangle$ using DMRG, we obtain the reduced one-body density matrix from the correlation matrix [145]

$$\rho_1^{i_1, j_1} = \frac{1}{N} \langle \Psi_0 | c_{i_1}^\dagger c_{j_1} | \Psi_0 \rangle . \quad (4.171)$$

From the eigenvalues ϵ_i of ρ_1 , we can then directly obtain the one-particle entanglement entropy

$$S_1 = - \sum_i \epsilon_i \ln \epsilon_i \quad (4.172)$$

$$S_\alpha = \frac{1}{1-\alpha} \ln \left[\sum_i \epsilon_i^\alpha \right] . \quad (4.173)$$

In order to obtain the waiting time dependence of the particle entanglement entropy after an interaction quench, we first use DMRG to find the ground state $|\Psi_0(0)\rangle$ for times $t \leq 0$ for free fermions, from which we compute the state $|\Psi(t)\rangle$ at time t after the quench using unitary time evolution

$$|\Psi(t)\rangle = e^{-iHt} |\Psi_0(0)\rangle . \quad (4.174)$$

In contrast to ED, we do not need to compute the full eigensystem of H to perform this evolution, but rather can construct an approximation of the time evolution operator for a short time step δt and apply it many times to estimate the time evolved state (see Sec. 4.5.2).

4.5.1 Advantages of tensor networks and stabilizing DMRG

For implementing the approximative numerics, we use the open source software package `ITensors.jl` [475] for the Julia programming language. `ITensors.jl` provides an implementation of tensor networks [489] and has several advantages for our purposes here:

- i) The package makes it easy to set up a lattice of L sites for fermions and automatically handles particle number conservation and commutation relations of operators.
- ii) The states are efficiently stored as matrix product states (MPSs) with efficient implementations of typical operations like inner products. Importantly, these states can be truncated to only keep a relevant subspace of the Hilbert space, which saves significant amounts of memory.
- iii) Operators can be encoded as so-called matrix product operators (MPOs) that can very efficiently be applied to MPSs where `ITensors` automatically handles the bookkeeping for indices. In addition, there are efficient algorithms for computing matrix elements and even for the correlation matrix $\langle \Psi | c_i^\dagger c_j | \Psi \rangle$ is already implemented. `ITensors` follows an easy-to-read notation, e.g. to construct an MPO for the Hamiltonian, and add a term $-tc_j^\dagger c_{j+1}$, one can write

```
H = AutoMPO()
H += -t, "Cdag", j, "C", j+1
```

and `ITensors` configures the MPO and handles fermionic commutation relations automatically.

- iv) `ITensors` comes with an implementation of the density matrix renormalization group algorithm (DMRG) [490–492]. DMRG is a well established, variational algorithm for approximating the ground state of a Hamiltonian that avoids considering the full Hilbert space by iterating over smaller subblocks of the system, iteratively improving the ground state estimate while also truncating the MPS to the relevant directions [492]. A detailed explanation of the DMRG algorithm and its implementation with MPS can be found in Ref. [492]. Besides handing an initial state to the DMRG algorithm as a starting point, the implementation in `ITensors` also allows us to define an orthonormal subspace to the desired ground state where orthogonality is enforced in each iteration.
- v) With `ITensorGPU.jl` there is also an additional package based on `ITensors` that allows certain operations to be easily performed using massively parallel GPUs.

For example, to implement a lattice of L sites, the J - V Hamiltonian, and compute the ground state using DMRG with `ITensors` one can simply write

```

using ITensors

# setup parameters
L = 20
N = 10
V = 1.0

# setup lattice, fermionic space, conserve particle numbers
sites = siteinds("Fermion", L; conserve_qns=true)

# any initial state with correct number of particles
state = Vector{String}([Bool(i <= N) ? "Occ" : "Emp" for i in 1:L])
shuffle!(state)
psi0 = MPS(sites, state)

# Hamiltonian as MPO
ampo = OpSum()
for j=1:L-1
    ampo += -t, "Cdag", j, "C", j+1
    ampo += -t, "Cdag", j+1, "C", j
    ampo += V, "N", j+1, "N", j
end
factor = (L/2 % 2 == 0) ? -1 : 1
ampo += -t*factor, "Cdag", L, "C", 1
ampo += -t*factor, "Cdag", 1, "C", L
ampo += V, "N", 1, "N", L
H = MPO(ampo, sites)

# DMRG parameters
sweeps = Sweeps(5)
maxdim!(sweeps, 10, 20, 100, 100, 200)
cutoff!(sweeps, 1e-10)
noise!(sweeps, 1e-5, 1e-8)

# get ground state from DMRG
energy, psi = dmrg(H, psi, sweeps; outputlevel=0)

```

However, there are problems that come with a naive implementation like this: (i) The `ITensors` networks are less efficient for (anti-) periodic boundary conditions, as they prefer to perform operations on lattice sites that are close together – an operation between the last and first sites is therefore very inefficient [475], and (ii) it can be hard to make DMRG converge to the true ground state, especially for large lattices without open boundary conditions. Here, already for systems with $N > 17$ fermions, a naive application of DMRG starting from a random state results in significant errors.

We thus stabilize DMRG with two tricks, described in the following:

Initial state

It is crucial to construct very good initial states, so that the algorithm starts as close as possible to the ground state. For this purpose, we combine a state which is a superposition of random states of the correct particle number, with a V/J dependent fraction of the corresponding $|\Psi_{V/J \rightarrow \pm\infty}\rangle$ state (Eqs. (4.28) and (4.36)).

Orthogonal subspace

The most important step for stabilizing convergence of DMRG to the ground state is to construct an orthogonal subspace to the ground state and enforce orthogonality to a basis in this subspace

during each sweep of DMRG. This feature has already been implemented in `ITensors.jl` with the intent to obtain excited states. The consideration of symmetry cycles Eq. (4.160) already reveals good candidates for orthogonal subspaces, because states with different q are orthogonal to each other. Using all states from blocks $q > 0$ is overkill, slowing down the DMRG algorithm and requiring huge amounts of memory, eliminating the advantages of the approximative method. We therefore only consider a subspace in which DMRG is most likely to converge if it misses the ground state. For $V/J > 0$ using the two states with maximal particle separation $|\psi_{>,\nu}\rangle = T^\nu(\prod_{j=1}^N c_{2j}^\dagger)|0\rangle$ this is:

$$|\Psi_{\perp,>}\rangle = \frac{1}{\sqrt{2}} [|\psi_{>,0}\rangle - |\psi_{>,1}\rangle] , \quad (4.175)$$

and for negative interactions there are L states with full clustered fermions $|\psi_{<,\nu}\rangle = T^\nu(\prod_{j=1}^N c_j^\dagger)|0\rangle$. Their span is given by

$$|\Psi_{\perp,<,q}\rangle = \frac{1}{\sqrt{N}} \sum_{\nu=0}^{L-1} \cos\left(\frac{2\pi\nu q}{L}\right) |\psi_{<,\nu}\rangle \quad (4.176)$$

$$|\Psi_{\perp,<,q+N}\rangle = \frac{1}{\sqrt{N}} \sum_{\nu=0}^{L-1} \sin\left(\frac{2\pi\nu q}{L}\right) |\psi_{<,\nu}\rangle . \quad (4.177)$$

4.5.2 Time evolution using GPUs

To efficiently perform time evolution of the initial state obtained with DMRG, we approximate the time evolution operator $e^{-iH\delta t}$ for a time step δt by using a symmetrized second order Trotter decomposition [493, 494]

$$e^{-i\delta t H} \approx e^{-i\delta t h_{1,2}/2} e^{-i\delta t h_{2,3}/2} \dots e^{-i\delta t h_{L,1}/2} \times \\ \times e^{-i\delta t h_{L,1}/2} e^{-i\delta t h_{L-1,L}/2} \dots e^{-i\delta t h_{1,2}/2} + \mathcal{O}(\delta t^3) , \quad (4.178)$$

where $h_{i,i+1} = -J(c_{i+1}^\dagger c_i + c_i^\dagger c_{i+1}) + V n_i n_{i+1}$. To derive Eq. (4.178) the J - V Hamiltonian

$$H = \sum_{i=1}^L h_{i,i+1} = \sum_{i \text{ even}} h_{i,i+1} + \sum_{i \text{ odd}} h_{i,i+1} \equiv H_{\text{even}} + H_{\text{odd}} \quad (4.179)$$

is split into the two internally commuting parts H_{even} and H_{odd} . The commutator $[H_{\text{even}}, H_{\text{odd}}]$ is neglected, which introduces an error $\mathcal{O}(\delta t^3)$ in Eq. (4.178). To maintain accuracy, it is therefore necessary to choose a short time step δt such that performing time evolution for a finite time interval t can require a large number $t/\delta t$ of time-consuming applications of the operator.

Application of a gate to an MPS can be significantly speeded up using GPUs, while construction of the gates and moving MPS and gates onto the GPU memory are time-consuming operations. Fortunately, we can construct the Trotter gates with time step δt once, move them to the GPU, and successively apply them to the MPS that has to be moved to the GPU only once. The time evolution using Trotter gates is therefore an ideal example for speeding up numerics by several orders of magnitude using GPUs instead of CPUs. For the implementation – the code including all necessary workarounds is open source and available online [474] – we use the `ITensorGPU.jl` extension to `ITensors` for which we were able to fix two interconnected bugs [495], which allows us to compute the waiting time dependence of the one-particle entanglement entropy for lattices with up to $L = 30$ sites with $N = 15$ fermions. This would be completely out of reach for ED calculations.

4.5.3 Higher order particle entanglement entropies

This subsection closely follows the publication: Harini Radhakrishnan, Matthias Thamm, Hatem Barghathi, Bernd Rosenow, and Adrian Del Maestro, *A Scaling Function for the Particle Entanglement Entropy of Fermions*, preprint arXiv:2302.09093 [488].

More generally, we can also compute the n -particle entanglement entropy from the n -RDM by first obtaining the ground state from DMRG, potentially further computing the time evolution after a quench, to obtain the state $|\Psi\rangle$ for which the n -RDM is given by [145]

$$\rho_n^{(i_1, \dots, i_n), (j_1, \dots, j_n)} = \langle \Psi | c_{i_1}^\dagger \cdots c_{i_n}^\dagger c_{j_1} \cdots c_{j_n} | \Psi \rangle / \binom{N}{n} . \quad (4.180)$$

In general, this requires computing $\binom{L}{n} \times \binom{L}{n}$ matrix elements and constructing a large number of MPOs, which is slow. However, we can utilize a few tricks to make computing the n -RDM feasible:

- (i) We map the operators $c_{i_1}^\dagger \cdots c_{i_n}^\dagger$ and $c_{j_1} \cdots c_{j_n}$ with $i_1 < i_2 < \dots < i_n$ and $j_1 < j_2 < \dots < j_n$ onto an integer basis, e.g. $c_1^\dagger c_3^\dagger c_4^\dagger \rightarrow 001101 = 13$. The advantages of the integer basis are that we can efficiently construct it and order the elements according to the integers for which locating an element in the ordered set of basis states scales only logarithmically with the number of elements. We can still translate back to the i_1, \dots, i_n indices to construct the operators in `ITensors` based on the site indices.
- (ii) We use the anti-commutation relations between the fermion operators to get the matrix elements for permutations of the $c_{i_1}^\dagger \cdots c_{i_n}^\dagger$ and $c_{j_1} \cdots c_{j_n}$. The parity of the permutation is then the sign of the corresponding matrix element of ρ_n , and the magnitude is the same for the matrix elements for all the permutations. It is even sufficient to only construct the *upper triangle* where $i_1 < i_2 < \dots < i_n$, $j_1 < j_2 < \dots < j_n$ and to normalize the singular values such that $\sum_n |\lambda_n|^2 = 1$.
- (iii) We use translational symmetry of $|\Psi\rangle$, i.e.

$$\rho_n^{(i_1, \dots, i_n), (j_1, \dots, j_n)} = (-1)^{\delta_{i_n, L} + \delta_{j_n, L}} \rho_n^{(i_1+1, \dots, i_n+1), (j_1+1, \dots, j_n+1)} \quad (4.181)$$

to further reduce the number of matrix elements we need to compute. Here, it is advantageous that we can both efficiently apply the translation operator to the fermion basis states (Sec. 4.4.2) and locate the resulting state in the basis with logarithmic time complexity.

These tricks allow us to compute the two-particle entanglement entropy for systems up to $L = 80$ lattice sites, the three-particle entanglement entropy for up to $L = 48$, and even the four-particle entanglement entropy for up to $L = 32$ sites.

4.5.4 Benchmarking DMRG with exact diagonalization

As DMRG is approximate, and the accuracy can be hard to estimate and control [492, 496], it is a priori not clear that the truncation error is small and that the true ground state is found. We therefore use ED as a benchmark to validate DMRG results and to estimate the size of the deviations. For this, we compute entropies with both DMRG and ED for some large systems that are still feasible with ED. We find excellent agreement when using the methods described above to stabilize DMRG. For example, Fig. 4.11 shows ED and DMRG results for the waiting

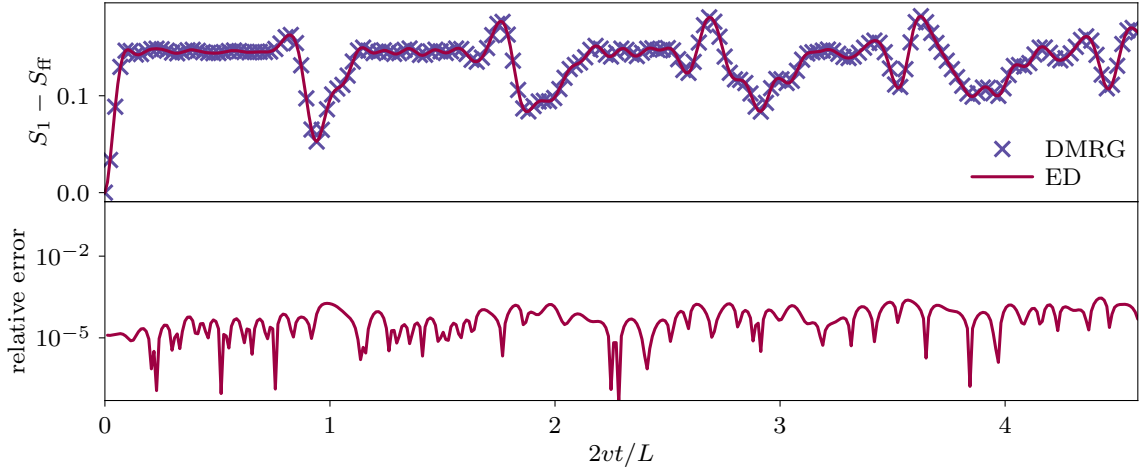


Figure 4.11: Comparison between ED and DMRG results for the waiting time dependence of the one-particle entanglement entropy S_1 after a quantum quench from free fermions to $V/J = -0.9$. We consider $N = 12$ fermions on $L = 24$ sites. The upper panel depicts the entropy S_1 obtained with ED (red line) and DMRG (blue crosses), and the lower panel shows the relative error of the DMRG approximation compared to the exact ED result. We find excellent agreement between DMRG and ED with a relative error well below 0.01% even for large waiting times after the quench.

time dependence of the one-particle entanglement entropy in a system of $N = 12$ fermions on $L = 24$ lattice sites after an interaction quench from free fermions to $V/J = -0.9$ (upper panel). The DMRG results are obtained with a Trotter evolution with time step $\delta t = 0.01$. Even for late times $t \approx 100$, the relative errors (lower panel) of the DMRG results are well below 0.01 %.

4.6 Numerical results and comparison to analytical bosonization results

Major parts of this section closely follow the publication: Matthias Thamm, Harini Radhakrishnan, Hatem Barghathi, Bernd Rosenow, and Adrian Del Maestro, *One-particle entanglement for one dimensional spinless fermions after an interaction quantum quench*, Physical Review B 106, 165116 (2022) [327].

4.6.1 Equilibrium case

Fig. 4.12 shows the one-particle entanglement entropy (crosses) calculated with DMRG for $N = 51$ for a large range of interaction strengths spanning all phases in the J - V model. For $V/J = -2$, the first order phase transition is clearly visible and $S_\alpha - S_{\text{ff}}$ remains stable, reaching the theoretical value $\ln(2)$ [145] for large negative V/J . For free fermions, where $V/J = 0$ in the LL phase, the one-particle entanglement entropy vanishes as expected [145]. Additionally, at the second-order phase transition into the charge density wave phase near $V/J = 2$, a change in the slope of the entropy is visible, which then slowly approaches the theoretical value $\ln(2)$ [145].

For comparison with field theory, we first estimate the thermodynamic limit $N \rightarrow \infty$ by finite size scaling of the numerical results for the one-particle entanglement entropy, with a general

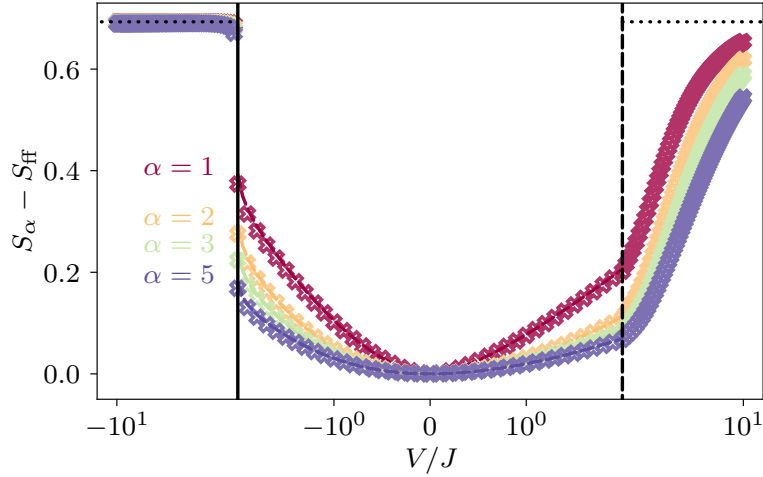


Figure 4.12: One-particle Rényi entanglement entropy S_α for different values of the Rényi index α as a function of the interaction strength V/J where S_{ff} is the one-particle entropy for free fermions. The crosses are obtained using DMRG for $N = 51$ on a lattice of $L = 102$ sites. Solid lines depict extrapolation to the thermodynamic limit from ED and DMRG data, and dashed horizontal lines show theory predictions for $|V/J| \rightarrow \infty$. Phase transitions in the J - V model are marked with vertical lines at $V/J = \pm 2$.

scaling form introduced by Haque *et al.* [145] and confirmed in subsequent works [152, 153]:

$$S_\alpha(N, V/J) = \ln(N) + A_\alpha(V/J) + \mathcal{O}(N^{-\lambda}) , \quad (4.182)$$

with $\lambda > 0$. In subsequent figures, we will focus on the behavior of the constant correction $A_\alpha(V/J)$ to the leading order logarithmic scaling.

For reliable finite size scaling, we calculate $S_\alpha - S_{\text{ff}}$ for systems with $N = 2, 3, \dots, 19$ fermions using ED and fermion numbers between $N = 17$ and $N = 51$ using DMRG and then extrapolate linearly to $1/N \rightarrow 0$ (white filled circles in Fig. 4.13). We find very good $1/N$ scaling and excellent agreement between exact ED (colored circles in Fig. 4.13) and approximate DMRG (crosses in Fig. 4.13) everywhere in the LL phase.

We perform this finite size scaling for all calculated interaction strengths V/J and plot the constant contribution to the one-particle entanglement entropies (A_α , circles) as a function of V/J in Fig. 4.14 along with the numerically integrated Luttinger liquid result from Eq. (4.129) and Eq. (4.130) (dashed lines) for a fixed interaction cutoff $\varepsilon = 0.84$ determined via fitting. We find excellent agreement between LL theory with this fixed cutoff and numerical results for the J - V model for small to moderate interaction strengths $-0.5 < V/J < 1.5$. Close to the phase transitions and especially for large negative interaction strengths $V/J \rightarrow -2$, where $\gamma_{\text{eq}} \rightarrow \infty$, significant deviations from the low energy LL theory are apparent.

To systematically study for which interactions the J - V model can be accurately described by the LL model, we fit the bosonization prediction for each interaction strength individually to the finite size scaled data for the von Neumann entropy $A_1(V/J)$ as defined in Eq. (4.182), to determine an effective interaction cutoff $\varepsilon_{\text{fit}}(V/J)$ (red circles in the main panel of Fig. 4.15). We find an extended region with $\varepsilon = 0.84$ (dashed, black line) for small negative and positive interactions V/J where the cutoff has minimal dependence on the interaction strength. With the obtained effective cutoff, we can fit the LL model at every point in the LL phase to the J - V model with excellent agreement as shown in the inset of Fig. 4.15 where we plot the one-particle

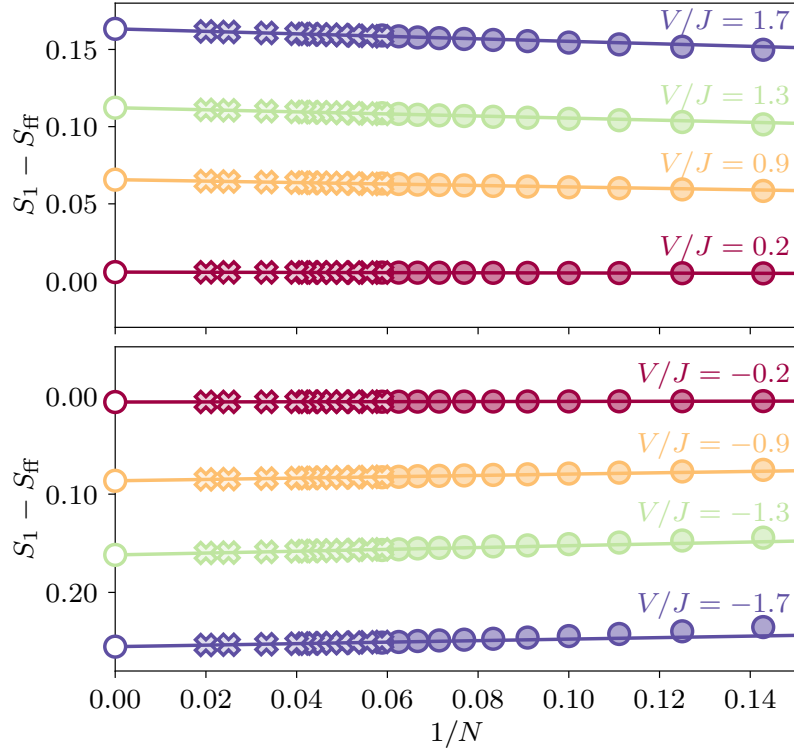


Figure 4.13: Finite size scaling of the equilibrium one-particle von Neumann entanglement entropy S_1 for various interaction strengths V/J where the free fermion contribution S_{ff} has been subtracted. Results obtained with DMRG (crosses) and ED (circles) are shown together along with a linear extrapolation to the thermodynamic limit $1/N \rightarrow 0$. ED provides access to lattices at half filling with up to $N = 19$ fermions and using DMRG lattices with more than $N = 51$ fermions can be studied.

entanglement entropies from numerics and for the effective interaction cutoff ε_{fit} . An interaction dependent cutoff for large interactions is also a consequence of approximating the q dependence of the exponent γ_{eq} by the cutoff $e^{-\varepsilon|q|}$ in field theory calculations (see Eq. (4.118)) in order to make the q sums analytically tractable.

4.6.2 Interaction quantum quench

We indeed observe the recurrence time $\Delta t = L/(2v)$ (Fig. 4.16 and Fig. 4.17a) predicted by the LL theory, which indicates that after the quench, density waves propagate with velocity v , (see Eq. (4.140)) through the lattice of length L , where the maximal distance between two points is $L/2$ due to the periodic boundary conditions. We show the waiting time dependence of the von Neumann entropy for several lattice sizes (solid lines) in Fig. 4.17a together with the steady state values (empty circles) obtained by averaging the entropy $S_1 - \ln(N)$ for times after the initial increase. Here, the entanglement entropy has plateaus with length proportional to L and regions, independent of the system size, where the entropy decreases and then increases to the next plateau. Convergence to the steady state can thus be understood from a running average (Fig. 4.17b), and in the thermodynamic limit, where the plateau size is infinite, the system reaches the steady state with entanglement entropy obtained by finite size scaling the entropy averages from the finite systems. Even for these relatively small systems, the fast

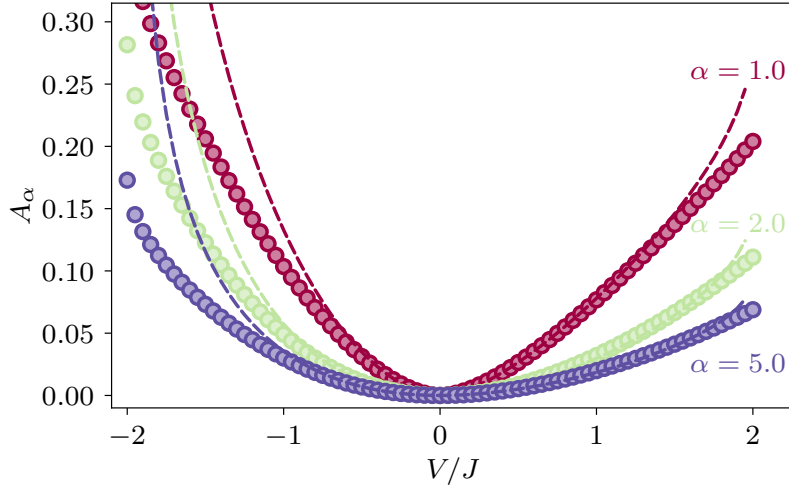


Figure 4.14: Interaction strength V/J dependence of the constant contribution to the one-particle Rényi entanglement entropies extrapolated to the thermodynamic limit A_α for Rényi indices $\alpha = 1, 2$, and 5 together with the prediction from bosonization for a fixed interaction cutoff $\varepsilon = 0.84$. We find very good agreement between the Luttinger liquid prediction and the numerical results for the J - V model in region $V/J \in [-0.5, 1.5]$.

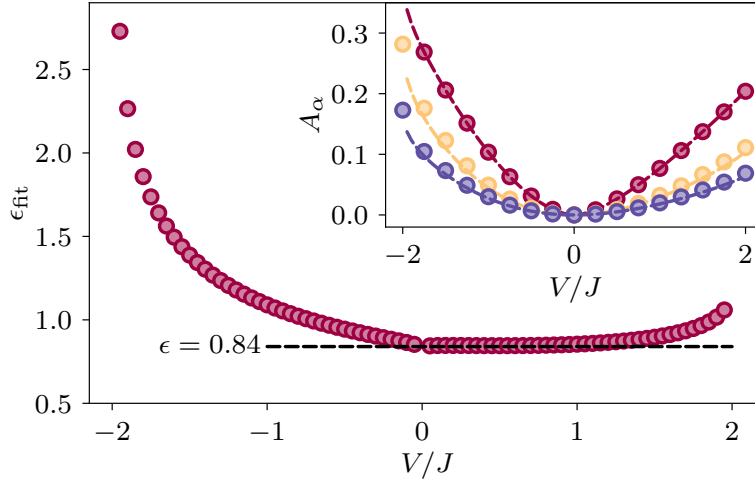


Figure 4.15: Effective interaction cutoff ε_{fit} as a function of the interaction strength V/J (red circles) obtained by fitting the Luttinger liquid prediction for the one-particle von Neumann entanglement entropy for each interaction region strength individually to the numerical data of the J - V model. We find an extended flat region of the effective cutoff $\varepsilon = 0.84$ (dashed, black line) that is nearly independent of interaction strength. The inset depicts numerical results for the Rényi entropies with $\alpha = 1$ (red circles), $\alpha = 2$ (yellow circles), and $\alpha = 5$ (blue circles) together with the field theory prediction using the fitted interaction dependent cutoff ε_{fit} .

decrease between consecutive steady state averages shows fast convergence to the thermodynamic limit. Such advantageous finite size scaling properties of the particle entanglement entropy were recently reported [161]. To estimate errors in the steady state averages, we use a blocking method [497] by consecutively averaging neighboring values in the time series and computing the error of the mean in each averaging step until it reaches a plateau. To further include errors

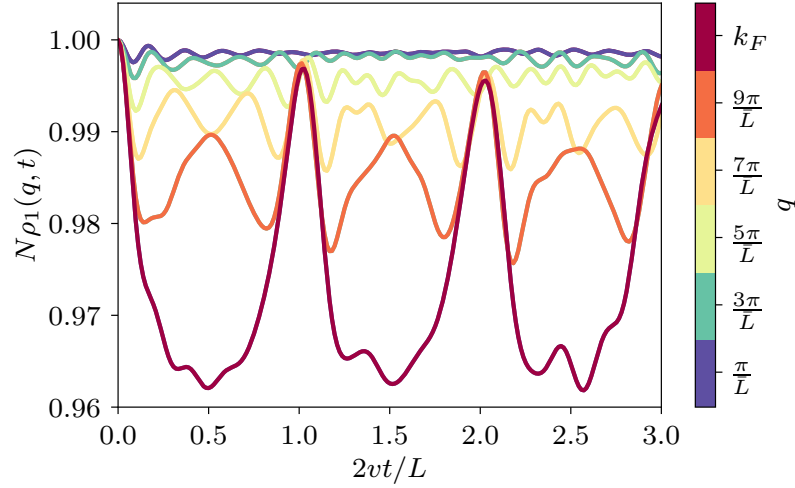


Figure 4.16: Dependence of eigenvalues of the one body density matrix ρ_1 on the rescaled waiting time $2vt/L$ after the quantum quench computed via exact diagonalization for a final interaction strength $V/J = -0.5$. We show a quarter of the spectrum with $0 < q < k_F$ for a system of $N = 12$ fermions on $L = 24$ lattice sites. The largest contributions to the one particle entanglement entropy come from the eigenvalues close to the Fermi levels which show the largest oscillation amplitude and the recurrence time $L/(2v)$ that appears in the entropy.

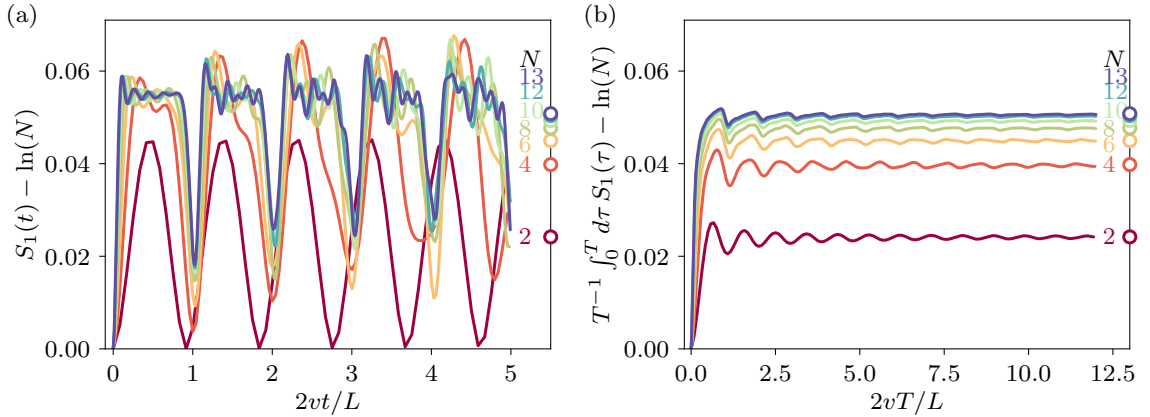


Figure 4.17: (a) One-particle von Neumann entanglement entropy as a function of the rescaled waiting time $2vt/L$ after the quantum quench at $t = 0$ for systems with $L = 2N$ sites at an interaction strength of $V/J = -0.5$. (b) Running average of the entropy in (a). We observe a steep increase of the entropy on a very short timescale after the quench and then recurrences with length N where the entropies oscillate around a constant steady state value (empty circles). The fast decrease in the distance between the steady state values with N suggests fast convergence to the thermodynamic limit. As the plateau length is proportional to the system length L and the length of the regions between the plateaus are independent of L , finite size scaling of the average will converge to the plateau height in the thermodynamic limit.

due to the finite time step and the endpoint of the time series, we additionally divide the time series into the individual N_b recurrence blocks with entropy averages M_i and add the error of the means $\text{mean}(M_i)/\sqrt{N_b}$, as well as the difference between the mean of the entropy time series

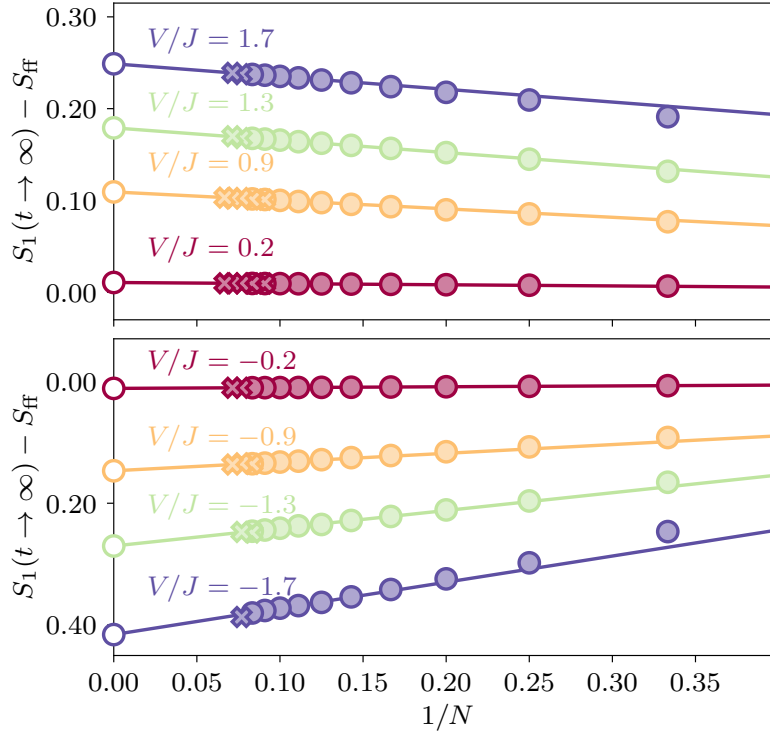


Figure 4.18: Finite size scaling of the steady state one-particle von Neumann entanglement entropy $S_1(t \rightarrow \infty)$ for various values of the post quench interaction strength V/J obtained with ED (filled circles) and with DMRG (filled crosses). We estimate the thermodynamic limit values of the entropy (empty circles) using linear extrapolation to $1/N \rightarrow 0$.

and the average of the M_i to the blocking error.

We compute the one-particle entanglement entropy for different interaction strengths V/J to again linearly extrapolate to the thermodynamic limit, $1/N \rightarrow 0$ (Fig. 4.18). In the case of small interactions, we can perform the time evolution on V100 GPUs for systems with up to $L = 30$ lattice sites, however, the ground states obtained with DMRG require more memory for larger interactions γ_{eq}^2 to perform calculations to the same accuracy. While for intermediate interactions, $V/J \geq 1.3$ and $-0.9 \leq V/J < 0$, we achieve results up to $L = 28$ sites, the calculation exceeds GPU memory available to us for $L \geq 28$ in the case of very strong interactions $V/J \leq -1.3$. Nonetheless, these additional points of the one-particle entanglement entropy obtained with the GPU accelerated tDMRG allow for relative improvements of the thermodynamic limit extrapolation from finite size scaling by up to 1.2% compared to using ED data alone. We find that even for these relatively small systems, linear extrapolation accurately describes the data in the whole Luttinger liquid phase. For all parameters where we compute one-particle entanglement entropies for both ED (circles) and DMRG (crosses), we find excellent agreement.

Performing the finite size scaling for all computed interaction strengths V/J , we obtain the interaction dependence of the steady state one-particle entanglement entropy in the thermodynamic limit (circles in Fig. 4.19) which we plot together with the entropy obtained from numerically computing the Fourier transform and numerically integrating the analytical steady state result from bosonization in Eq. (4.157) (dashed line, Fig. 4.19) for a fixed interaction cutoff $\varepsilon = 0.84$ determined in the ground state. We observe very similar agreement between results for the LL model and numerical results for the J - V model as in the equilibrium case Fig. 4.14

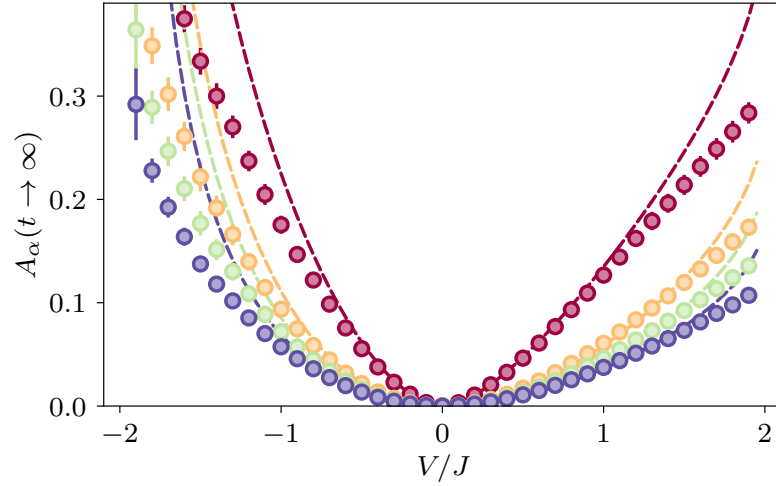


Figure 4.19: One-particle entanglement entropies extrapolated from the numerical steady state estimates to the thermodynamic limit as a function of the post quench interaction strength V/J (circles) for Rényi indices $\alpha = 1$ (red), $\alpha = 2$ (yellow), $\alpha = 3$ (green), and $\alpha = 5$ (blue). Dashed lines depict the corresponding theory predictions from the steady states after the quench in the Luttinger liquid model using a fixed interaction cutoff $\varepsilon = 0.84$ obtained from ground state calculations. Similar to the equilibrium case, we find good agreement between LL prediction and numerical data for moderate interaction strengths.

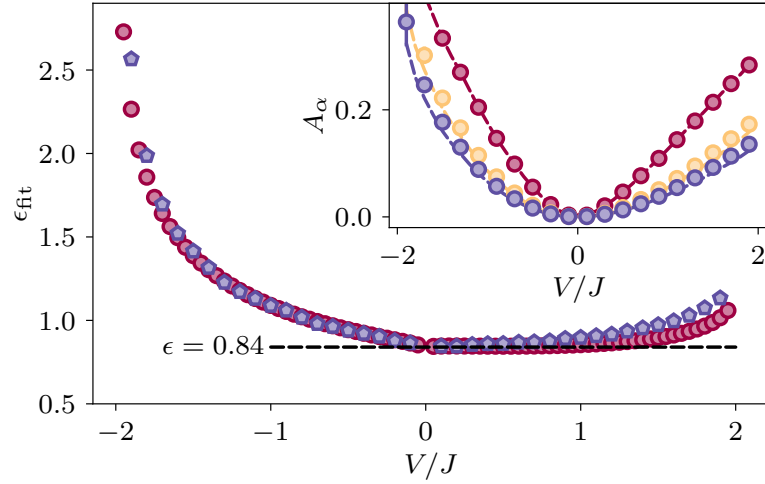


Figure 4.20: Interaction dependence of the effective cutoff ε_{fit} (blue pentagons) obtained by fitting the steady state of the Luttinger liquid model at each interaction strength V/J individually to the numerical data of the von Neumann entanglement entropy. For comparison, we show again the effective cutoff obtained from the ground state case (red circles) and find very good agreement with a quasi-plateau in the region $0 < V/J < 1$. The inset depicts numerical data for Rényi entropies with $\alpha = 1$ (red circles), $\alpha = 2$ (yellow circles), and $\alpha = 5$ (blue circles) together with the fitted field theory steady state predictions for the effective cutoff ε_{fit} .

when using the same cutoff.

We again fit an effective interaction cutoff (blue pentagons in Fig. 4.20) at each interaction strength V/J separately to match the LL solution to the von Neumann entropy from the J - V

model (red circles in the inset of Fig. 4.20). We find very good agreement with the interaction cutoff determined for the equilibrium ground state case (red circles in the main panel of Fig. 4.20) which suggests that the parameter ε of the LL calculation can be fixed by numerical analysis of the J - V model in the region $-0.5 < J/V < 1.5$, where the low energy LL approximation is most accurate resulting in $\varepsilon = 0.84$.

4.7 Conclusions

In this chapter, we presented a comprehensive study of the one-body reduced density matrix and the corresponding one-particle entanglement entropy for the J - V model of one-dimensional interacting fermions. We considered two cases: (i) the equilibrium case for the ground state of the J - V model at finite interaction strength and (ii) the waiting time dependence after an interaction quantum quench from free fermions to finite interactions. For both cases, we performed large scale numerical computations using exact diagonalization, where exploiting translational, particle-hole, and reflection symmetry allowed us to consider systems up to $L = 38$ lattice sites for case (i) and $L = 13$ sites for case (ii) at half filling. Using appropriate initial states and an orthogonal subspace to the ground state based on analytical considerations, we were further able to stabilize DMRG results for the system with (anti-)periodic boundary conditions for lattices with up to $L = 102$ sites in the ground state. By using time dependent DMRG on massively parallelized GPUs, we even performed the time evolution for up to $L = 30$ lattice sites. Being able to obtain reliable numerical results for such large systems, allowed us to perform accurate finite size scaling to the thermodynamic limit.

Within the Luttinger liquid phase, we complemented the numerical results by a bosonization calculation, where an interaction cutoff is introduced to account for the short range interactions of the J - V model. In the thermodynamic limit $L \rightarrow \infty$, we compared numerical and analytical results which allowed us to self-consistently determine the interaction cutoff. Using this cutoff, we find excellent agreement between bosonization and numerical results for small interaction strengths even for the steady state after the quench, such that the Luttinger liquid bosonization calculation provide quantitatively accurate results far beyond the computation of scaling exponents.

In addition, the methods and the implementation of the numerical computations open up new avenues for future research: It might be interesting to study the bosonization calculation for the two-particle entanglement entropy. Further, the numerical implementations also give access to higher order n -particle entanglement entropy for large systems, such that scaling of the entropy with n and N can be studied. Ultimately, it would be interesting to expand the entanglement entropy in terms of higher order reduced density matrices which would involve the one-particle entanglement entropy as a starting point [327].

5. Summary

We would like to use this final section to summarize the main results: In this dissertation, we approached the study of mesoscopic systems from very different directions, using several methods: Random matrix theory allowed us to gain insight into the weight matrices of deep neural networks:

- (i) By comparing the spectra of neural network weight matrices to universal RMT predictions, we found excellent agreement of the level spacing distribution of unfolded singular values with the Wigner surmise, and we showed that the level number variance grows only logarithmically as predicted for random matrices. This is supported by a careful analysis using Kolmogorov-Smirnov tests for several self-trained networks and even for large pre-trained networks – covering different network architectures. We compared the Marcenko-Pastur distribution to the singular value spectra and the Porter-Thomas distribution to the singular vector entries which revealed that small and intermediate singular values and corresponding vectors follow the RMT predictions, and for large values they deviate.
- (ii) When performing the same analysis for networks trained in various learning regimes, we found that the universal predictions are unchanged while comparing the spectra and eigenvectors with the Marcenko-Pastur and Porter-Thomas distributions respectively allows us to distinguish the learning regimes without knowledge of the initial weights.
- (iii) By training additional networks with label noise and systematically setting singular values to zero, starting from the smallest ones – reducing the rank of the matrices – and monitoring the training and generalization accuracy shows that: small singular values and corresponding vectors have no importance to the network’s performance at all, intermediate values can be important for the training accuracy in presence of label noise by encoding memorized labels, and large singular values and vectors store the information about the underlying rule.
- (iv) Based on the observation of a random bulk and a few large singular values encoding the rule which are stored separated from the memorized noise, we proposed a noise filtering algorithm that both removes the singular values encoding only noise and reverts the level repulsion of the large singular values due to the random bulk by down shifting the large singular values. We find significant improvements of the generalization performance for networks trained with label noise when using the filtering technique. These improvements vanish if the network is trained with a schedule that promotes severe overfitting such that the boundary between information and noise in the spectra is blurred.

By using a scattering matrix formalism adapted to include the effects of superconductivity, finite temperature, and Coulomb blockade, we studied the coherent transmission amplitude through a semiconductor-superconductor hybrid nano-wire hosting Majorana zero modes embedded into an arm of an electron interferometer:

- (i) By combining an analytical and numerical analysis of one-dimensional Majorana wires, we were able to explain the occurrence of an amplitude maximum as a function of the Zeeman field at the onset of the topological phase that had recently been measured experimentally [130]. When entering the topological phase, there is a concurrence between two correlation

lengths for the localization of the Majorana zero modes that is first dominated by an s-wave correlation length that diverges at the phase transition and transitions at the region of the maximum into a p-wave correlation length, which is dominant for large fields where it asymptotically increases proportionally to the Zeeman field.

- (ii) As the transmission amplitude is determined by the weight of the Majorana wave functions at the wire ends – where they are localized with localization length independent of the wire length – and the inverse of the charging energy which is proportional to the inverse wire length, the transmission amplitude maximum increases linearly with the wire length in the presence of Majorana zero modes. For an extended, trivial level, the wave functions' weight decreases linearly with the wire length such that the transmission amplitude is independent of wire length.
- (iii) We showed that higher levels have little influence on the transmission amplitude in the presence of Majorana zero modes, and that the transmission amplitude at finite temperatures is even able to distinguish Majorana zero modes from trivial pseudo-Majorana modes. In the latter case, a pair of pseudo-Majorana levels can mimic the magnetic field dependence of the direct conductance through the wire; however, for the coherent transmission amplitude measured with the interferometer setup, there are two contributions with equal magnitude but opposite signs that interfere destructively.
- (iv) Even though these results are robust to small amounts of disorder, the topological phase – and with it the Majorana modes – can be destroyed by strong disorder that is prevalent in the hybrid wires in experiments.

By placing an array of gates in proximity to the nano-wire, we made a fruitful connection to the field of Machine Learning by using the CMA-ES algorithm to tune the gate voltages in order to maximize the amplitude of coherent transmission in the wire:

- (i) We find that optimization in the absence of disorder significantly improves the localization of the Majorana zero modes.
- (ii) The algorithm is capable of learning disorder profiles and even to restore Majorana modes that were fully destroyed by strong disorder by using a feasible number of about 20 gates.
- (iii) The underlying CMA-ES machine learning algorithm is system agnostic such that it can be applied to a large variety of scenarios. We discuss convergence of the algorithm, the challenges when choosing a metric for optimization, and the choice of realistic parameters which will be useful for applying the method to any system in an experiment

Last but not least, by combining large scale exact diagonalization, approximate DMRG, and an analytical field theoretic bosonization calculation, we were able to study the one-particle entanglement entropy in one-dimensional systems of interacting fermions:

- (i) We provide a detailed study of the one-particle entanglement entropy in the one-dimensional J - V model both in the ground state and after an interaction quantum quench.
- (ii) Utilizing the symmetries of the system and state-of-the art numerical methods, we are able to compute the ground state one-particle entanglement entropy on a periodic lattice with exact diagonalization for up to $N = 19$ fermions on $N = 38$ lattice sites and even up to $N = 51$ fermions on $L = 102$ lattice sites using DMRG. For the time evolution after

the quench, we were able to consider up to $N = 13$ fermions on $L = 26$ sites with exact diagonalization and by using massively parallel GPUs, we pushed the time evolution with the DMRG code to up to $N = 15$ fermions on $L = 30$ sites for small interactions.

- (iii) By pushing the boundaries of the numerical exact diagonalization and density matrix renormalization group computations, we are able to accurately scale to the thermodynamic limit where we made contact to the analytic, field theory model. This allowed us to fix an interaction cutoff required in the continuum bosonization calculation to account for the short range interaction of the lattice model, such that the bosonization result provides accurate predictions for the one-body reduced density matrix in the Luttinger liquid phase.
- (iv) We provide the code for exact diagonalization in the ground state, exact diagonalization after the quench, as well as the DMRG code for the equilibrium case and the GPU assisted DMRG code for the time evolution after the quench as open source software [474].

Besides the direct results, the tools created for obtaining them open the way for several future research directions: The filtering technique for mitigating noise in weight matrices of neural networks may be further extended to more realistic models for how noise enters the weight matrices, and the methods for locating information in the weights might allow novel training algorithms to be designed which may ultimately improve training performance and speed and thus save resources already during training. The code developed for the numerical studies of the particle entanglement entropy is capable of computing more than just the one-particle entanglement entropy. It might therefore be possible to extract scaling relations of the particle entanglement with the system and partition size, and to facilitate a deeper understanding of particle entanglement and its applications. Tuning an array of gates using Machine Learning to stabilize features of a system and to counteract disorder has potentially many applications in all kind of nano-scale electronic devices. As disorder is currently a major limiting factor for realizing qubits based on Majorana zero modes, Machine Learning approaches – similar to the one we have shown here – might even help along the way to harvesting the full potential of topological quantum computing.

Bibliography

- [1] E. P. Wigner. On the statistical distribution of the widths and spacings of nuclear resonance levels. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 47, pages 790–798. Cambridge University Press, 1951.
- [2] E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Annals of Mathematics*, 62:548–564, 1955.
- [3] E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions ii. *Annals of Mathematics*, 65:203–207, 1957.
- [4] E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions i. In *The Collected Works of Eugene Paul Wigner*, pages 524–540. Springer, 1993.
- [5] E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions ii. In *The Collected Works of Eugene Paul Wigner*, pages 541–545. Springer, 1993.
- [6] A. M. G. Mayer and J. H. D. Jensen. Elementary theory of nuclear shell structure. *Wiley, New York*, 1955.
- [7] T. Papenbrock and H. A. Weidenmüller. Colloquium: Random matrices and chaos in nuclear spectra. *Reviews of Modern Physics*, 79(3):997, 2007.
- [8] R. U. Haq, A. Pandey, and O. Bohigas. Fluctuation properties of nuclear energy levels: Do theory and experiment agree? *Physical Review Letters*, 48(16):1086, 1982.
- [9] O. Bohigas, M.-J. Giannoni, and C. Schmit. Characterization of chaotic quantum spectra and universality of level fluctuation laws. *Physical review letters*, 52(1):1, 1984.
- [10] V. Zelevinsky, B. A. Brown, N. Frazier, and M. Horoi. The nuclear shell model as a testing ground for many-body quantum chaos. *Physics reports*, 276(2-3):85–176, 1996.
- [11] T. Guhr, A. Müller-Groeling, and H. A. Weidenmüller. Random-matrix theories in quantum physics: common concepts. *Physics Reports*, 299(4-6):189–425, 1998. ISSN 03701573.
- [12] H. A. Weidenmüller and G. E. Mitchell. Random matrices and chaos in nuclear physics: Nuclear structure. *Reviews of Modern Physics*, 81(2):539, 2009.
- [13] J. Wishart. The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, pages 32–52, 1928.
- [14] A. T. James. Normal multivariate analysis and the orthogonal group. *The Annals of Mathematical Statistics*, 25(1):40–75, 1954.
- [15] A. T. James. The distribution of the latent roots of the covariance matrix. *The Annals of Mathematical Statistics*, pages 151–158, 1960.
- [16] A. T. James. Distributions of matrix variates and latent roots derived from normal samples. *The Annals of Mathematical Statistics*, 35(2):475–501, 1964.

- [17] Y. Aharonov and D. Bohm. Significance of Electromagnetic Potentials in the Quantum Theory. *Physical Review*, 115(3):485, 1959.
- [18] C. E. Porter. Statistical theories of spectra: fluctuations. Technical report, 1965.
- [19] T. A. Brody, J. Flores, J. B. French, P. A. Mello, A. Pandey, and S. S. M. Wong. Random-matrix physics: spectrum and strength fluctuations. *Reviews of Modern Physics*, 53(3):385–479, 1981.
- [20] J. Garg. *Statistical Properties of Nuclei: Proceedings of the International Conference on Statistical Properties of Nuclei, Held at Albany, New York, August 23–27, 1971*. Springer Science & Business Media, 2012.
- [21] H. A. Weidenmüller. Transport theories of heavy-ion reactions. *Progress in particle and nuclear physics*, 3:49–126, 1980.
- [22] J. Moser. *Stable and random motions in dynamical systems: With special emphasis on celestial mechanics*, volume 1. Princeton university press, 2001.
- [23] H. G. Schuster. Deterministic chaos, physik, 1984.
- [24] T. H. Seligman, J. J. M. Verbaarschot, and M. R. Zirnbauer. Quantum spectra and transition from regular to chaotic classical motion. *Physical review letters*, 53(3):215, 1984.
- [25] S. Iida, H. A. Weidenmüller, and J. A. Zuk. Wave propagation through disordered media and universal conductance fluctuations. *Physical review letters*, 64(5):583, 1990.
- [26] F. Calogero. Solution of the one-dimensional n-body problems with quadratic and/or inversely quadratic pair potentials. *Journal of Mathematical Physics*, 12(3):419–436, 1971.
- [27] B. Sutherland. Exact results for a quantum many-body problem in one dimension. *Physical Review A*, 4(5):2019, 1971.
- [28] B. Sutherland. Exact results for a quantum many-body problem in one dimension. ii. *Physical Review A*, 5(3):1372, 1972.
- [29] B. D. Simons, P. A. Lee, and B. L. Altshuler. Exact pair correlation of the one-dimensional quantum gas with $1/r^2$ repulsion derived from the symplectic dyson ensemble. *Physical Review B*, 48(15):11450, 1993.
- [30] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, and H. E. Stanley. Universal and nonuniversal properties of cross correlations in financial time series. *Physical Review Letters*, 83:1471–1474, Aug 1999.
- [31] L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters. Noise Dressing of Financial Correlation Matrices. *Physical Review Letters*, 83(7):1467–1470, 1999.
- [32] L. Laloux, P. Cizeau, M. Potters, and J.-P. Bouchaud. Random Matrix Theory and Financial Correlations. *International Journal of Theoretical and Applied Finance*, 03(03):391–397, 2000. ISSN 0219-0249.

- [33] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, T. Guhr, and H. E. Stanley. Random matrix approach to cross correlations in financial data. *Physical Review E*, 65(6 Pt 2):066126, 2002. ISSN 1539-3755.
- [34] B. Rosenow, V. Plerou, P. Gopikrishnan, and H. E. Stanley. Portfolio optimization and the random magnet problem. *Europhysics Letters (EPL)*, 59(4):500–506, aug 2002.
- [35] C. Louart, Z. Liao, and R. Couillet. A random matrix approach to neural networks. *The Annals of Applied Probability*, 28(2):1190 – 1248, 2018.
- [36] J. Pennington and P. Worah. Nonlinear random matrix theory for deep learning. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(12):124005, dec 2019.
- [37] C. H. Martin and M. W. Mahoney. Implicit self-regularization in deep neural networks: Evidence from random matrix theory and implications for learning. *Journal of Machine Learning Research*, 22(165):1–73, 2021.
- [38] W. Samek, G. Montavon, S. Lapuschkin, C. J. Anders, and K.-R. Müller. Explaining deep neural networks and beyond: A review of methods and applications. *Proceedings of the IEEE*, 109(3):247–278, 2021.
- [39] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals. Understanding deep learning requires rethinking generalization. *arXiv:1611.03530*, 2016.
- [40] S. Liu, Z. Zhu, Q. Qu, and C. You. Robust training under label noise by over-parameterization. *preprint arXiv:2202.14026*, 2022.
- [41] H. Song, M. Kim, D. Park, Y. Shin, and J.-G. Lee. Learning from noisy labels with deep neural networks: A survey. *IEEE Transactions on Neural Networks and Learning Systems*, 2022.
- [42] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.
- [43] I. Goodfellow, Y. Bengio, and A. Courville. *Deep learning*. MIT press, 2016.
- [44] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105, 2012.
- [45] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, et al. Mastering the game of go without human knowledge. *Nature*, 550(7676):354–359, 2017.
- [46] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová. Machine learning and the physical sciences. *Reviews of Modern Physics*, 91(4):045002, 2019.
- [47] Y. Bahri, J. Kadmon, J. Pennington, S. S. Schoenholz, J. Sohl-Dickstein, and S. Ganguli. Statistical mechanics of deep learning. *Annual Review of Condensed Matter Physics*, 11: 501–528, 2020.
- [48] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.

- [49] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo. Neural-network quantum state tomography. *Nature Physics*, 14(5):447–450, 2018. ISSN 1745-2473.
- [50] C. Cao, S.-Y. Hou, N. Cao, and B. Zeng. Supervised learning in Hamiltonian reconstruction from local measurements on eigenstates. *Journal of Physics: Condensed Matter*, 33(6):064002, 2020.
- [51] J. Carrasquilla and R. G. Melko. Machine learning phases of matter. *Nature Physics*, 13(5):431–434, 2017.
- [52] Evert P. L. van Nieuwenburg, Ye-Hua Liu, and Sebastian D. Huber. Learning phase transitions by confusion. *Nature Physics*, 13(5):435–439, 2017. ISSN 1745-2473.
- [53] K. Ch’ng, J. Carrasquilla, R. G. Melko, and E. Khatami. Machine learning phases of strongly correlated fermions. *Physical Review X*, 7(3):031038, 2017.
- [54] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst. Machine learning quantum phases of matter beyond the fermion sign problem. *Scientific reports*, 7(1):1–10, 2017.
- [55] P. Huembeli, A. Dauphin, and P. Wittek. Identifying quantum phase transitions with adversarial neural networks. *Physical Review B*, 97:134109, Apr 2018.
- [56] T. Karzig, C. Knapp, R. M. Lutchyn, P. Bonderson, M. B. Hastings, C. Nayak, J. Alicea, K. Flensberg, S. Plugge, Y. Oreg, C. M. Marcus, and M. H. Freedman. Scalable designs for quasiparticle-poisoning-protected topological quantum computation with Majorana zero modes. *Physical Review B*, 95(23):235305, 2017.
- [57] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, Brandao, F. G. S. L., D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019. ISSN 1476-4687.
- [58] D. T. Lennon, H. Moon, L. C. Camenzind, L. Yu, D. M. Zumbühl, G. A. D. Briggs, M. A. Osborne, E. A. Laird, and N. Ares. Efficiently measuring a quantum device using machine learning. *npj Quantum Information*, 5(1):1–8, 2019. ISSN 2056-6387.
- [59] Y. Oreg and F. von Oppen. Majorana Zero Modes in Networks of Cooper-Pair Boxes: Topologically Ordered States and Topological Quantum Computation. *Annual Review of Condensed Matter Physics*, 11(1):397–420, 2020. ISSN 1947-5454.
- [60] N. Ares. Machine learning as an enabler of qubit scalability. *Nature Reviews Materials*, 6(10):870–871, 2021. ISSN 2058-8437.

- [61] T. A. Baart, P. T. Eendebak, C. Reichl, W. Wegscheider, and L. M. K. Vandersypen. Computer-automated tuning of semiconductor double quantum dots into the single-electron regime. *Applied Physics Letters*, 108(21):213104, 2016. ISSN 0003-6951.
- [62] D. L. Craig, H. Moon, F. Fedele, D. T. Lennon, B. van Straaten, F. Vigneau, L. C. Camenzind, D. M. Zumbühl, G. A. D. Briggs, M. A. Osborne, D. Sejdinovic, and N. Ares. Bridging the reality gap in quantum devices with physics-aware machine learning, 2021.
- [63] J. Ziegler, T. McJunkin, E. S. Joseph, S. S. Kalantre, B. Harpt, D. E. Savage, M. G. Lagally, M. A. Eriksson, J. M. Taylor, and J. P. Zwolak. Toward Robust Autotuning of Noisy Quantum Dot Devices.
- [64] A. Frees, J. K. Gamble, D. R. Ward, R. Blume-Kohout, M. A. Eriksson, M. Friesen, and S. N. Coppersmith. Compressed Optimization of Device Architectures for Semiconductor Quantum Devices. *Physical Review Applied*, 11(2):024063, 2019.
- [65] S. S. Kalantre, J. P. Zwolak, S. Ragole, X. Wu, N. M. Zimmerman, M. D. Stewart, and J. M. Taylor. Machine learning techniques for state recognition and auto-tuning in quantum dots. *npj Quantum Information*, 5(1):1–10, 2019. ISSN 2056-6387.
- [66] H. Moon, D. T. Lennon, J. Kirkpatrick, N. M. van Esbroeck, L. C. Camenzind, L. Yu, F. Vigneau, D. M. Zumbühl, G. A. D. Briggs, M. A. Osborne, D. Sejdinovic, E. A. Laird, and N. Ares. Machine learning enables completely automatic tuning of a quantum device faster than human experts. *Nature Communications*, 11(1):4161, 2020. ISSN 2041-1723.
- [67] H.-C. Ruiz Euler, M. N. Boon, J. T. Wildeboer, B. van de Ven, T. Chen, H. Broersma, P. A. Bobbert, and W. G. van der Wiel. A deep-learning approach to realizing functionality in nanoelectronic devices. *Nature Nanotechnology*, 15(12):992–998, 2020. ISSN 1748-3395.
- [68] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O’Brien. Quantum computers. *Nature*, 464(7285):45–53, 2010. ISSN 1476-4687.
- [69] R. Hanson, L. P. Kouwenhoven, J. R. Petta, S. Tarucha, and L. M. K. Vandersypen. Spins in few-electron quantum dots. *Reviews of Modern Physics*, 79(4):1217–1265, 2007.
- [70] C. Kloeffel and D. Loss. Prospects for Spin-Based Quantum Computing in Quantum Dots. *Annual Review of Condensed Matter Physics*, 4(1):51–81, 2013. ISSN 1947-5454.
- [71] L. M. K. Vandersypen, H. Bluhm, J. S. Clarke, A. S. Dzurak, R. Ishihara, A. Morello, D. J. Reilly, L. R. Schreiber, and M. Veldhorst. Interfacing spin qubits in quantum dots and donors—hot, dense, and coherent. *npj Quantum Information*, 3(1):1–10, 2017. ISSN 2056-6387.
- [72] T. Botzem, M. D. Shulman, S. Foletti, S. P. Harvey, O. E. Dial, P. Bethke, P. Cerfontaine, R. P. G. McNeil, D. Mahalu, V. Umansky, A. Ludwig, A. Wieck, D. Schuh, D. Bougeard, A. Yacoby, and H. Bluhm. Tuning Methods for Semiconductor Spin Qubits. *Physical Review Applied*, 10(5), 2018.
- [73] J. D. Teske, S. S. Humpohl, R. Otten, P. Bethke, P. Cerfontaine, J. Dedden, A. Ludwig, A. D. Wieck, and H. Bluhm. A machine learning approach for automated fine-tuning of semiconductor spin qubits. *Applied Physics Letters*, 114(13):133102, 2019. ISSN 0003-6951.

- [74] A. R. Mills, M. M. Feldman, C. Monical, P. J. Lewis, K. W. Larson, A. M. Mounce, and J. R. Petta. Computer-automated tuning procedures for semiconductor quantum dot arrays. *Applied Physics Letters*, 115(11):113501, 2019. ISSN 0003-6951.
- [75] R. Durrer, B. Kratochwil, J. V. Koski, A. J. Landig, C. Reichl, W. Wegscheider, T. Ihn, and E. Grepova. Automated Tuning of Double Quantum Dots into Specific Charge States Using Neural Networks. *Physical Review Applied*, 13(5):054019, 2020.
- [76] N. M. van Esbroeck, D. T. Lennon, H. Moon, V. Nguyen, F. Vigneau, L. C. Camenzind, L. Yu, D. M. Zumbühl, G. A. D. Briggs, D. Sejdinovic, and N. Ares. Quantum device fine-tuning using unsupervised embedding learning. *New Journal of Physics*, 22(9):095003, 2020. ISSN 1367-2630.
- [77] F. Fedele, A. Chatterjee, S. Fallahi, G. C. Gardner, M. J. Manfra, and F. Kuemmeth. Simultaneous Operations in a Two-Dimensional Array of Singlet-Triplet Qubits. *PRX Quantum*, 2(4):040306, 2021. ISSN 2691-3399.
- [78] O. Krause, A. Chatterjee, F. Kuemmeth, and E. van Nieuwenburg. Learning coulomb diamonds in large quantum dot arrays. *arXiv preprint arXiv:2205.01443*, 2022.
- [79] J. Alicea, Y. Oreg, G. Refael, F. von Oppen, and M. P. A. Fisher. Non-Abelian statistics and topological quantum information processing in 1d wire networks. *Nature Physics*, 7(5):412, 2011. ISSN 1745-2481.
- [80] D. J. Clarke, J. D. Sau, and S. Tewari. Majorana fermion exchange in quasi-one-dimensional networks. *Physical Review B*, 84(3):035120, 2011. ISSN 2469-9950.
- [81] T. Hyart, B. van Heck, I. C. Fulga, M. Burrello, A. R. Akhmerov, and C. W. J. Beenakker. Flux-controlled quantum computation with Majorana fermions. *Physical Review B*, 88(3):035121, 2013. ISSN 2469-9950.
- [82] S. Das Sarma, M. Freedman, and C. Nayak. Majorana zero modes and topological quantum computation. *npj Quantum Information*, 1(1):1–13, 2015. ISSN 2056-6387.
- [83] D. Aasen, M. Hell, R. V. Mishmash, A. Higginbotham, J. Danon, M. Leijnse, T. S. Jespersen, J. A. Folk, C. M. Marcus, K. Flensberg, and J. Alicea. Milestones Toward Majorana-Based Quantum Computing. *Physical Review X*, 6(3):031016, 2016. ISSN 2160-3308.
- [84] R. M. Lutchyn, E. P. A. M. Bakkers, L. P. Kouwenhoven, P. Krogstrup, C. M. Marcus, and Y. Oreg. Majorana zero modes in superconductor–semiconductor heterostructures. *Nature Reviews Materials*, 3(5):52–68, 2018. ISSN 2058-8437.
- [85] S. Vijay and L. Fu. Teleportation-based quantum information processing with Majorana zero modes. *Physical Review B*, 94(23):235446, 2016. ISSN 2469-9950.
- [86] A. Y. Kitaev. Unpaired Majorana fermions in quantum wires. *Physics-Uspekhi*, 44(10S):131–136, 2001. ISSN 1063-7869.
- [87] R. M. Lutchyn, J. D. Sau, and S. Das Sarma. Majorana fermions and a topological phase transition in semiconductor-superconductor heterostructures. *Physical Review Letters*, 105(7):077001, 2010.

- [88] Y. Oreg, G. Refael, and F. von Oppen. Helical Liquids and Majorana Bound States in Quantum Wires. *Physical Review Letters*, 105(17):177002, 2010.
- [89] J. D. Sau, S. Tewari, R. M. Lutchyn, T. D. Stanescu, and S. Das Sarma. Non-Abelian quantum order in spin-orbit-coupled semiconductors: Search for topological Majorana particles in solid-state systems. *Physical Review B*, 82(21):214509, 2010.
- [90] V. Mourik, K. Zuo, S. M. Frolov, S. R. Plissard, E. P. A. M. Bakkers, and L. P. Kouwenhoven. Signatures of Majorana fermions in hybrid superconductor-semiconductor nanowire devices. *Science (New York, N.Y.)*, 336(6084):1003–1007, 2012. ISSN 1095-9203.
- [91] A. Das, Y. Ronen, Y. Most, Y. Oreg, M. Heiblum, and H. Shtrikman. Zero-bias peaks and splitting in an Al–InAs nanowire topological superconductor as a signature of Majorana fermions. *Nature Physics*, 8(12):887–895, 2012. ISSN 1745-2481.
- [92] M. T. Deng, C. L. Yu, G. Y. Huang, M. Larsson, P. Caroff, and H. Q. Xu. Anomalous zero-bias conductance peak in a Nb–InSb nanowire–Nb hybrid device. *Nano letters*, 12(12):6414–6419, 2012.
- [93] F. Nichele, A. C. C. Drachmann, A. M. Whiticar, E. C. T. O’Farrell, H. J. Suominen, A. Fornieri, T. Wang, G. C. Gardner, C. Thomas, A. T. Hatke, P. Krogstrup, M. J. Manfra, K. Flensberg, and C. M. Marcus. Scaling of Majorana Zero-Bias Conductance Peaks. *Physical review letters*, 119(13):136803, 2017.
- [94] L. P. Rokhinson, X. Liu, and J. K. Furdyna. The fractional a.c. Josephson effect in a semiconductor–superconductor nanowire as a signature of Majorana particles. *Nature Physics*, 8(11):795–799, 2012. ISSN 1745-2481.
- [95] S. M. Albrecht, A. P. Higginbotham, M. Madsen, F. Kuemmeth, T. S. Jespersen, J. Nygård, P. Krogstrup, and C. M. Marcus. Exponential protection of zero modes in Majorana islands. *Nature*, 531(7593):206, 2016. ISSN 1476-4687.
- [96] J. Klinovaja, P. Stano, and D. Loss. Transition from fractional to majorana fermions in rashba nanowires. *Physical review letters*, 109(23):236801, 2012.
- [97] S. Boutin, J. Camirand Lemyre, and I. Garate. Majorana bound state engineering via efficient real-space parameter optimization. *Physical Review B*, 98(21):214512, 2018.
- [98] N. Mohanta, T. Zhou, J.-W. Xu, J. E. Han, A. D. Kent, J. Shabani, I. Žutić, and A. Matos-Abiague. Electrical Control of Majorana Bound States Using Magnetic Stripes. *Physical Review Applied*, 12(3):034048, 2019.
- [99] S. Turcotte, S. Boutin, J. C. Lemyre, I. Garate, and M. Pioro-Ladrière. Optimized micro-magnet geometries for Majorana zero modes in low g -factor materials. *Physical Review B*, 102(12):125425, 2020.
- [100] A. Melo, T. Tanev, and A. R. Akhmerov. Greedy optimization of the geometry of majorana josephson junctions. *arXiv preprint arXiv:2205.05689*, 2022.
- [101] H. Zhang, C.-X. Liu, S. Gazibegovic, D. Xu, J. A. Logan, G. Wang, N. van Loo, J. D. S. Bommer, M. W. A. de Moor, D. Car, et al. Retraction note: Quantized majorana conductance. *Nature*, 591(7851):E30–E30, 2021.

- [102] H. Zhang, Ö. Gül, S. Conesa-Boj, M. P. Nowak, M. Wimmer, K. Zuo, V. Mourik, F. K. de Vries, J. van Veen, M. W. A. de Moor, J. D. S. Bommer, D. J. van Woerkom, D. Car, S. R. Plissard, E. P. A. M. Bakkers, M. Quintero-Pérez, M. C. Cassidy, S. Koelling, S. Goswami, K. Watanabe, T. Taniguchi, and L. P. Kouwenhoven. Ballistic superconductivity in semiconductor nanowires. *Nature Communications*, 8(1):16025, 2017. ISSN 2041-1723.
- [103] S. Ahn, H. Pan, B. Woods, T. D. Stanescu, and S. Das Sarma. Estimating disorder and its adverse effects in semiconductor Majorana nanowires. *arXiv:2109.00007*, 2021.
- [104] S. Das Sarma and H. Pan. Disorder-induced zero-bias peaks in Majorana nanowires. *Physical Review B*, 103(19):195158, 2021.
- [105] P. Yu, J. Chen, M. Gomanko, G. Badawy, Bakkers, E. P. A. M., K. Zuo, V. Mourik, and S. M. Frolov. Non-Majorana states yield nearly quantized conductance in proximatized nanowires. *Nature Physics*, 17(4):482–488, 2021. ISSN 1745-2481.
- [106] Microsoft Quantum: M. Aghaee, A. Akkala, Z. Alam, R. Ali, A. A. Ramirez, M. Andrzejczuk, A. E. Antipov, M. Astafev, Bela B., J. Becker, et al. Inas-al hybrid devices passing the topological gap protocol. *arXiv preprint arXiv:2207.02472*, 2022.
- [107] S. Takei, B. M. Fregoso, H.-Y. Hui, A. M. Lobos, and S. Das Sarma. Soft superconducting gap in semiconductor Majorana nanowires. *Physical review letters*, 110(18):186803, 2013.
- [108] D. Bagrets and A. Altland. Class D spectral peak in Majorana quantum wires. *Physical review letters*, 109(22):227005, 2012.
- [109] D. I. Pikulin, J. P. Dahlhaus, M. Wimmer, H. Schomerus, and C. W. J. Beenakker. A zero-voltage conductance peak from weak antilocalization in a Majorana nanowire. *New Journal of Physics*, 14(12):125011, 2012. ISSN 1367-2630.
- [110] J. Liu, A. C. Potter, K. T. Law, and P. A. Lee. Zero-Bias Peaks in the Tunneling Conductance of Spin-Orbit-Coupled Superconducting Wires with and without Majorana End-States. *Physical Review Letters*, 109(26):267002, 2012.
- [111] H. Pan and S. Das Sarma. Physical mechanisms for zero-bias conductance peaks in Majorana nanowires. *Physical Review Research*, 2(1):013377, 2020.
- [112] G. Kells, D. Meidan, and P. W. Brouwer. Near-zero-energy end states in topologically trivial spin-orbit coupled superconducting nanowires with a smooth confinement. *Physical Review B*, 86(10):100503(R), 2012. ISSN 2469-9950.
- [113] E. Prada, P. San-Jose, and R. Aguado. Transport spectroscopy of NS nanowire junctions with Majorana fermions. *Physical Review B*, 86(18):180503(R), 2012. ISSN 2469-9950.
- [114] D. Rainis, L. Trifunovic, J. Klinovaja, and D. Loss. Towards a realistic transport modeling in a superconducting nanowire with Majorana fermions. *Physical Review B*, 87(2):024515, 2013.
- [115] J. Cayao, E. Prada, P. San-Jose, and R. Aguado. SNS junctions in nanowires with spin-orbit coupling: Role of confinement and helicity on the subgap spectrum. *Physical Review B*, 91(2):024514, 2015. ISSN 2469-9950.

- [116] P. San-Jose, J. Cayao, E. Prada, and R. Aguado. Majorana bound states from exceptional points in non-topological superconductors. *Scientific reports*, 6(1):1–13, 2016.
- [117] J. Chen, P. Yu, J. Stenger, M. Hocevar, D. Car, S. R. Plissard, E. P. A. M. Bakkers, T. D. Stanescu, and S. M. Frolov. Experimental phase diagram of zero-bias conductance peaks in superconductor/semiconductor nanowire devices. *Science advances*, 3(9):e1701476, 2017.
- [118] C.-X. Liu, J. D. Sau, T. D. Stanescu, and S. Das Sarma. Andreev bound states versus Majorana bound states in quantum dot-nanowire-superconductor hybrid structures: Trivial versus topological zero-bias conductance peaks. *Physical Review B*, 96(7):075161, 2017.
- [119] F. Peñaranda, R. Aguado, P. San-Jose, and E. Prada. Quantifying wave-function overlaps in inhomogeneous majorana nanowires. *Physical Review B*, 98(23):235406, 2018.
- [120] J. Avila, F. Peñaranda, E. Prada, P. San-Jose, and R. Aguado. Non-hermitian topology as a unifying framework for the andreev versus majorana states controversy. *Communications Physics*, 2(1):1–8, 2019.
- [121] C.-K. Chiu and S. Das Sarma. Fractional Josephson effect with and without Majorana zero modes. *Physical Review B*, 99(3):035312, 2019.
- [122] J. Chen, B. D. Woods, P. Yu, M. Hocevar, D. Car, S. R. Plissard, Bakkers, E. P. A. M., T. D. Stanescu, and S. M. Frolov. Ubiquitous Non-Majorana Zero-Bias Conductance Peaks in Nanowire Devices. *Physical review letters*, 123(10):107703, 2019.
- [123] B. D. Woods, J. Chen, S. M. Frolov, and T. D. Stanescu. Zero-energy pinning of topologically trivial bound states in multiband semiconductor-superconductor nanowires. *Physical Review B*, 100(12):125407, 2019.
- [124] A. Vuik, B. Nijholt, A. Akhmerov, and M. Wimmer. Reproducing topological properties with quasi-Majorana states. *SciPost Physics*, 7(5):061, 2019. ISSN 2542-4653.
- [125] O. Dmytruk, D. Loss, and J. Klinovaja. Pinning of andreev bound states to zero energy in two-dimensional superconductor-semiconductor rashba heterostructures. *Physical Review B*, 102(24):245431, 2020.
- [126] E. Prada, P. San-Jose, M. W. A. de Moor, A. Geresdi, E. J. H. Lee, J. Klinovaja, D. Loss, J. Nygård, R. Aguado, and L. P. Kouwenhoven. From andreev to majorana bound states in hybrid superconductor–semiconductor nanowires. *Nature Reviews Physics*, 2(10):575–594, 2020.
- [127] M. Valentini, F. Peñaranda, A. Hofmann, M. Brauns, R. Hauschild, P. Krogstrup, P. San-Jose, E. Prada, R. Aguado, and G. Katsaros. Nontopological zero-bias peaks in full-shell nanowires induced by flux-tunable andreev states. *Science*, 373(6550):82–88, 2021.
- [128] H. Zhang, M. W. A. de Moor, J. D. S. Bommer, Di Xu, G. Wang, N. van Loo, C.-X. Liu, S. Gazibegovic, J. A. Logan, D. Car, Veld, R. L. M. Op het, P. J. van Veldhoven, S. Koelling, M. A. Verheijen, M. Pendharkar, D. J. Pennachio, B. Shojaei, J. S. Lee, C. J. Palmstrøm, E. P. A. M. Bakkers, S. Das Sarma, and L. P. Kouwenhoven. Large zero-bias peaks in InSb-Al hybrid semiconductor-superconductor nanowire devices.
- [129] L. Fu. Electron teleportation via Majorana bound states in a mesoscopic superconductor. *Physical review letters*, 104(5):056402, 2010.

- [130] A. M. Whiticar, A. Fornieri, E. C. T. O’Farrell, A. C. C. Drachmann, T. Wang, C. Thomas, S. Gronin, R. Kallaher, G. C. Gardner, M. J. Manfra, C. M. Marcus, and F. Nichele. Coherent transport through a Majorana island in an Aharonov-Bohm interferometer. *Nature Communications*, 11(1):3212, 2020. ISSN 2041-1723.
- [131] M. Thamm and B. Rosenow. Transmission amplitude through a Coulomb blockaded Majorana wire. *Physical Review Research*, 3(2):023221, 2021.
- [132] M. Hell, K. Flensberg, and M. Leijnse. Distinguishing Majorana bound states from localized Andreev bound states by interferometry. *Physical Review B*, 97(16):161401(R), 2018. ISSN 2469-9950.
- [133] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. *Reviews of modern physics*, 81(2):865, 2009.
- [134] R. Jozsa. Entanglement and quantum computation. *arXiv preprint quant-ph/9707034*, 1997.
- [135] P. Calabrese and J. Cardy. Time Dependence of Correlation Functions Following a Quantum Quench. *Phys. Rev. Lett.*, 96(13):136801, 2006.
- [136] P. Calabrese and J. Cardy. Evolution of entanglement entropy in one-dimensional systems. *J. Stat. Mech.: Theor. Exp.*, 2005(04):P04010, 2005.
- [137] V. Alba and P. Calabrese. Entanglement and thermodynamics after a quantum quench in integrable systems. *Proc. Nat. Acad. Sci.*, 114(30):7947, 2017.
- [138] M. Srednicki. Chaos and quantum thermalization. *Phys. Rev. E*, 50(2):888, 1994.
- [139] M. Rigol, V. Dunjko, and M. Olshanii. Thermalization and its mechanism for generic isolated quantum systems. *Nature*, 452(7189):854, 2008.
- [140] A. Polkovnikov, K. Sengupta, A. Silva, and M. Vengalattore. Colloquium: Nonequilibrium dynamics of closed interacting quantum systems. *Rev. Mod. Phys.*, 83(3):863, 2011.
- [141] L. D’Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol. From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics. *Adv. in Phys.*, 65(3):239, 2016.
- [142] P. Zanardi. Quantum entanglement in fermionic lattices. *Phys. Rev. A*, 65(4):042101, 2002.
- [143] Y. Shi. Quantum entanglement of identical particles. *Phys. Rev. A*, 67(2):024301, 2003.
- [144] O. S. Zozulya, M. Haque, and K. Schoutens. Particle partitioning entanglement in itinerant many-particle systems. *Phys. Rev. A*, 78(4):042326, 2008.
- [145] M. Haque, O. S. Zozulya, and K. Schoutens. Entanglement between particle partitions in itinerant many-particle states. *J. Phys. A: Math. Theor.*, 42(50):504012, 2009.
- [146] M. Haque, O. Zozulya, and K. Schoutens. Entanglement Entropy in Fermionic Laughlin States. *Phys. Rev. Lett.*, 98(6):060401, 2007.

- [147] O. S. Zozulya, M. Haque, K. Schoutens, and E. H. Rezayi. Bipartite entanglement entropy in fractional quantum Hall states. *Phys. Rev. B*, 76(12):125310, 2007.
- [148] R. Santachiara, F. Stauffer, and D. C. Cabra. Entanglement properties and momentum distributions of hard-core anyons on a ring. *J. Stat. Mech.: Theor. Exp.*, 2007(05):05003, 2007.
- [149] H. Katsura and Y. Hatsuda. Entanglement entropy in the Calogero–Sutherland model. *J. Phys. A: Math. Theor.*, 40(46):13931, 2007.
- [150] C. M. Herdman, P. N. Roy, R. G. Melko, and A. Del Maestro. Particle entanglement in continuum many-body systems via quantum Monte Carlo. *Phys. Rev. B*, 89(14):140501(R), 2014.
- [151] C. M. Herdman, Stephen Inglis, P. N. Roy, R. G. Melko, and A. Del Maestro. Path-integral Monte Carlo method for Rényi entanglement entropies. *Phys. Rev. E*, 90(1):013308, 2014.
- [152] C. M. Herdman and A. Del Maestro. Particle partition entanglement of bosonic Luttinger liquids. *Phys. Rev. B*, 91(18):184507, 2015.
- [153] H. Barghathi, E. Casiano-Diaz, and A. Del Maestro. Particle partition entanglement of one dimensional spinless fermions. *J. Stat. Mech.: Theor. Exp.*, 2017(8):083108, 2017.
- [154] L. Rammelmüller, W. J. Porter, J. Braun, and J. E. Drut. Evolution from few- to many-body physics in one-dimensional fermi systems: One- and two-body density matrices and particle-partition entanglement. *Phys. Rev. A*, 96(3):033635, September 2017.
- [155] F. Iemini, T. O. Maciel, and R. O. Vianna. Entanglement of indistinguishable particles as a probe for quantum phase transitions in the extended Hubbard model. *Phys. Rev. B*, 92(7):075423, 2015.
- [156] D. L. B. Ferreira, T. O. Maciel, R. O. Vianna, and F. Iemini. Quantum correlations, entanglement spectrum, and coherence of the two-particle reduced density matrix in the extended Hubbard model. *Phys. Rev. B*, 105(11):115145, 2022.
- [157] S. Pu, A. C. Balram, and Z. Papić. Local density of states and particle entanglement in topological quantum fluids, 2022.
- [158] N. Killoran, M. Cramer, and M. B. Plenio. Extracting Entanglement from Identical Particles. *Phys. Rev. Lett.*, 112(15):150501, 2014.
- [159] A. Polkovnikov, E. Altman, and E. Demler. Interference between independent fluctuating condensates. *Proc. Nat. Acad. Sci.*, 103(16):6125, 2006.
- [160] V. Gritsev, E. Altman, E. Demler, and A. Polkovnikov. Full quantum distribution of contrast in interference experiments between interacting one-dimensional Bose liquids. *Nature Phys.*, 2(10):705, 2006.
- [161] A. Del Maestro, H. Barghathi, and B. Rosenow. Equivalence of spatial and particle entanglement growth after a quantum quench. *Phys. Rev. B*, 104(19):195101, 2021.
- [162] S. Eggert. One-dimensional quantum wires: A pedestrian approach to bosonization, 2009.

- [163] A. Iucci and M. A. Cazalilla. Quantum quench dynamics of the Luttinger model. *Phys. Rev. A*, 80(6):063619, 2009.
- [164] M. A. Cazalilla. Effect of Suddenly Turning on Interactions in the Luttinger Model. *Phys. Rev. Lett.*, 97(15):156403, 2006.
- [165] O. Bohigas, R. U. Haq, and A. Pandey. Nuclear data for science and technology. *Reidel, Dordrecht*, page 809, 1983.
- [166] C. W. J. Beenakker. Random-matrix theory of quantum transport. *Reviews of modern physics*, 69(3):731, 1997.
- [167] C. W. J. Beenakker. Random-matrix theory of majorana fermions and topological superconductors. *Reviews of Modern Physics*, 87(3):1037, 2015.
- [168] K. B. Efetov and A. I. Larkin. Kinetics of a quantum particle in long metallic wires. *Sov. Phys. JETP*, 58:444–451, 1983.
- [169] O. N. Dorokhov. Transmission coefficient and the localization length of an electron in n bound disordered chains. *Soviet Journal of Experimental and Theoretical Physics Letters*, 36:318, 1982.
- [170] P. A. Mello, P. Pereyra, and N. Kumar. Macroscopic approach to multichannel disordered conductors. *Annals of Physics*, 181(2):290–317, 1988.
- [171] J. P. Keating and F. Mezzadri. Random matrix theory and entanglement in quantum spin chains. *Communications in mathematical physics*, 252(1):543–579, 2004.
- [172] J. P. Keating and F. Mezzadri. Entanglement in quantum spin chains, symmetry classes of random matrices, and conformal field theory. *Physical review letters*, 94(5):050501, 2005.
- [173] M. Potters, J.-P. Bouchaud, and L. Laloux. Financial applications of random matrix theory: Old laces and new pieces. *arXiv preprint physics/0507111*, 2005.
- [174] A. M. Tulino, S. Verdú, et al. Random matrix theory and wireless communications. *Foundations and Trends® in Communications and Information Theory*, 1(1):1–182, 2004.
- [175] A. Edelman and N. R. Rao. Random matrix theory. *Acta numerica*, 14:233–297, 2005.
- [176] F. Luo, Y. Yang, J. Zhong, H. Gao, L. Khan, D. K. Thompson, and J. Zhou. Constructing gene co-expression networks and predicting functions of unknown genes by random matrix theory. *BMC Bioinformatics*, 8(1):299, 2007.
- [177] Y. Deng, Y.-H. Jiang, Y. Yang, Z. He, F. Luo, and J. Zhou. Molecular ecological network analyses. *BMC Bioinformatics*, 13(1):113, 2012.
- [178] D. Paul and A. Aue. Random matrix theory in statistics: A review. *Journal of Statistical Planning and Inference*, 150:1–29, 2014.
- [179] M. Z. Rácz and J. Richey. A smooth transition from wishart to goe. *Journal of Theoretical Probability*, 32(2):898–906, 2019.
- [180] B. Dietz and K. Zyczkowski. Level-spacing distributions beyond the wigner surmise. *Zeitschrift für Physik B Condensed Matter*, 84(1):157–158, 1991.

- [181] P. J. Forrester and N. S. Witte. Exact wigner surmise type evaluation of the spacing distribution in the bulk of the scaled random matrix ensembles. *Letters in Mathematical Physics*, 53(3):195–200, 2000.
- [182] M. L. Mehta. *Random matrices*, volume v. 142 of *Pure and applied mathematics*. Elsevier/Academic Press, Amsterdam and San Diego, CA, 3rd ed. edition, 2004. ISBN 0120884097.
- [183] G. Livan, M. Novaes, and P. Vivo. Introduction to random matrices theory and practice. *Monograph Award*, page 63, 2018.
- [184] V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1(4):457–483, 1967. ISSN 0025-5734.
- [185] A. M. Sengupta and P. P. Mitra. Distributions of singular values for some random matrices. *Physical Review E*, 60(3):3389, 1999.
- [186] M. Staats. Random Matrix Analysis Of Deep Neural Networks. Master’s thesis, Leipzig University, 2022.
- [187] C. E. Porter and R. G. Thomas. Fluctuations of nuclear reaction widths. *Physical Review*, 104(2):483, 1956.
- [188] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.
- [189] X. Glorot and Y. Bengio. Understanding the difficulty of training deep feedforward neural networks. *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, pages 249–256, 2010. ISSN 1938-7228.
- [190] A. Krizhevsky, G. Hinton, et al. Learning multiple layers of features from tiny images. *Tech Report*, 2009.
- [191] F. Rosenblatt. *The perceptron, a perceiving and recognizing automaton Project Para*. Cornell Aeronautical Laboratory, 1957.
- [192] F. Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958.
- [193] J. Schmidhuber. Deep learning in neural networks: An overview. *Neural networks*, 61: 85–117, 2015.
- [194] A. G. Ivakhnenko and V. G. Lapa. Cybernetic predicting devices. ccm information corporation. *First working Deep Learners with many layers, learning internal representations*, 1965.
- [195] A. G. Ivakhnenko and V. G. Lapa. *Cybernetics and forecasting techniques*, volume 8. American Elsevier Publishing Company, 1967.

- [196] A. G. Ivakhnenko. The group method of data handling, a rival of the method of stochastic approximation. *Soviet Automatic Control*, 13(3):43–55, 1968.
- [197] A. G. Ivakhnenko. Polynomial theory of complex systems. *IEEE transactions on Systems, Man, and Cybernetics*, (4):364–378, 1971.
- [198] P. Werbos. Beyond regression:" new tools for prediction and analysis in the behavioral sciences. *Ph. D. dissertation, Harvard University*, 1974.
- [199] D. B. Parker. Learning logic technical report tr-47. *Center of Computational Research in Economics and Management Science, Massachusetts Institute of Technology, Cambridge, MA*, 1985.
- [200] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *nature*, 323(6088):533–536, 1986.
- [201] S. Geman, E. Bienenstock, and R. Doursat. Neural Networks and the Bias/Variance Dilemma. *Neural Computation*, 4(1):1–58, 1992. ISSN 0899-7667.
- [202] P. Sermanet, K. Kavukcuoglu, S. Chintala, and Y. LeCun. Pedestrian detection with unsupervised multi-stage feature learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3626–3633, 2013.
- [203] I. Sutskever, O. Vinyals, and Q. V. Le. Sequence to sequence learning with neural networks. *Advances in neural information processing systems*, 27, 2014.
- [204] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998. ISSN 0018-9219.
- [205] Y. LeCun, P. Haffner, L. Bottou, and Y. Bengio. Object Recognition with Gradient-Based Learning. In *Shape, Contour and Grouping in Computer Vision*, pages 319–345. Springer, Berlin, Heidelberg, 1999.
- [206] R. Pascanu, Y. N. Dauphin, S. Ganguli, and Y. Bengio. On the saddle point problem for non-convex optimization. *preprint arXiv:1405.4604*, 2014.
- [207] Y. Dauphin, R. Pascanu, C. Gulcehre, K. Cho, S. Ganguli, and Y. Bengio. Identifying and attacking the saddle point problem in high-dimensional non-convex optimization. *preprint arXiv:1406.2572*, 2014.
- [208] I. J. Goodfellow, O. Vinyals, and A. M. Saxe. Qualitatively characterizing neural network optimization problems. *preprint arXiv:1412.6544*, 2015.
- [209] B. Neyshabur, S. Bhojanapalli, D. McAllester, and N. Srebro. Exploring Generalization in Deep Learning. *preprint arXiv:1706.08947*, 2017.
- [210] D. Soudry and E. Hoffer. Exponentially vanishing sub-optimal local minima in multilayer neural networks. *preprint arXiv:1702.05777*, 2017.
- [211] S. Ma, R. Bassily, and M. Belkin. The Power of Interpolation: Understanding the Effectiveness of SGD in Modern Over-parametrized Learning. *International Conference on Machine Learning*, pages 3325–3334, 2018. ISSN 2640-3498.

- [212] K. Kawaguchi, L. P. Kaelbling, and Y. Bengio. Generalization in deep learning. *arXiv preprint arXiv:1710.05468*, 2017.
- [213] K. Simonyan and A. Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. *preprint arXiv:1409.1556*, 2014.
- [214] X. Zhai, A. Kolesnikov, N. Houlsby, and L. Beyer. Scaling Vision Transformers. *preprint arXiv:2106.04560*, 2021.
- [215] M. Belkin, D. Hsu, S. Ma, and S. Mandal. Reconciling modern machine-learning practice and the classical bias-variance trade-off. *Proceedings of the National Academy of Sciences*, 116(32):15849–15854, 2019. ISSN 1091-6490.
- [216] P. Nakkiran, G. Kaplun, Y. Bansal, T. Yang, B. Barak, and I. Sutskever. Deep double descent: Where bigger models and more data hurt. *preprint arXiv:1912.02292*, 2019.
- [217] M. Belkin, D. Hsu, and J. Xu. Two Models of Double Descent for Weak Features. *SIAM Journal on Mathematics of Data Science*, 2(4):1167–1180, 2020.
- [218] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, et al. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. *arXiv preprint arXiv:1712.01815*, 2017.
- [219] O. Vinyals, I. Babuschkin, W. M. Czarnecki, M. Mathieu, A. Dudzik, J. Chung, D. H. Choi, R. Powell, T. Ewalds, P. Georgiev, et al. Grandmaster level in starcraft ii using multi-agent reinforcement learning. *Nature*, 575(7782):350–354, 2019.
- [220] J. Lee, Y. Bahri, R. Novak, S. S. Schoenholz, J. Pennington, and J. Sohl-Dickstein. Deep neural networks as gaussian processes. *preprint arXiv:1711.00165*, 2017.
- [221] A. G de G Matthews, M. Rowland, J. Hron, R. E. Turner, and Z. Ghahramani. Gaussian process behaviour in wide deep neural networks. *preprint arXiv:1804.11271*, 2018.
- [222] O. Ben-David and Z. Ringel. The role of a layer in deep neural networks: a gaussian process perspective. *preprint arXiv:1902.02354*, 2019.
- [223] B. Hanin. Random neural networks in the infinite width limit as gaussian processes. *preprint arXiv:2107.01562*, 2021.
- [224] Z. Allen-Zhu, Y. Li, and Z. Song. A convergence theory for deep learning via over-parameterization. In *International Conference on Machine Learning*, pages 242–252. PMLR, 2019.
- [225] J. Lee, L. Xiao, S. Schoenholz, Y. Bahri, R. Novak, J. Sohl-Dickstein, and J. Pennington. Wide neural networks of any depth evolve as linear models under gradient descent. *Advances in neural information processing systems*, 32:8572–8583, 2019.
- [226] G. Naveh, O. Ben-David, H. Sompolinsky, and Z. Ringel. Predicting the outputs of finite networks trained with noisy gradients. *preprint arXiv:2004.01190*, 2020.
- [227] B. Frénay and M. Verleysen. Classification in the presence of label noise: a survey. *IEEE transactions on neural networks and learning systems*, 25(5):845–869, 2013.

- [228] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 2016.
- [229] A. K. Lampinen and S. Ganguli. An analytic theory of generalization dynamics and transfer learning in deep linear networks. *preprint arXiv:1809.10374*, 2018.
- [230] J. Pennington, S. Schoenholz, and S. Ganguli. The emergence of spectral universality in deep networks. In Amos Storkey and Fernando Perez-Cruz, editors, *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, volume 84 of *Proceedings of Machine Learning Research*, pages 1924–1932. PMLR, 09–11 Apr 2018.
- [231] J. Pennington, S. Schoenholz, and S. Ganguli. Resurrecting the sigmoid in deep learning through dynamical isometry: theory and practice. *Advances in neural information processing systems*, 30, 2017.
- [232] N. P. Baskerville, D. Granzio, and J. P. Keating. Applicability of Random Matrix Theory in Deep Learning. *preprint arXiv:2102.06740*.
- [233] D. Granzio. Beyond random matrix theory for deep networks. *preprint arXiv:2006.07721*, 2020.
- [234] D. Granzio, S. Zohren, and S. Roberts. Learning rates as a function of batch size: A random matrix theory approach to neural network training. *arXiv preprint arXiv:2006.09092*, 2020.
- [235] C. H. Martin and M. W. Mahoney. Implicit self-regularization in deep neural networks: Evidence from random matrix theory and implications for learning. *Journal of Machine Learning Research*, 22(165):1–73, 2021.
- [236] P. Cizeau and J. P. Bouchaud. Theory of lévy matrices. *Phys. Rev. E*, 50:1810–1822, Sep 1994.
- [237] E. Tarquini, G. Biroli, and M. Tarzia. Level statistics and localization transitions of lévy matrices. *Phys. Rev. Lett.*, 116:010601, Jan 2016.
- [238] C. H. Martin, T. (S.) Peng, and M. W. Mahoney. Predicting trends in the quality of state-of-the-art neural networks without access to training or testing data. *Nature Communications*, 12(1):4122, 2021.
- [239] N. Hansen, A. Ostermeier, and A. Gawelczyk. On the adaptation of arbitrary normal mutation distributions in evolution strategies: The generating set adaptation. In *ICGA*, pages 57–64. Citeseer, 1995.
- [240] N. Hansen and A. Ostermeier. Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. In *Proceedings of IEEE international conference on evolutionary computation*, pages 312–317. IEEE, 1996.
- [241] N. Hansen and A. Ostermeier. Convergence properties of evolution strategies with the derandomized covariance matrix adaptation: The (/i,)-es. *Eufit*, 97:650–654, 1997.
- [242] N. Hansen and A. Ostermeier. Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation*, 9(2):159–195, 2001. ISSN 1063-6560.

- [243] N. Hansen, S. D. Müller, and P. Koumoutsakos. Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evolutionary Computation*, 11(1):1–18, 2003. ISSN 1063-6560.
- [244] N. Hansen and S. Kern. Evaluating the cma evolution strategy on multimodal test functions. In *International Conference on Parallel Problem Solving from Nature*, pages 282–291. Springer, 2004.
- [245] A. Auger and N. Hansen. A restart cma evolution strategy with increasing population size. In *2005 IEEE congress on evolutionary computation*, volume 2, pages 1769–1776. IEEE, 2005.
- [246] N. Hansen. The CMA Evolution Strategy: A Comparing Review. In Jose A. Lozano, Endika Bengoetxea, Iñaki Inza, and Pedro Larrañaga, editors, *Towards a New Evolutionary Computation*, Studies in Fuzziness and Soft Computing, pages 75–102. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, 2006. ISBN 978-3-540-32494-2.
- [247] G. A. Jastrebski and D. V. Arnold. Improving evolution strategies through active covariance matrix adaptation. In *2006 IEEE international conference on evolutionary computation*, pages 2814–2821. IEEE, 2006.
- [248] C. Igel, T. Suttorp, and N. Hansen. A computational efficient covariance matrix update and a (1+ 1)-cma for evolution strategies. In *Proceedings of the 8th annual conference on Genetic and evolutionary computation*, pages 453–460, 2006.
- [249] J. A. Lozano, E. Bengoetxea, I. Inza, and P. Larrañaga, editors. *Towards a New Evolutionary Computation: Advances in the Estimation of Distribution Algorithms*, volume 192 of *Studies in Fuzziness and Soft Computing*. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, 2006. ISBN 978-3-540-32494-2.
- [250] I. Loshchilov and F. Hutter. CMA-ES for Hyperparameter Optimization of Deep Neural Networks. *arxiv:1604.07269*, 2016.
- [251] M. Willjuice Iruthayarajan and S. Baskar. Covariance matrix adaptation evolution strategy based design of centralized PID controller. *Expert Systems with Applications*, 37(8): 5775–5781, 2010. ISSN 0957-4174.
- [252] I. Loshchilov, M. Schoenauer, and M. Sebag. Bi-population cma-es algorithms with surrogate models and line searches. In *Proceedings of the 15th annual conference companion on Genetic and evolutionary computation*, pages 1177–1184, 2013.
- [253] Nikolaus Hansen. The cma evolution strategy: A tutorial. *arXiv preprint arXiv:1604.00772*, 2016.
- [254] N. Hansen, Y. Akimoto, and P. Baudis. CMA-ES/pycma on Github. Zenodo, 10.5281/zenodo.2559634, February 2019.
- [255] M. Sato and Y. Ando. Topological superconductors: a review. *Reports on Progress in Physics*, 80(7):076501, 2017. ISSN 0034-4885.
- [256] D. A. Ivanov. Non-abelian statistics of half-quantum vortices in p-wave superconductors. *Physical Review Letters*, 86(2):268, 2001.

- [257] P. Adrien M. Dirac. The quantum theory of the electron. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 117 (778):610–624, 1928.
- [258] P. A. M. Dirac. The quantum theory of the electron. part ii. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 118(779):351–361, 1928.
- [259] T. Pashby. Dirac’s prediction of the positron: A case study for the current realism debate. *Perspectives on Science*, 20(4):440–475, 2012.
- [260] J. J. Sakurai and J. Napolitano. *Modern Quantum Mechanics*. Addison-Wesley, 2011. ISBN 9780805382914.
- [261] D. F. Moyer. Evaluations of dirac’s electron, 1928–1932. *American journal of physics*, 49 (11):1055–1062, 1981.
- [262] C. D. Anderson. The positive electron. *Physical review*, 43(6):491, 1933.
- [263] E. Majorana. Teoria simmetrica dell’elettrone e del positrone. *Il Nuovo Cimento (1924-1942)*, 14(4):171, 1937. ISSN 1827-6121.
- [264] C. W. J. Beenakker. Search for Majorana Fermions in Superconductors. *Annual Review of Condensed Matter Physics*, 4(1):113–136, 2013. ISSN 1947-5454.
- [265] B. Zocher. *Signatures of topological superconductivity in semiconductors with strong spin-orbit coupling*. PhD thesis, Leipzig University, 2013.
- [266] G. Moore and N. Read. Nonabelions in the fractional quantum hall effect. *Nuclear Physics B*, 360(2):362–396, 1991. ISSN 0550-3213.
- [267] N. Read and D. Green. Paired states of fermions in two dimensions with breaking of parity and time-reversal symmetries and the fractional quantum Hall effect. *Physical Review B*, 61(15):10267, 2000. ISSN 2469-9950.
- [268] K. Sengupta, I. Žutić, H.-J. Kwon, V. M. Yakovenko, and S. Das Sarma. Midgap edge states and pairing symmetry of quasi-one-dimensional organic superconductors. *Physical Review B*, 63(14):144531, 2001. ISSN 2469-9950.
- [269] A. Stern, F. von Oppen, and E. Mariani. Geometric phases and quantum entanglement as building blocks for non-abelian quasiparticle statistics. *Physical Review B*, 70(20):205338, 2004. ISSN 2469-9950.
- [270] M. Stone and S.-B. Chung. Fusion rules and vortices in px+ipy superconductors. *Physical Review B*, 73(1):014505, 2006. ISSN 2469-9950.
- [271] S. Das Sarma, C. Nayak, and S. Tewari. Proposal to stabilize and detect half-quantum vortices in strontium ruthenate thin films: Non-abelian braiding statistics of vortices in a px+ipy superconductor. *Physical Review B*, 73(22):220502, 2006. ISSN 2469-9950.
- [272] A. R. Akhmerov, J. Nilsson, and C. W. J. Beenakker. Electrically detected interferometry of majorana fermions in a topological insulator. *Physical Review Letters*, 102(21):216404, 2009.

- [273] V. Gurarie, L. Radzihovsky, and A. V. Andreev. Quantum phase transitions across a p-wave feshbach resonance. *Physical Review Letters*, 94(23):230403, 2005.
- [274] S. Tewari, S. Das Sarma, C. Nayak, C. Zhang, and P. Zoller. Quantum computation using vortices and majorana zero modes of a px+ipy superfluid of fermionic cold atoms. *Physical Review Letters*, 98(1):010506, 2007.
- [275] N. Goldman, I. Satija, P. Nikolic, A. Bermudez, M. A. Martin-Delgado, M. Lewenstein, and I. B. Spielman. Realistic time-reversal invariant topological insulators with neutral atoms. *Physical Review Letters*, 105(25):255302, 2010.
- [276] J. D. Sau, R. Sensarma, S. Powell, I. B. Spielman, and S. Das Sarma. Chiral rashba spin textures in ultracold fermi gases. *Physical Review B*, 83(14):140510, 2011. ISSN 2469-9950.
- [277] J. D. Sau and S. Das Sarma. Realizing a robust practical majorana chain in a quantum-dot-superconductor linear array. *Nature Communications*, 3(1):1–6, 2012. ISSN 2041-1723.
- [278] I. C. Fulga, A. Haim, A. R. Akhmerov, and Y. Oreg. Adaptive tuning of majorana fermions in a quantum dot chain. *New Journal of Physics*, 15(4):045020, 2013. ISSN 1367-2630.
- [279] T.-P. Choy, J. M. Edge, A. R. Akhmerov, and C. W. J. Beenakker. Majorana fermions emerging from magnetic nanoparticles on a superconductor without spin-orbit coupling. *Physical Review B*, 84(19):195442, 2011. ISSN 2469-9950.
- [280] I. Martin and A. F. Morpurgo. Majorana fermions in superconducting helical magnets. *Physical Review B*, 85(14):144505, 2012. ISSN 2469-9950.
- [281] M. Kjaergaard, K. Wölms, and K. Flensberg. Majorana fermions in superconducting nanowires without spin-orbit coupling. *Physical Review B*, 85(2):020503, 2012.
- [282] S. Nadj-Perge, I. K. Drozdov, B. A. Bernevig, and A. Yazdani. Proposal for realizing majorana fermions in chains of magnetic atoms on a superconductor. *Physical Review B*, 88(2):020407, 2013. ISSN 2469-9950.
- [283] C. Chamon, M. O. Goerbig, R. Moessner, and L. F. Cugliandolo. *Topological Aspects of Condensed Matter Physics: Lecture Notes of the Les Houches Summer School: Volume 103, August 2014*. OUP Oxford, 2017. ISBN 9780191088797.
- [284] R. Balian and N. R. Werthamer. Superconductivity with pairs in a relative p wave. *Physical Review*, 131(4):1553–1564, 1963.
- [285] K. Ishida, H. Mukuda, Y. Kitaoka, K. Asayama, Z. Q. Mao, Y. Mori, and Y. Maeno. Spin-triplet superconductivity in sr2ruo4 identified by ^{17}O knight shift. *Nature*, 396(6712):658–660, 1998. ISSN 1476-4687.
- [286] C. Kallin and J. Berlinsky. Chiral superconductors. *Reports on Progress in Physics*, 79(5):054502, 2016.
- [287] Y. Maeno, S. Kittaka, T. Nomura, S. Yonezawa, and K. Ishida. Evaluation of spin-triplet superconductivity in sr2ruo4. *Journal of the Physical Society of Japan*, 81(1):011009, 2011.
- [288] H. Meissner. Superconductivity of contacts with interposed barriers. *Physical Review*, 117(3):672–680, 1960.

- [289] L. N. Cooper. Superconductivity in the neighborhood of metallic contacts. *Physical review letters*, 6(12):689–690, 1961. ISSN 0031-9007.
- [290] Y.-J. Doh, J. A. van Dam, A. L. Roest, Bakkers, E. P. A. M., L. P. Kouwenhoven, and S. de Franceschi. Tunable supercurrent through semiconductor nanowires. *Science*, 309(5732):272–275, 2005. ISSN 1095-9203.
- [291] L. Fu and C. L. Kane. Superconducting proximity effect and majorana fermions at the surface of a topological insulator. *Physical review letters*, 100(9):096407, 2008.
- [292] B. J. Wees and H. Takayanagi. The superconducting proximity effect in semiconductor-superconductor systems: Ballistic transport, low dimensionality and sample specific properties. In Lydia L. Sohn, Leo P. Kouwenhoven, and Gerd Schön, editors, *Mesoscopic Electron Transport*, pages 469–501. Springer Netherlands, Dordrecht, 2010. ISBN 978-90-481-4906-3.
- [293] F. Harper, A. Pushp, and R. Roy. Majorana braiding in realistic nanowire Y-junctions and tuning forks. *Physical Review Research*, 1(3):033207, 2019.
- [294] M. Freedman, A. Kitaev, M. Larsen, and Z. Wang. Topological quantum computation. *Bulletin of the American Mathematical Society*, 40(1):31–38, 2003.
- [295] K. Flensberg. Non-abelian operations on majorana fermions via single-charge control. *Physical review letters*, 106(9):090503, 2011.
- [296] S. Bravyi. Universal quantum computation with the 5/2 fractional quantum hall state. *Phys. Rev. A*, 73:042313, Apr 2006.
- [297] F. Hassler, A. R. Akhmerov, C. Y. Hou, and C. W. J. Beenakker. Anyonic interferometry without anyons: how a flux qubit can read out a topological qubit. *New Journal of Physics*, 12(12):125002, 2010.
- [298] K. T. Law, Patrick A. Lee, and T. K. Ng. Majorana fermion induced resonant Andreev reflection. *Physical Review Letters*, 103(23):237001, 2009.
- [299] M. T. Deng, S. Vaitiekėnas, E. B. Hansen, J. Danon, M. Leijnse, K. Flensberg, J. Nygård, P. Krogstrup, and C. M. Marcus. Majorana bound state in a coupled quantum-dot hybrid-nanowire system. *Science (New York, N.Y.)*, 354(6319):1557–1562, 2016.
- [300] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. Das Sarma. Non-abelian anyons and topological quantum computation. *Reviews of Modern Physics*, 80(3):1083, 2008.
- [301] A. Stern. Anyons and the quantum Hall effect—A pedagogical review. *Annals of Physics*, 323(1):204–249, 2008. ISSN 0003-4916.
- [302] D. I. Pikulin, B. van Heck, T. Karzig, E. A. Martinez, B. Nijholt, T. Laeven, G. W. Winkler, J. D. Watson, S. Heedt, M. Temurhan, et al. Protocol to identify a topological superconducting phase in a three-terminal device. *arXiv preprint arXiv:2103.12217*, 2021.
- [303] M. Thamm and B. Rosenow. Machine learning optimization of majorana hybrid nanowires. *arXiv preprint arXiv:2208.02182*, 2022.

- [304] C. Drukier, H.-G. Zirnstein, B. Rosenow, A. Stern, and Y. Oreg. Evolution of the transmission phase through a Coulomb-blockaded Majorana wire. *Physical Review B*, 98(16):161401(R), 2018. ISSN 2469-9950.
- [305] M. Thamm. Scattering Matrix Analysis of the Transmission Phase Through a Coulomb-Blockaded Nanowire. Master’s thesis, Universität Leipzig, Germany, 2018.
- [306] R. Islam, R. Ma, P. M. Preiss, M. Eric Tai, A. Lukin, M. Rispoli, and M. Greiner. Measuring entanglement entropy in a quantum many-body system. *Nature*, 528(7580):77–83, 2015.
- [307] J. Eisert, M. Cramer, and M. B. Plenio. Area laws for the entanglement entropy-a review. *arXiv preprint arXiv:0808.3773*, 2008.
- [308] A. Del Maestro, H. Barghathi, and B. Rosenow. Equivalence of spatial and particle entanglement growth after a quantum quench. *Physical Review B*, 104(19):195101, 2021.
- [309] B.-B. Wei. Linking phase transitions and quantum entanglement at arbitrary temperature. *Physical Review A*, 97(4):042115, 2018.
- [310] Y. Shi. Quantum entanglement in second-quantized condensed matter systems. *Journal of Physics A: Mathematical and General*, 37(26):6807, 2004.
- [311] A. J. Coleman. Structure of Fermion Density Matrices. *Rev. Mod. Phys.*, 35(3):668, 1963.
- [312] G. D. Mahan. *Many-Particle Physics*. Physics of Solids and Liquids. Springer US, 2013. ISBN 9781475757149.
- [313] I. Peschel and V. Eisler. Reduced density matrices and entanglement entropy in free lattice models. *Journal of physics a: mathematical and theoretical*, 42(50):504003, 2009.
- [314] S.-H. Lin, B. Sbierski, F. Dorfner, C. Karrasch, and F. Heidrich-Meisner. Many-body localization of spinless fermions with attractive interactions in one dimension. *SciPost Physics*, 4(1), January 2018.
- [315] M. Hopjan, F. Heidrich-Meisner, and V. Alba. Scaling properties of a spatial one-particle density-matrix entropy in many-body localized systems. *Phys. Rev. B*, 104(3):035129, 2021.
- [316] L.-A. Wu, M. S. Sarandy, and D. A. Lidar. Quantum phase transitions and bipartite entanglement. *Physical review letters*, 93(25):250404, 2004.
- [317] G. C. Ghirardi and L. Marinatto. General criterion for the entanglement of two indistinguishable particles. *Phys. Rev. A*, 70:012109, Jul 2004.
- [318] A. P. Balachandran, T. R. Govindarajan, A. R. de Queiroz, and A. F. Reyes-Lega. Entanglement and particle identity: A unifying approach. *Phys. Rev. Lett.*, 110:080503, Feb 2013.
- [319] M. C. Tichy, F. Mintert, and A. Buchleitner. Essential entanglement for atomic and molecular physics. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 44(19):192001, 2011.

- [320] F. Aikebaier, T. Ojanen, and J. L. Lado. Extracting electronic many-body correlations from local measurements with artificial neural networks. *arXiv preprint arXiv:2206.02388*, 2022.
- [321] J. G. Kirkwood and E. M. Boggs. The Radial Distribution Function in Liquids. *J. Chem. Phys.*, 10(6):394, 1942.
- [322] H. S. Green. *The Molecular Theory of Fluids*. Norh-Holland, Amsterdam, 1952.
- [323] J. Des Cloizeaux. A Soluble Fermi–Gas Model. Validity of Transformations of the Bogoliubov Type. *Journal of Mathematical Physics*, 7(12):2136–2144, 1966. ISSN 0022-2488.
- [324] C. N. Yang and C. P. Yang. One-Dimensional Chain of Anisotropic Spin-Spin Interactions. I. Proof of Bethe’s Hypothesis for Ground State in a Finite System. *Physical Review*, 150(1):321–327, 1966.
- [325] L. Gotta, L. Mazza, P. Simon, and G. Roux. Pairing in spinless fermions and spin chains with next-nearest neighbor interactions. *Physical Review Research*, 3(1):013114, 2021.
- [326] T. Giamarchi. *Quantum physics in one dimension*, volume 121 of *Oxford science publications*. Clarendon Press, Oxford, repr edition, 2010. ISBN 9780198525004.
- [327] M. Thamm, H. Radhakrishnan, H. Barghathi, B. Rosenow, and A. Del Maestro. One-particle entanglement for one-dimensional spinless fermions after an interaction quantum quench. *Physical Review B*, 106(16):165116, 2022.
- [328] M. Thamm, M. Staats, and B. Rosenow. Random matrix analysis of deep neural network weight matrices. *Phys. Rev. E*, 106:054124, Nov 2022.
- [329] M. Staats, M. Thamm, and B. Rosenow. Boundary between noise and information applied to filtering neural network weight matrices. *arXiv preprint arXiv:2206.03927*, 2022.
- [330] L. N. Cooper. Bound electron pairs in a degenerate fermi gas. *Physical Review*, 104(4):1189–1190, 1956.
- [331] S. Bose. Plancks gesetz und lichtquantenhypothese. *Zeitschrift für Physik*, 26(1):178–181, 1924.
- [332] A. Einstein. Quantentheorie des einatomigen idealen gases. *Koenigliche Preußische Akademie der Wissenschaften, Sitzungsberichte*, pages 261–267, 1924.
- [333] J. F. Annett. *Superconductivity, Superfluids and Condensates*. Oxford Univ. Press, Oxford, 9 edition, 2004. ISBN 978-0-198-50756-7.
- [334] D. Goldhaber-Gordon, H. Shtrikman, D. Mahalu, D. Abusch-Magder, U. Meirav, and M. A. Kastner. Kondo effect in a single-electron transistor. *Nature*, 391(6663):156–159, 1998. ISSN 1476-4687.
- [335] E. J. H. Lee, X. Jiang, R. Aguado, G. Katsaros, C. M. Lieber, and S. de Franceschi. Zero-bias anomaly in a nanowire quantum dot coupled to superconductors. *Physical review letters*, 109(18):186802, 2012.
- [336] M. Sigrist and K. Ueda. Phenomenological theory of unconventional superconductivity. *Reviews of Modern Physics*, 63(2):239–311, 1991.

- [337] J. G. Valatin. Comments on the theory of superconductivity. *Il Nuovo Cimento (1924-1942)*, 7(6):843–857, 1958. ISSN 1827-6121.
- [338] N. N. Bogoljubov. On a new method in the theory of superconductivity. *Il Nuovo Cimento (1924-1942)*, 7(6):794–805, 1958. ISSN 1827-6121.
- [339] A. Altland and M. R. Zirnbauer. Nonstandard symmetry classes in mesoscopic normal-superconducting hybrid structures. *Physical review. B, Condensed matter*, 55(2):1142–1161, 1997. ISSN 0163-1829.
- [340] A. P. Schnyder, S. Ryu, A. Furusaki, and A. W. W. Ludwig. Classification of topological insulators and superconductors in three spatial dimensions. *Physical Review B*, 78(19):195125, 2008. ISSN 2469-9950.
- [341] A. Kitaev, V. Lebedev, and M. Feigel'man. Periodic table for topological insulators and superconductors. *AIP Conference Proceedings*, 1134(1):22–30, 2009. ISSN 0094-243X.
- [342] C.-K. Chiu, J. C. Y. Teo, A. P. Schnyder, and S. Ryu. Classification of topological quantum matter with symmetries. *Reviews of Modern Physics*, 88(3):035005, 2016.
- [343] A. J. Heeger, S. Kivelson, J. R. Schrieffer, and W.-P. Su. Solitons in conducting polymers. *Reviews of Modern Physics*, 60(3):781–850, 1988.
- [344] A. Kitaev. Anyons in an exactly solved model and beyond. *Annals of Physics*, 321(1):2–111, 2006. ISSN 0003-4916.
- [345] A. Cayley. Sur les déterminants gauches. (suite du mémoire t. xxxii. p. 119). *Journal für die reine und angewandte Mathematik*, 38:93–96, 1849. ISSN 0075-4102.
- [346] P. Ghosh, J. D. Sau, S. Tewari, and S. Das Sarma. Non-abelian topological order in non-centrosymmetric superconductors with broken time-reversal symmetry. *Physical Review B*, 82(18):184525, 2010. ISSN 2469-9950.
- [347] S. Tewari and J. D. Sau. Topological invariants for spin-orbit coupled superconductor nanowires. *Physical review letters*, 109(15):150408, 2012. ISSN 0031-9007.
- [348] L. Fidkowski and A. Kitaev. Topological phases of fermions in one dimension. *Physical Review B*, 83(7):075103, 2011. ISSN 2469-9950.
- [349] Y. Tanaka, M. Sato, and N. Nagaosa. Symmetry and topology in superconductors –odd-frequency pairing and edge states–. *Journal of the Physical Society of Japan*, 81(1):011013, 2012. ISSN 0031-9015.
- [350] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. Theory of superconductivity. *Physical Review*, 108(5):1175–1204, 1957.
- [351] B. A. Bernevig and T. L. Hughes. *Topological Insulators and Topological Superconductors*. Princeton University Press, 2013. ISBN 9780691151755.
- [352] P. G. De Gennes. *Superconductivity Of Metals And Alloys*. Basic Books, 1994. ISBN 9780201408423.
- [353] L. N. Cooper, R. L. Mills, and A. M. Sessler. Possible superfluidity of a system of strongly interacting fermions. *Physical Review*, 114(6):1377–1382, 1959.

- [354] L. P. Pitaevskii. On the superfluidity of liquid he-3. *SOVIET PHYSICS JETP-USSR*, 10 (6):1267–1275, 1960.
- [355] K. A. Brueckner, To. Soda, P. W. Anderson, and P. Morel. Level structure of nuclear matter and liquid he3. *Physical Review*, 118(5):1442–1446, 1960.
- [356] V. J. Emery and A. M. Sessler. Possible phase transition in liquid he3. *Physical Review*, 119(1):43–49, 1960.
- [357] E. F. Talantsev, K. Iida, T. Ohmura, T. Matsumoto, W. P. Crump, N. M. Strickland, S. C. Wimbush, and H. Ikuta. p-wave superconductivity in iron-based superconductors. *Scientific Reports*, 9(1):14245, 2019. ISSN 2045-2322.
- [358] E. Hassinger, P. Bourgeois-Hope, H. Taniguchi, S. R. de Cotret, G. Grissonnanche, M. S. Anwar, Y. Maeno, N. Doiron-Leyraud, and Louis Taillefer. Vertical line nodes in the superconducting gap structure of sr2ruo4. *Physical Review X*, 7(1):011032, 2017. ISSN 2160-3308.
- [359] A. Di Bernardo, O. Millo, M. Barbone, H. Alpern, Y. Kalcheim, U. Sassi, A. K. Ott, D. de Fazio, D. Yoon, M. Amado, A. C. Ferrari, J. Linder, and J. W. A. Robinson. p-wave triggered superconductivity in single-layer graphene on an electron-doped oxide superconductor. *Nature communications*, 8(1):14024, 2017. ISSN 2041-1723.
- [360] O. Viyuela, L. Fu, and M. A. Martin-Delgado. Chiral topological superconductors enhanced by long-range interactions. *Physical review letters*, 120(1):017001, 2018. ISSN 0031-9007.
- [361] J. D. Sau, R. M. Lutchyn, S. Tewari, and S. Das Sarma. Generic new platform for topological quantum computation using semiconductor heterostructures. *Physical review letters*, 104(4):040502, 2010.
- [362] M. Yazawa, M. Koguchi, and K. Hiruma. Heteroepitaxial ultrafine wire-like growth of inas on gaas substrates. *Applied Physics Letters*, 58(10):1080–1082, 1991. ISSN 0003-6951.
- [363] S. Tsukamoto, Y. Nagamune, M. Nishioka, and Y. Arakawa. Fabrication of gaas quantum wires on epitaxially grown v grooves by metal-organic chemical-vapor deposition. *Journal of Applied Physics*, 71(1):533–535, 1992. ISSN 0021-8979.
- [364] S. Koshiha, H. Noge, H. Akiyama, T. Inoshita, Y. Nakamura, A. Shimizu, Y. Nagamune, M. Tsuchiya, H. Kano, H. Sakaki, and K. Wada. Formation of gaas ridge quantum wire structures by molecular beam epitaxy on patterned substrates. *Applied Physics Letters*, 64(3):363–365, 1994. ISSN 0003-6951.
- [365] S. Debal and B. Kramer. Rashba effect and magnetic field in semiconductor quantum wires. *Physical Review B*, 71(11):115322, 2005.
- [366] A. C. Potter and P. A. Lee. Majorana end states in multiband microstructures with Rashba spin-orbit coupling. *Physical Review B*, 83(9):094525, 2011. ISSN 2469-9950.
- [367] X.-Q. Li and L. Xu. Nonlocality of Majorana zero modes and teleportation: Self-consistent treatment based on the Bogoliubov-de Gennes equation. *Physical Review B*, 101(20):205401, 2020. ISSN 2469-9950.

- [368] S. Das Sarma, Jay D. Sau, and Tudor D. Stanescu. Splitting of the zero-bias conductance peak as smoking gun evidence for the existence of the Majorana mode in a superconductor-semiconductor nanowire. *Physical Review B*, 86(22):220506(R), 2012. ISSN 2469-9950.
- [369] L. Matthiessen. *Grundzüge der antiken und modernen Algebra der litteralen Gleichungen*. Teubner, Leipzig, 1878.
- [370] T. D. Stanescu and S. Tewari. Robust low-energy andreev bound states in semiconductor-superconductor structures: Importance of partial separation of component majorana bound states. *Physical Review B*, 100(15):155429, 2019. ISSN 2469-9950.
- [371] M. Cheng, R. M. Lutchyn, V. Galitski, and S. Das Sarma. Splitting of majorana-fermion modes due to intervortex tunneling in a px+ipy superconductor. *Physical Review Letters*, 103(10):107001, 2009.
- [372] C. Fleckenstein, F. Domínguez, N. Traverso Ziani, and B. Trauzettel. Decaying spectral oscillations in a majorana wire with finite coherence length. *Physical Review B*, 97(15):155425, 2018. ISSN 2469-9950.
- [373] C. Zeng, C. Moore, A. M. Rao, T. D. Stanescu, and S. Tewari. Analytical solution of the finite-length kitaev chain coupled to a quantum dot. *Physical Review B*, 99(9):107, 2019. ISSN 2469-9950.
- [374] C.-X. Liu, J. D. Sau, and S. D. Sarma. Distinguishing topological majorana bound states from trivial andreev bound states: Proposed tests through differential tunneling conductance spectroscopy. *Physical Review B*, 97(21):214502, 2018. ISSN 2469-9950.
- [375] C. Moore, T. D. Stanescu, and S. Tewari. Two-terminal charge tunneling: Disentangling majorana zero modes from partially separated andreev bound states in semiconductor-superconductor heterostructures. *Physical Review B*, 97(16):165302, 2018. ISSN 2469-9950.
- [376] O. A. Awoga, J. Cayao, and A. M. Black-Schaffer. Supercurrent detection of topologically trivial zero-energy states in nanowire junctions. *Physical Review Letters*, 123(11):117001, 2019. ISSN 0031-9007.
- [377] A. C. Potter and P. A. Lee. Multichannel generalization of kitaev’s majorana end states and a practical route to realize them in thin films. *Physical Review Letters*, 105(22):227003, 2010.
- [378] T. D. Stanescu and S. Tewari. Majorana fermions in semiconductor nanowires: fundamentals, modeling, and experiment. *Journal of Physics: Condensed Matter*, 25(23):233201, may 2013.
- [379] R. P. Feynman. *Feynman lectures on physics*. Addison Wesley Longman, Boston, MA, 1970. ISBN 0201021153.
- [380] R. Winkler, S. J. Papadakis, E. P. Poortere, and M. Shayegan. Spin-orbit coupling in two-dimensional electron and hole systems. In Bernhard Kramer, editor, *Advances in solid state physics*, volume 41 of *Advances in Solid State Physics Volume 41*, pages 211–223. Springer, Berlin, Heidelberg Germany and New York, 2001. ISBN 978-3-540-42000-2.

- [381] G. W. Winkler, A. E. Antipov, B. Van Heck, A. A. Soluyanov, L. I. Glazman, M. Wimmer, and R. M. Lutchyn. Unified numerical approach to topological semiconductor-superconductor heterostructures. *Physical Review B*, 99(24):245408, 2019.
- [382] B. Nijholt and A. R. Akhmerov. Orbital effect of magnetic field on the majorana phase diagram. *Phys. Rev. B*, 93:235434, Jun 2016.
- [383] M. V. Berry. Quantal phase factors accompanying adiabatic changes. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 392(1802):45–57, 1984.
- [384] D. R. Hofstadter. Energy levels and wave functions of bloch electrons in rational and irrational magnetic fields. *Phys. Rev. B*, 14:2239–2249, Sep 1976.
- [385] R. Peierls. Zur theorie des diamagnetismus von leitungselektronen. *Zeitschrift für Physik*, 80(11):763–791, 1933.
- [386] P. Wójcik and M. P. Nowak. Durability of the superconducting gap in Majorana nanowires under orbital effects of a magnetic field. *Physical Review B*, 97(23):235445, 2018.
- [387] B. Zocher, M. Horsdal, and B. Rosenow. Proposed detection of the topological phase in ring-shaped semiconductor-superconductor nanowires using coulomb blockade transport. *Phys. Rev. Lett.*, 109:227001, Nov 2012.
- [388] W. Ehrenberg and R. E. Siday. The refractive index in electron optics and the principles of dynamics. *Proceedings of the Physical Society. Section B*, 62(1):8–21, jan 1949.
- [389] C. W. Groth, M. Wimmer, A. R. Akhmerov, and X. Waintal. Kwant: a software package for quantum transport. *New Journal of Physics*, 16(6):063065, 2014. ISSN 1367-2630.
- [390] A. Aharony, O. Entin-Wohlman, B. I. Halperin, and Y. Imry. Phase measurement in the mesoscopic Aharonov-Bohm interferometer. *Physical Review B*, 66(11):115311, 2002. ISSN 2469-9950.
- [391] L. Onsager. Reciprocal Relations in Irreversible Processes. I. *Physical Review*, 37(4):405, 1931.
- [392] M. Büttiker. Four-Terminal Phase-Coherent Conductance. *Physical Review Letters*, 57(14):1761, 1986.
- [393] O. Entin-Wohlman, A. Aharony, Y. Imry, Y. Levinson, and A. Schiller. Broken Unitarity and Phase Measurements in Aharonov-Bohm Interferometers. *Physical Review Letters*, 88(16):166801, 2002.
- [394] H. A. Weidenmüller. Transmission phase of an isolated coulomb blockade resonance. *Phys. Rev. B*, 65:245322, Jun 2002.
- [395] T. Hecht, A. Weichselbaum, Y. Oreg, and J. von Delft. Interplay of mesoscopic and kondo effects for transmission amplitude of few-level quantum dots. *Phys. Rev. B*, 80:115330, Sep 2009.
- [396] R. Landauer. Spatial Variation of Currents and Fields Due to Localized Scatterers in Metallic Conduction. *IBM Journal of Research and Development*, 1(3):223–231, 1957. ISSN 0018-8646.

- [397] S. Datta. *Electronic Transport in Mesoscopic Systems*. Cambridge Studies in Semiconductor Physi. Cambridge University Press, 1997. ISBN 9780521599436.
- [398] T. Ihn. *Semiconductor Nanostructures: Quantum states and electronic transport*. Semiconductor Nanostructures: Quantum States and Electronic Transport. OUP Oxford, 2009. ISBN 9780199534432.
- [399] L. P. Kouwenhoven, N. C. Van der Vaart, A. T. Johnson, W. Kool, C. J. P. M. Harmans, J. G. Williamson, A. A. M. Staring, and C. T. Foxon. Single electron charging effects in semiconductor quantum dots. *Zeitschrift für Physik B Condensed Matter*, 85(3):367–373, 1991.
- [400] H. Grabert and M. H. Devoret. *Single charge tunneling: Coulomb blockade phenomena in nanostructures*, volume 294. Springer Science & Business Media, 2013.
- [401] D. V. Averin and K. K. Likharev. Single electronics: A correlated transfer of single electrons and cooper pairs in systems of small tunnel junctions. In *Modern Problems in Condensed Matter Sciences*, volume 30, pages 173–271. Elsevier, 1991.
- [402] Y. V. Nazarov and Y. M. Blanter. *Quantum Transport: Introduction to Nanoscience*. Cambridge University Press, 2009. ISBN 9780521832465.
- [403] D. V. Averin and Y. V. Nazarov. Virtual electron diffusion during quantum tunneling of the electric charge. *Phys. Rev. Lett.*, 65:2446–2449, Nov 1990.
- [404] R. A. Žak and K. Flensberg. Coulomb blockade of a three-terminal quantum dot. *Phys. Rev. B*, 77:045329, Jan 2008.
- [405] C. W. J. Beenakker. Theory of coulomb-blockade oscillations in the conductance of a quantum dot. *Phys. Rev. B*, 44:1646–1656, Jul 1991.
- [406] P. Lafarge, P. Joyez, D. Esteve, C. Urbina, and M. H. Devoret. Measurement of the even-odd free-energy difference of an isolated superconductor. *Phys. Rev. Lett.*, 70:994–997, Feb 1993.
- [407] M. T. Tuominen, J. M. Hergenrother, T. S. Tighe, and M. Tinkham. Experimental evidence for parity-based $2e$ periodicity in a superconducting single-electron tunneling transistor. *Phys. Rev. Lett.*, 69:1997–2000, Sep 1992.
- [408] F. Braun and J. von Delft. Superconductivity in ultrasmall metallic grains. *Physical Review B*, 59(14):9527, 1999. ISSN 2469-9950.
- [409] T. D. Stanescu, J. D. Sau, R. M. Lutchyn, and S. Das Sarma. Proximity effect at the superconductor–topological insulator interface. *Phys. Rev. B*, 81:241310, Jun 2010.
- [410] J. Alicea. New directions in the pursuit of Majorana fermions in solid state systems. *Reports on Progress in Physics*, 75(7):076501, 2012. ISSN 0034-4885.
- [411] M. Hell, M. Leijnse, and K. Flensberg. Two-Dimensional Platform for Networks of Majorana Bound States. *Physical review letters*, 118(10):107701, 2017. ISSN 0031-9007.
- [412] M. Tinkham. *Introduction to superconductivity*. Dover books on physics. Dover Publ, Mineola, NY, 2. ed. edition, 2004. ISBN 9780486435039.

- [413] H. Boumrar, M. Hamidi, H. Zenia, and S. Lounis. Equivalence of wave function matching and green's functions methods for quantum transport: generalized fisher-lee relation. *Journal of physics. Condensed matter : an Institute of Physics journal*, 32(35):355302, 2020.
- [414] C. Mahaux and H. A. Weidenmüller. Comparison between the R-matrix and eigenchannel methods. *Physical Review*, 170(4):847, 1968.
- [415] E. A. Lynton, B. Serin, and M. Zucker. The superconductive critical temperature and the electronic specific heat of impurle tin. *Journal of Physics and Chemistry of Solids*, 3(3):165–174, 1957. ISSN 0022-3697.
- [416] A. B. Pippard. The effect of alloying on the superconducting transition temperature of tin. *Journal of Physics and Chemistry of Solids*, 3(3):175–177, 1957. ISSN 0022-3697.
- [417] P. W. Anderson. Theory of dirty superconductors. *Journal of Physics and Chemistry of Solids*, 11(1):26–30, 1959. ISSN 0022-3697.
- [418] L. J. Buchholtz and G. Zwicknagl. Identification of p-wave superconductors. *Physical Review B*, 23(11):5788, 1981. ISSN 2469-9950.
- [419] P. Hirschfeld, D. Vollhardt, and P. Wölfle. Resonant impurity scattering in heavy fermion superconductors. *Solid State Communications*, 59(3):111–115, 1986. ISSN 0038-1098.
- [420] S. Schmitt-Rink, K. Miyake, and C. M. Varma. Transport and thermal properties of heavy-fermion superconductors: A unified picture. *Physical Review Letters*, 57(20):2575, 1986.
- [421] K. Maki and E. Puchkaryov. Impurity scattering in isotropic p-wave superconductors. *EPL (Europhysics Letters)*, 45(2):263, 1999. ISSN 0295-5075.
- [422] B. Zocher, M. Horsdal, and B. Rosenow. Robustness of topological order in semiconductor–superconductor nanowires in the Coulomb blockade regime. *New Journal of Physics*, 15(8):085003, 2013. ISSN 1367-2630.
- [423] J. P. T. Stenger, N. T. Bronn, D. J. Egger, and D. Pekker. Simulating the dynamics of braiding of Majorana zero modes using an IBM quantum computer. *Physical Review Research*, 3(3):033171, 2021.
- [424] N. Hansen, yoshihikoueno, ARF1, K. Nozawa, M. Chan, Y. Akimoto, and D. Brockhoff. *CMA-ES/pycma: r3.1.0*. Zenodo, 2021.
- [425] B. Bauer, T. Karzig, R. Mishmash, A. Antipov, and J. Alicea. Dynamics of Majorana-based qubits operated with an array of tunable gates. *SciPost Physics*, 5(1):004, 2018. ISSN 2542-4653.
- [426] A. R. Mills, D. M. Zajac, M. J. Gullans, F. J. Schupp, T. M. Hazard, and J. R. Petta. Shuttling a single charge across a one-dimensional array of silicon quantum dots. *Nature communications*, 10(1):1–6, 2019.
- [427] A. Pöschl. *Nonlocal transport signatures of Andreev Bound States*. PhD thesis, Center for Quantum Devices, Niels Bohr Institute, University of Copenhagen, 2022.

- [428] F. Pientka, G. Kells, A. Romito, P. W. Brouwer, and F. Von Oppen. Enhanced zero-bias majorana peak in the differential tunneling conductance of disordered multisubband quantum-wire/superconductor junctions. *Physical review letters*, 109(22):227006, 2012.
- [429] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals. Understanding deep learning (still) requires rethinking generalization. *Communications of the ACM*, 64(3):107–115, 2021. ISSN 0001-0782.
- [430] G. De Palma, B. T. Kiani, and S. Lloyd. Random deep neural networks are biased towards simple functions. *preprint arXiv:1812.10156*, 2018.
- [431] G. Valle-Perez, C. Q. Camargo, and A. A. Louis. Deep learning generalizes because the parameter-function map is biased towards simple functions. *preprint arXiv:1805.08522*, 2018.
- [432] O. Cohen, O. Malka, and Z. Ringel. Learning curves for overparametrized deep neural networks: A field theory perspective. *Physical Review Research*, 3(2):023034, 2021.
- [433] A. K. Lampinen and S. Ganguli. An analytic theory of generalization dynamics and transfer learning in deep linear networks. *preprint arXiv:1809.10374*, 2018.
- [434] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala. Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019.
- [435] Michael A. N. *Neural Networks and Deep Learning*. Determination Press, 2015.
- [436] Y. Long, Y. Gong, Z. Xiao, and Q. Liu. Accurate object localization in remote sensing images based on convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 55(5):2486–2498, 2017.
- [437] D. Scherer, A. Müller, and S. Behnke. Evaluation of pooling operations in convolutional architectures for object recognition. In *International conference on artificial neural networks*, pages 92–101. Springer, 2010.
- [438] D. C. Ciresan, U. Meier, J. Masci, L. M. Gambardella, and J. Schmidhuber. Flexible, high performance convolutional neural networks for image classification. In *Twenty-second international joint conference on artificial intelligence*, 2011.
- [439] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee, 2009.
- [440] K. K. Pal and K. S. Sudeep. Preprocessing for image classification by convolutional neural networks. In *2016 IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology (RTEICT)*, pages 1778–1781, 2016.
- [441] D. E. Rumelhart, R. Durbin, R. Golden, and Y. Chauvin. Backpropagation: The basic theory. *Backpropagation: Theory, architectures and applications*, pages 1–34, 1995.

- [442] M. Thamm, M. Staats, and B. Rosenow. All code, scripts and data used in this work are included in a Zenodo archive: <https://zenodo.org/record/7075360>. *Zenodo*, 2022.
- [443] Y. Idelbayev and M. A. Carreira-Perpinán. Low-rank compression of neural nets: Learning the rank of each layer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8049–8059, 2020.
- [444] H. Sedghi, V. Gupta, and P. M. Long. The singular values of convolutional layers. In *International Conference on Learning Representations*, 2019.
- [445] A. Clauset, C. R. Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM review*, 51(4):661–703, 2009.
- [446] M. Brack, J. Damgaard, A. S. Jensen, H. C. Pauli, V. M. Strutinsky, and C. Y. Wong. Funny Hills: The Shell-Correction Approach to Nuclear Shell Effects and Its Applications to the Fission Process. *Reviews of Modern Physics*, 44(2):320–405, 1972.
- [447] B. Efron. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1), 1979. ISSN 0090-5364.
- [448] Bradley Efron and Robert Tibshirani. *An introduction to the bootstrap*, volume 57 of *Mono-graphs on statistics and applied probability*. Chapman & Hall, Boca Raton, F., [nachdr.] edition, 1998. ISBN 0412042312.
- [449] B. Efron. Second Thoughts on the Bootstrap. *Statistical Science*, 18(2):135–140, 2003. ISSN 08834237.
- [450] D. Berger. `empyiricalRMT` 0.4.1 python package: <https://pypi.org/project/empyiricalRMT/>. *pypi package*, 2022.
- [451] L. Denby and C. L. Mallows. Computing sciences and statistics: Proceedings of the 23rd symposium on the interface, edited by em keramidas. *Interface Foundation, Fairfax Station, VA*, pages 54–57, 1991.
- [452] S. S. Du, X. Zhai, B. Póczos, and A. Singh. Gradient descent provably optimizes over-parameterized neural networks. *arXiv preprint arXiv:1810.02054*, 2018.
- [453] C. Li and C. J. Shi. Constrained optimization based low-rank approximation of deep neural networks. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 732–747, 2018.
- [454] S. Du, J. Lee, H. Li, L. Wang, and X. Zhai. Gradient descent finds global minima of deep neural networks. In *International conference on machine learning*, pages 1675–1685. PMLR, 2019.
- [455] D. Zou, Y. Cao, D. Zhou, and Q. Gu. Gradient descent optimizes over-parameterized deep relu networks. *Machine learning*, 109(3):467–492, 2020.
- [456] L. Chizat, E. Oyallon, and F. Bach. On lazy training in differentiable programming. *Advances in Neural Information Processing Systems*, 32, 2019.
- [457] R. Novak, L. Xiao, J. Hron, J. Lee, A. A. Alemi, J. Sohl-Dickstein, and S. S. Schoenholz. Neural tangents: Fast and easy infinite neural networks in python. In *International Conference on Learning Representations*, 2020.

- [458] T. Guhr and A. Müller-Groeling. Spectral correlations in the crossover between gue and poisson regularity: on the identification of scales. *Journal of Mathematical Physics*, 38(4): 1870–1887, 1997.
- [459] B. M. Hill. A simple general approach to inference about the tail of a distribution. *The annals of statistics*, pages 1163–1174, 1975.
- [460] V. Akgiray and G. G. Booth. The siable-law model of stock returns. *Journal of Business & Economic Statistics*, 6(1):51–57, 1988.
- [461] B. N. Cheng and S. T. Rachev. Multivariate stable futures prices. *Mathematical Finance*, 5(2):133–153, 1995.
- [462] P. Gopikrishnan, V. Plerou, L. A. Nunes Amaral, M. Meyer, and H. E. Stanley. Scaling of the distribution of fluctuations of financial market indices. *Physical Review E*, 60(5): 5305, 1999.
- [463] T. Lux. The stable paretian hypothesis and the frequency of large returns: an examination of major german stocks. *Applied financial economics*, 6(6):463–475, 1996.
- [464] C. Quintos, Z. Fan, and P. C. B. Phillips. Structural change tests in tail behaviour and the asian crisis. *The Review of Economic Studies*, 68(3):633–663, 2001.
- [465] N. H. Chan, S.-J. Deng, L. Peng, and Z. Xia. Interval estimation of value-at-risk based on garch models with heavy-tailed innovations. *Journal of Econometrics*, 137(2):556–576, 2007.
- [466] J. B. Hill. On tail index estimation for dependent, heterogeneous data. *Econometric Theory*, 26(5):1398–1436, 2010.
- [467] J. Alstott, E. Bullmore, and D. Plenz. powerlaw: a python package for analysis of heavy-tailed distributions. *PloS one*, 9(1):e85777, 2014.
- [468] J. Bun, J.-P. Bouchaud, and M. Potters. Cleaning large correlation matrices: Tools from random matrix theory. *Physics Reports*, 666:1–109, 2017. ISSN 0370-1573.
- [469] R. Schäfer, N. F. Nilsson, and T. Guhr. Power mapping with dynamical adjustment for improved portfolio optimization. *Quantitative Finance*, 10(1):107–119, 2010.
- [470] F. Benaych-Georges and R. R. Nadakuditi. The singular values and vectors of low rank perturbations of large rectangular random matrices. *Journal of Multivariate Analysis*, 111: 120–135, 2012.
- [471] V. Alba and P. Calabrese. Entanglement dynamics after quantum quenches in generic integrable systems. *SciPost Physics*, 4(3):017, 2018.
- [472] A. Del Maestro, H. Barghathi, and B. Rosenow. Measuring postquench entanglement entropy through density correlations. *Physical Review Research*, 4(2):L022023, 2022.
- [473] A. Del Maestro. DelMaestroGroup/papers-code-EntanglementQuantumQuench: <https://github.com/DelMaestroGroup/papers-code-EntanglementQuantumQuench/tree/v2.0>. <https://github.com/DelMaestroGroup>, 2022.

- [474] M. Thamm, H. Radhakrishnan, H. Barghathi, B. Rosenow, and A. Del Maestro. All code, scripts and data used in this work are included in a GitHub repository: <https://github.com/DelMaestroGroup/papers-code-OneParticleEntanglementEntropy>. <https://github.com/DelMaestroGroup>, 2022.
- [475] M. Fishman, S. R. White, and E. M. Stoudenmire. The ITensor software library for tensor network calculations, 2020.
- [476] T. Giamarchi. *Quantum Physics in One Dimension*. Clarendon Press, Oxford, U. K., 2004. ISBN 9780198525004.
- [477] H. Bethe. Zur theorie der metalle. *Zeitschrift für Physik*, 71(3):205–226, 1931.
- [478] N. N. Bogoljubov. On a new method in the theory of superconductivity. *Il Nuovo Cimento (1955-1965)*, 7(6):794–805, 1958.
- [479] A. G. Del Maestro and M. J. P. Gingras. Quantum spin fluctuations in the dipolar Heisenberg-like rare earth pyrochlores. *Journal of Physics: Condensed Matter*, 16(20):3339–3353, 2004. ISSN 0953-8984.
- [480] C. N. Yang. Concept of Off-Diagonal Long-Range Order and the Quantum Phases of Liquid He and of Superconductors. *Rev. Mod. Phys.*, 34(4):694, 1962.
- [481] B. C. Carlson and J. M. Keller. Eigenvalues of Density Matrices. *Phys. Rev.*, 121(3):659, 1961.
- [482] F. Sasaki. Eigenvalues of Fermion Density Matrices. *Phys. Rev.*, 138(5B):1338, 1965.
- [483] M. Lemm. On the entropy of fermionic reduced density matrices, 2017.
- [484] M. Lemm and M. M. Wilde. Information-theoretic limitations on approximate quantum cloning and broadcasting. *Phys. Rev. A*, 96(1):012304, 2017.
- [485] S.-A. Cheong and C. L. Henley. Many-body density matrices for free fermions. *Phys. Rev. B*, 69(7):075111, 2004.
- [486] C. Lanczos. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *J. Res. Natl. Bur. Stand.*, 45(4):255, 1950.
- [487] H. Barghathi, C. Usadi, M. Beck, and A. Del Maestro. Compact unary coding for bosonic states as efficient as conventional binary encoding for fermionic states. *Phys. Rev. B*, 105(12):L121116, 2022.
- [488] Harini Radhakrishnan, Matthias Thamm, Hatem Barghathi, Bernd Rosenow, and Adrian Del Maestro. A scaling function for the particle entanglement entropy of fermions. *preprint arXiv:2302.09093*, 2023.
- [489] J. Biamonte and V. Bergholm. Tensor networks in a nutshell. *arXiv preprint arXiv:1708.00006*, 2017.
- [490] S. R. White. Density matrix formulation for quantum renormalization groups. *Physical review letters*, 69(19):2863, 1992.

- [491] S. R. White. Density-matrix algorithms for quantum renormalization groups. *Physical review b*, 48(14):10345, 1993.
- [492] U. Schollwöck. The density-matrix renormalization group. *Reviews of modern physics*, 77(1):259, 2005.
- [493] M. Suzuki. Generalized Trotter’s formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems. *Communications in Mathematical Physics*, 51(2):183–190, 1976. ISSN 1432-0916.
- [494] S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck, and C. Hubig. Time-evolution methods for matrix-product states. *Ann. Phys.*, 411:167998, 2019. ISSN 00034916.
- [495] The merge request for our bug fixes in ITensorGPU.jl can be found here: <https://github.com/ITensor/ITensors.jl/pull/871>.
- [496] C. Hubig, J. Haegeman, and U. Schollwöck. Error estimates for extrapolations with matrix-product states. *Phys. Rev. B*, 97:045125, Jan 2018.
- [497] H. Flyvbjerg and H. G. Petersen. Error estimates on averages of correlated data. *The Journal of Chemical Physics*, 91(1):461–466, 1989. ISSN 0021-9606.

Original publications and author contributions

- [E1] Matthias Thamm and Bernd Rosenow, *Transmission amplitude through a Coulomb blockaded Majorana wire*, Phys. Rev. Research **3**, 023221 (2021)

Matthias Thamm performed the numerical and analytical calculations, wrote the code for the numerical computations, created all figures, and wrote the manuscript. Bernd Rosenow supervised the work and edited the manuscript.

- [E2] Matthias Thamm and Bernd Rosenow, *Machine learning optimization of Majorana hybrid nanowires*, preprint **arXiv:2208.02182** (accepted for Physical Review Letters)

Matthias Thamm performed the numerical and analytical calculations, wrote the code for the numerical computations, created all figures, and wrote the manuscript. Bernd Rosenow supervised the work and edited the manuscript.

- [E3] Matthias Thamm¹, Max Staats¹, and Bernd Rosenow, *Random matrix analysis of deep neural network weight matrices*, Phys. Rev. E **106**, 054124(2022)

Work on this project has been performed as a true collaboration between Max Staats and Matthias Thamm, such that it is difficult to disentangle individual contributions. Exceptions are that Matthias Thamm performed the analysis and wrote the code for the section about networks in different learning regimes, and that Matthias Thamm prepared the code that is published alongside the manuscript. Bernd Rosenow supervised the work and edited the manuscript.

- [E4] Max Staats¹, Matthias Thamm¹, and Bernd Rosenow, *Boundary between noise and information applied to filtering neural network weight matrices*, preprint **arXiv:2206.03927** (2022)

Work on this project has been performed as a true collaboration between Max Staats and Matthias Thamm, such that it is difficult to disentangle individual contributions. One exception is the analysis of different reshaping techniques for convolutional weight matrices, which has been primarily performed by Max Staats. Bernd Rosenow supervised the work and edited the manuscript.

- [E5] Matthias Thamm, Harini Radhakrishnan, Hatem Barghathi, Bernd Rosenow, and Adrian Del Maestro, *One-particle entanglement for one dimensional spinless fermions after an interaction quantum quench*, Phys. Rev. B **106**, 165116 (2022)

Harini Radhakrishnan and Matthias Thamm performed the Luttinger liquid analytical calculation together. Hatem Barghathi and Adrian Del Maestro wrote a previous version of the code for exact diagonalization. Matthias Thamm added improvements to the exact diagonalization code and wrote the code for the approximate DMRG computations with input from Hatem Barghathi. Matthias Thamm performed the numerical computations, the analysis of the data, the comparison between analytics and numerics, and created the figures. Adrian Del Maestro and Bernd Rosenow wrote the introduction of the manuscript. Adrian Del Maestro wrote the conclusion section of the manuscript. Hatem Barghathi wrote the section about the time dependence of the spectrum of the post quench 1-RDM and the appendix about the definition of the action of symmetry operators based on fermion operators. Matthias Thamm wrote all other sections with input from Harini

¹Equal author contribution

Radhakrishnan in the section about the analytical calculation. All authors edited the manuscript. Adrian Del Maestro and Bernd Rosenow supervised the work.

- [E6] Harini Radhakrishnan, Matthias Thamm, Hatem Barghathi, Bernd Rosenow, and Adrian Del Maestro, *A Scaling Function for the Particle Entanglement Entropy of Fermions*, preprint **arXiv:2302.09093** (2023)

Harini Radhakrishnan performed the analytical analysis with input from Hatem Barghathi. Harini Radhakrishnan performed the data analysis and created the figures. Harini Radhakrishnan wrote the main part of the manuscript. Matthias Thamm wrote the DMRG code and extended it to compute n -particle entanglement with input from Hatem Barghathi. Matthias Thamm ran the DMRG computations. Matthias Thamm wrote the section about technical details on the numerical methods used. Adrian Del Maestro wrote the introductory part of the manuscript. Adrian Del Maestro, Harini Radhakrishnan, and Hatem Barghathi wrote the discussion section. All authors edited the manuscript. Adrian Del Maestro and Bernd Rosenow supervised the work.

Acknowledgement

First and foremost, I would like to thank my supervisor Prof. Dr. Bernd Rosenow for his excellent advice, patience, and the opportunities as well as the freedom he gave me. His passion for physics, his extensive knowledge, and his analytical way of thinking were always very inspiring, and I am very grateful for the fruitful and professional collaboration.

He also made possible my research stay at the University of Tennessee in Knoxville, where the inspiring collaboration with the research group of Prof. Adrian Del Maestro further ignited my enthusiasm for research. Adrian's helpful advice and his eye for detail were great motivations for me. In addition, I would like to thank Hatem Barghathi for many fruitful discussions – his incredible passion and broad coding skills are truly awesome. Also, a big thank you to Harini Radhakrishnan for the numerous helpful but also fun discussions. I look forward for future collaboration with Adrian's group.

Furthermore, I would like to thank Max Staats for the fruitful and constructive collaboration and the numerous great discussions. I hope that we will continue our collaboration in the future.

I acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG) under Grant No. 406116891 within the Research Training Group RTG 2522/1. I would like to thank the RTG for the financial support, the interesting events, but also the lot of freedom it has given to its members. I further acknowledge the University of Tennessee, Knoxville and the Institute of Advanced Materials and Manufacturing for the hospitality during my research visit.

Of course, the good environment and climate at the Institute for Theoretical Physics in Leipzig and in our research group have also been crucial. For this I would like to thank, among others: Niels, Casper, Heinrich-Gregor, Giovanni, Assem, Andreas, Roman, and Max, as well as Frau Voigt, Frau Moreno, Frau Menge, and Frau Thieme.

I am especially grateful to Harini Radhakrishnan for proofreading this thesis.

Of course, creating this thesis would not have been possible without the support of my friends and family. I am especially grateful to my parents, my wife, and my son for their immeasurable support.

Appendices

A.1 Mean field approximation for the superconductor Hamiltonian

We start from the Hamiltonian Eq. (2.1) [255, 284, 336]

$$H_{\text{int}} = \frac{1}{2} \sum_{\mathbf{k}\mathbf{k}'} \sum_{\substack{\sigma_1\sigma_2 \\ \sigma_3\sigma_4}} V_{\sigma_1\sigma_2\sigma_3\sigma_4}(\mathbf{k}, \mathbf{k}') c_{-\mathbf{k}\sigma_1}^\dagger c_{\mathbf{k}\sigma_2}^\dagger c_{\mathbf{k}'\sigma_3} c_{-\mathbf{k}'\sigma_4} , \quad (\text{A.1})$$

and perform a mean field approximation by introducing operators

$$\delta_{\mathbf{k}\sigma_1\sigma_2} = c_{\mathbf{k}\sigma_1} c_{-\mathbf{k}\sigma_2} - \langle c_{\mathbf{k}\sigma_1} c_{-\mathbf{k}\sigma_2} \rangle , \quad (\text{A.2})$$

$$\delta_{\mathbf{k}\sigma_1\sigma_2}^\dagger = c_{-\mathbf{k}\sigma_2}^\dagger c_{\mathbf{k}\sigma_1}^\dagger - \langle c_{-\mathbf{k}\sigma_2}^\dagger c_{\mathbf{k}\sigma_1}^\dagger \rangle . \quad (\text{A.3})$$

We neglect powers of second order or higher in these operators. Because of the fermionic commutator relations $\{c_{\mathbf{k}\sigma}, c_{\mathbf{k}'\sigma'}\} = 0 = \{c_{\mathbf{k}\sigma}^\dagger, c_{\mathbf{k}'\sigma'}^\dagger\}$, $\{c_{\mathbf{k}\sigma}^\dagger, c_{\mathbf{k}'\sigma'}\} = \delta_{\mathbf{k}\mathbf{k}'}\delta_{\sigma\sigma'}$, the pair interaction fulfills [255, 336]

$$V_{\sigma_1\sigma_2\sigma_3\sigma_4}(\mathbf{k}, \mathbf{k}') = -V_{\sigma_2\sigma_1\sigma_3\sigma_4}(-\mathbf{k}, \mathbf{k}') = -V_{\sigma_1\sigma_2\sigma_4\sigma_3}(\mathbf{k}, -\mathbf{k}') = V_{\sigma_4\sigma_3\sigma_2\sigma_1}(\mathbf{k}', \mathbf{k}) . \quad (\text{A.4})$$

We use the new operators in the Hamiltonian, which yields

$$H_{\text{int}} = \frac{1}{2} \sum_{\mathbf{k}\mathbf{k}'} \sum_{\substack{\sigma_1\sigma_2 \\ \sigma_3\sigma_4}} V_{\sigma_1\sigma_2\sigma_3\sigma_4}(\mathbf{k}, \mathbf{k}') \left(\delta_{\mathbf{k}\sigma_2\sigma_1}^\dagger + \langle c_{-\mathbf{k}\sigma_1}^\dagger c_{\mathbf{k}\sigma_2}^\dagger \rangle \right) \left(\delta_{\mathbf{k}'\sigma_3\sigma_4} + \langle c_{\mathbf{k}'\sigma_3} c_{-\mathbf{k}'\sigma_4} \rangle \right) . \quad (\text{A.5})$$

$$\begin{aligned} &= \frac{1}{2} \sum_{\mathbf{k}\mathbf{k}'} \sum_{\substack{\sigma_1\sigma_2 \\ \sigma_3\sigma_4}} V_{\sigma_1\sigma_2\sigma_3\sigma_4} \left[\delta_{\mathbf{k}\sigma_2\sigma_1}^\dagger \langle c_{\mathbf{k}'\sigma_3} c_{-\mathbf{k}'\sigma_4} \rangle + \delta_{\mathbf{k}'\sigma_3\sigma_4} \langle c_{-\mathbf{k}\sigma_1}^\dagger c_{\mathbf{k}\sigma_2}^\dagger \rangle \right. \\ &\quad \left. + \langle c_{-\mathbf{k}\sigma_1}^\dagger c_{\mathbf{k}\sigma_2}^\dagger \rangle \langle c_{\mathbf{k}'\sigma_3} c_{-\mathbf{k}'\sigma_4} \rangle \right] + \mathcal{O}(\delta^2) . \end{aligned} \quad (\text{A.6})$$

Using Eq. (A.2) and Eq. (A.3), we reinsert the linear terms such that

$$\begin{aligned} H_{\text{int}} &= \frac{1}{2} \sum_{\mathbf{k}\mathbf{k}'} \sum_{\substack{\sigma_1\sigma_2 \\ \sigma_3\sigma_4}} V_{\sigma_1\sigma_2\sigma_3\sigma_4}(\mathbf{k}, \mathbf{k}') \left[c_{-\mathbf{k}\sigma_1}^\dagger c_{\mathbf{k}\sigma_2}^\dagger \langle c_{\mathbf{k}'\sigma_3} c_{-\mathbf{k}'\sigma_4} \rangle - \langle c_{-\mathbf{k}\sigma_1}^\dagger c_{\mathbf{k}\sigma_2}^\dagger \rangle \langle c_{\mathbf{k}'\sigma_3} c_{-\mathbf{k}'\sigma_4} \rangle \right. \\ &\quad \left. + c_{\mathbf{k}'\sigma_3} c_{-\mathbf{k}'\sigma_4} \langle c_{-\mathbf{k}\sigma_1}^\dagger c_{\mathbf{k}\sigma_2}^\dagger \rangle - \langle c_{\mathbf{k}'\sigma_3} c_{-\mathbf{k}'\sigma_4} \rangle \langle c_{-\mathbf{k}\sigma_1}^\dagger c_{\mathbf{k}\sigma_2}^\dagger \rangle \right. \\ &\quad \left. + \langle c_{-\mathbf{k}\sigma_1}^\dagger c_{\mathbf{k}\sigma_2}^\dagger \rangle \langle c_{\mathbf{k}'\sigma_3} c_{-\mathbf{k}'\sigma_4} \rangle \right] \end{aligned} \quad (\text{A.7})$$

$$\begin{aligned} &= \frac{1}{2} \sum_{\mathbf{k}\mathbf{k}'} \sum_{\sigma\sigma'} \left[\sum_{\tilde{\mathbf{k}}} \sum_{\tilde{\sigma}\tilde{\sigma}'} V_{\sigma\sigma'\tilde{\sigma}\tilde{\sigma}'}(\mathbf{k}, \tilde{\mathbf{k}}) \langle c_{\tilde{\mathbf{k}}\tilde{\sigma}} c_{-\tilde{\mathbf{k}}\tilde{\sigma}'} \rangle c_{-\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma'}^\dagger \right. \\ &\quad \left. + \sum_{\tilde{\mathbf{k}}} \sum_{\tilde{\sigma}\tilde{\sigma}'} V_{\tilde{\sigma}\tilde{\sigma}'\sigma\sigma'}(\tilde{\mathbf{k}}, \mathbf{k}) \langle c_{-\tilde{\mathbf{k}}\tilde{\sigma}}^\dagger c_{\tilde{\mathbf{k}}\tilde{\sigma}'}^\dagger \rangle c_{\mathbf{k}\sigma} c_{-\mathbf{k}\sigma'} \right] - \text{const.} + \mathcal{O}(\delta^2) . \end{aligned} \quad (\text{A.8})$$

In the first term, we change $\mathbf{k} \rightarrow -\mathbf{k}$ which leaves the sums unchanged and use $V_{\sigma\sigma'\tilde{\sigma}\tilde{\sigma}'}(-\mathbf{k}, \tilde{\mathbf{k}}) = -V_{\sigma'\sigma\tilde{\sigma}\tilde{\sigma}'}(\mathbf{k}, \tilde{\mathbf{k}})$. In the second term, we change $\tilde{\mathbf{k}} \rightarrow -\tilde{\mathbf{k}}$ which leaves the sums unchanged and use that $V_{\tilde{\sigma}\tilde{\sigma}'\sigma\sigma'}(-\tilde{\mathbf{k}}, \mathbf{k}) = -V_{\sigma'\sigma\tilde{\sigma}\tilde{\sigma}'}(\mathbf{k}, \tilde{\mathbf{k}})$. Additionally, we commute the operators in the second term both in the expectation value and outside, which yields two minus signs. Therefore, the Hamiltonian is given by

$$H_{\text{int}} = \frac{1}{2} \sum_{\mathbf{k}\mathbf{k}'} \sum_{\sigma\sigma'} \left[- \sum_{\tilde{\mathbf{k}}} \sum_{\tilde{\sigma}\tilde{\sigma}'} V_{\sigma'\sigma\tilde{\sigma}\tilde{\sigma}'}(\mathbf{k}, \tilde{\mathbf{k}}) \langle c_{\tilde{\mathbf{k}}\tilde{\sigma}} c_{-\tilde{\mathbf{k}}\tilde{\sigma}'} \rangle c_{\mathbf{k}\sigma}^\dagger c_{-\mathbf{k}\sigma'}^\dagger \right. \\ \left. - \sum_{\tilde{\mathbf{k}}} \sum_{\tilde{\sigma}\tilde{\sigma}'} V_{\sigma'\sigma\tilde{\sigma}\tilde{\sigma}'}(\mathbf{k}, \tilde{\mathbf{k}}) \langle c_{-\tilde{\mathbf{k}}\tilde{\sigma}'}^\dagger c_{\tilde{\mathbf{k}}\tilde{\sigma}}^\dagger \rangle c_{-\mathbf{k}\sigma'} c_{\mathbf{k}\sigma} \right] - \text{const.} + \mathcal{O}(\delta^2) . \quad (\text{A.9})$$

We define the pair potential [255, 336]

$$\Delta_{\sigma\sigma'}(\mathbf{k}) = - \sum_{\tilde{\mathbf{k}}} \sum_{\tilde{\sigma}\tilde{\sigma}'} V_{\sigma'\sigma\tilde{\sigma}\tilde{\sigma}'}(\mathbf{k}, \tilde{\mathbf{k}}) \langle c_{\tilde{\mathbf{k}}\tilde{\sigma}} c_{-\tilde{\mathbf{k}}\tilde{\sigma}'} \rangle , \quad (\text{A.10})$$

$$\Delta_{\sigma\sigma'}^\dagger(\mathbf{k}) = - \sum_{\tilde{\mathbf{k}}} \sum_{\tilde{\sigma}\tilde{\sigma}'} V_{\sigma'\sigma\tilde{\sigma}\tilde{\sigma}'}(\mathbf{k}, \tilde{\mathbf{k}}) \langle c_{-\tilde{\mathbf{k}}\tilde{\sigma}'}^\dagger c_{\tilde{\mathbf{k}}\tilde{\sigma}}^\dagger \rangle , \quad (\text{A.11})$$

with $\Delta_{\sigma\sigma'}(\mathbf{k}) = -\Delta_{\sigma'\sigma}(-\mathbf{k})$ such that the Hamiltonian can be written as

$$H = \sum_{\mathbf{k}\sigma\sigma'} h_{\sigma\sigma'}(\mathbf{k}) c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma'} + \frac{1}{2} \sum_{\mathbf{k}\sigma\sigma'} \left[\Delta_{\sigma\sigma'}(\mathbf{k}) c_{\mathbf{k}\sigma}^\dagger c_{-\mathbf{k}\sigma'}^\dagger + \Delta_{\sigma\sigma'}^\dagger(\mathbf{k}) c_{-\mathbf{k}\sigma'} c_{\mathbf{k}\sigma} \right] - \text{const.} + \mathcal{O}(\delta^2) . \quad (\text{A.12})$$

A.2 Zero energy solutions at a domain wall in a spin-polarized one dimensional p-wave superconductor

For a one dimensional spin polarized p-wave superconductor with a domain wall $\mu(x) = \mu \text{sgn}(x)$, we found that the BdG equation in Sec. 2.2.1 is given by Eq. (2.54)

$$\begin{pmatrix} -\hbar^2/(2m) \partial_x^2 - \mu(x) & -\Delta \partial_x \\ \Delta \partial_x & \hbar^2/(2m) \partial_x^2 + \mu(x) \end{pmatrix} \begin{pmatrix} u_0(x) \\ v_0(x) \end{pmatrix} = 0 . \quad (\text{A.13})$$

This is

$$-\frac{\hbar^2}{2m} u_0''(x) - \mu(x) u_0(x) - \Delta v_0'(x) = 0 \quad (\text{A.14})$$

$$\frac{\hbar^2}{2m} v_0''(x) + \mu(x) v_0(x) + \Delta u_0'(x) = 0 . \quad (\text{A.15})$$

Moreover, the wave function and its first derivative are continuous. We therefore calculate the solution for $x > 0$ and $x < 0$ separately and then use the continuity for joining them at $x = 0$. In addition, the particle-hole operator $\mathcal{P} = \tau_x \mathcal{K}$ with $\mathcal{P}^2 = 1$ transforms the Hamiltonian as $\mathcal{P} \mathcal{H}_{\text{BdG}}(x) \mathcal{P}^{-1} = -\mathcal{H}_{\text{BdG}}^T(x)$ where the basis is given by $\psi^\dagger(x) = (\psi^\dagger(x), \psi(x))$ and $\partial_x^T = -\partial_x$, $\partial_x^* = \partial_x$. From this we find that for every eigenstate $\xi_E(x)$ with eigenvalue E , i.e. $\mathcal{H}_{\text{BdG}}(x) \xi_E(x) = E \xi_E(x)$, there exists an eigenstate $\mathcal{P} \xi_E(x)$ with $\mathcal{P} \mathcal{H}_{\text{BdG}}(x) \mathcal{P}^{-1} \mathcal{P} \xi_E(x) = E \mathcal{P} \xi_E(x) = -\mathcal{H}_{\text{BdG}}(x) \mathcal{P} \xi_E(x) = E \mathcal{P} \xi_E(x)$. Thus, $\mathcal{P} \xi_E(x)$ has the eigenvalue $-E$. If a single zero energy solution exists, this implies that $\mathcal{P} \xi_E(x) = \xi_E(x)$. Assuming there are two $E = 0$

solutions with $\mathcal{P}\xi_{E=0} \neq \xi_{E=0}$, we can construct the solutions $\tilde{\xi}_{\pm} = (\mathcal{P}\xi_{E=0} \pm \xi_{E=0})/\sqrt{2}$, which satisfy $\mathcal{P}\tilde{\xi}_+ = \tilde{\xi}_+$ and $\mathcal{P}\tilde{\xi}_- = -\tilde{\xi}_-$. In the case of our zero energy solution $(u_0(x), v_0(x))^T$, this means $\mathcal{P}(u_0(x), v_0(x))^T = (v_0^*(x), u_0^*(x))^T$. We therefore look for solutions with $u_0(x) = \pm v_0^*(x)$. Thus, both BdG equations are equivalent and given by

$$-\frac{\hbar^2}{2m}u_0''(x) - \mu(x)u_0(x) \mp \Delta(u_0')^*(x) = 0. \quad (\text{A.16})$$

This is a damped harmonic oscillator in x . To obtain a solution for all x , we write $\mu(x)$, while keeping in mind that it is a constant for both $x > 0$ and $x < 0$. We use the ansatz $u_0(x) = e^{Ax}$ and get

$$0 = -\frac{\hbar^2}{2m}A^2 - \mu(x) \mp \Delta A \quad (\text{A.17})$$

$$\Rightarrow A = (\mp_1) \frac{m\Delta}{\hbar^2} \pm_2 \sqrt{\frac{m^2\Delta^2}{\hbar^4} - \frac{2m\mu(x)}{\hbar^2}}, \quad (\text{A.18})$$

where \mp_1 corresponds to $u_0 = \mp v_0$ and \pm_2 to the two solutions for A . Here, the system can be overdamped for $\mu < m\Delta/(2\hbar^2)$ and underdamped if $\mu > m\Delta^2/(2\hbar^2)$. This yields the solutions

$$u_0(x > 0) = e^{\mp|x|/\xi} \left(C_1 \exp \left[\sqrt{\xi^{-2} - \frac{2m\mu}{\hbar^2}}|x| \right] + C_2 \exp \left[-\sqrt{\xi^{-2} - \frac{2m\mu}{\hbar^2}}|x| \right] \right) \quad (\text{A.19})$$

$$u_0(x < 0) = e^{\pm|x|/\xi} \left(C_3 \exp \left[-\sqrt{\xi^{-2} + \frac{2m\mu}{\hbar^2}}|x| \right] + C_4 \exp \left[\sqrt{\xi^{-2} + \frac{2m\mu}{\hbar^2}}|x| \right] \right) \quad (\text{A.20})$$

with $\xi = \hbar^2/(m\Delta)$, where the signs \pm correspond to $u_0(x) = \pm v_0^*(x)$.

a) Overdamped case $\mu < m\Delta/(2\hbar^2)$

In this case, requiring normalizability restricts possible solutions to

$$u_0(x > 0) = v_0(x > 0) = e^{-x/\xi} \left(C_1 \exp \left[\sqrt{\xi^{-2} - \frac{2m\mu}{\hbar^2}}x \right] + C_2 \exp \left[-\sqrt{\xi^{-2} - \frac{2m\mu}{\hbar^2}}x \right] \right) \quad (\text{A.21})$$

$$u_0(x < 0) = \begin{cases} -v_0(x < 0) = e^{+x/\xi} C_3 \exp \left[\sqrt{\xi^{-2} + \frac{2m\mu}{\hbar^2}}x \right] & \text{case i) or} \\ +v_0(x < 0) = e^{-x/\xi} C_3 \exp \left[\sqrt{\xi^{-2} + \frac{2m\mu}{\hbar^2}}x \right] & \text{case ii) } \end{cases} \quad (\text{A.22})$$

The continuity condition for $u_0(x=0)$ and $v_0(x=0)$ yields in case i) that $C_1 + C_2 = C_3$ and $C_1 + C_2 = -C_3$ with the solution $C_3 = 0$ and $C_1 = -C_2$. But in this case, continuity for $u_0'(x=0)$ then requires $C_1 = 0$ such that case i) is not a solution. In case ii), continuity of $u_0(x=0)$ and $v_0(x=0)$ yields

$$C_1 + C_2 = C_3 \quad (\text{A.23})$$

and from the first derivatives, we obtain

$$\left(-\frac{1}{\xi} + \sqrt{\xi^{-2} - \frac{2m\mu}{\hbar^2}} \right) C_1 + \left(-\frac{1}{\xi} - \sqrt{\xi^{-2} - \frac{2m\mu}{\hbar^2}} \right) C_2 = \left(-\frac{1}{\xi} + \sqrt{\xi^{-2} + \frac{2m\mu}{\hbar^2}} \right) (C_1 + C_2). \quad (\text{A.24})$$

This yields

$$C_2 = C_1 \frac{-\hbar^2}{2m\mu} \left(\xi^{-2} - \sqrt{\xi^{-4} - \frac{4m^2\mu^2}{\hbar^2}} \right) \quad (\text{A.25})$$

$$C_3 = C_1 \left[1 - \frac{-\hbar^2}{2m\mu} \left(\xi^{-2} - \sqrt{\xi^{-4} - \frac{4m^2\mu^2}{\hbar^2}} \right) \right] . \quad (\text{A.26})$$

Thus, we find the zero energy, $E = 0$, solution in the overdamped case $\mu < m\Delta^2/(2\hbar^2)$ as

$$\begin{pmatrix} u_0(x) \\ v_0(x) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tilde{N} \begin{cases} e^{-x/\xi} \left[e^{\kappa_- x} - \hbar^2/(2m\mu) \left(\xi^{-2} - \sqrt{\xi^{-4} - 4m^2\mu^2/\hbar^2} \right) e^{-\kappa_- x} \right] & \text{for } x > 0 \\ \left[1 - \hbar^2/(2m\mu) \left(\xi^{-2} - \sqrt{\xi^{-4} - 4m^2\mu^2/\hbar^2} \right) \right] e^{-x/\xi} e^{\kappa_+ x} & \text{for } x < 0 \end{cases} , \quad (\text{A.27})$$

with $\kappa_{\pm} = \sqrt{\xi^{-2} \pm 2m\mu/\hbar^2}$ and $\xi = \hbar^2/(m\Delta)$. Here, \tilde{N} is a normalization constant. We depict this solution in Fig. A1.

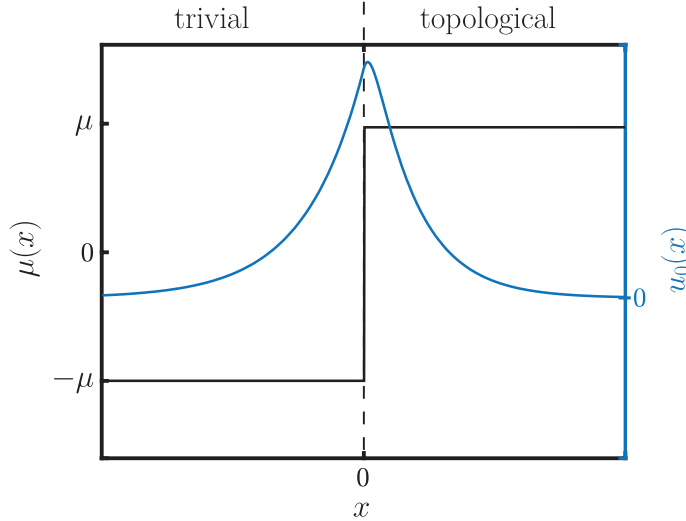


Figure A1: Zero energy wave function (blue) for the domain wall zero mode in a one dimensional spinless p-wave superconductor in the overdamped case Eq. (A.27). The black line is the chemical potential. The wave function is exponentially localized on the domain wall.

b) Underdamped case $\mu > m\Delta/(2\hbar^2)$

In this case we define abbreviations

$$K = \sqrt{\frac{2m\mu}{\hbar^2} - \xi^{-2}} \text{ and } \kappa = \sqrt{\xi^{-2} + \frac{2m\mu}{\hbar^2}} \quad (\text{A.28})$$

and obtain the general solutions

$$u_0(x > 0) = v_0(x > 0) = e^{-x/\xi} (C_1 \cos(Kx) + C_2 \sin(Kx)) \quad (\text{A.29})$$

$$u_0(x < 0) = \begin{cases} -v_0(x < 0) = e^{+x/\xi} C_3 e^{\kappa x} & \text{case i) or} \\ +v_0(x < 0) = e^{-x/\xi} C_3 e^{\kappa x} & \text{case ii) } \end{cases} . \quad (\text{A.30})$$

The continuity conditions for $u_0(x)$ and $v_0(x)$ at $x = 0$ yields in case i) $C_1 = C_3 = -C_3 = 0$ and for $u'_0(x)$ then $C_2 = 0$ such that this is not a solution. For case ii) continuity of $u_0(x)$ yields

$$C_1 = C_3 \quad (\text{A.31})$$

and for $u'_0(x)$ we find $-\xi^{-1}C_1 + KC_2 = (-\xi^{-1} + \kappa)C_3$

$$\Rightarrow C_2 = \frac{\kappa}{K}C_1. \quad (\text{A.32})$$

Thus, we get the $E = 0$ solution in the underdamped case $\mu > m\Delta^2/(2\hbar^2)$ as

$$\begin{pmatrix} u_0(x) \\ v_0(x) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tilde{N} \begin{cases} e^{-x/\xi} [\cos(Kx) + \kappa/K \sin(Kx)] & \text{for } x > 0 \\ e^{-x/\xi} e^{\kappa x} & \text{for } x < 0 \end{cases}, \quad (\text{A.33})$$

with $\xi = \hbar^2/(m\Delta)$, $K = \sqrt{2m\mu/(\hbar^2) - \xi^{-2}}$ and $\kappa = \sqrt{\xi^{-2} + 2m\mu/(\hbar^2)}$. Here, \tilde{N} is a normalization constant. The wave function is plotted in main text Fig. 2.1.

A.3 Eigenstates of the normal-conducting Rashba wire

Here we calculate the eigenvectors of the Rashba wire Hamiltonian for $\Delta = 0$ and show how they can be transformed into the form used in the main text Eq. (2.85). From the Hamiltonian

$$H = \begin{pmatrix} \tilde{k}^2 - \tilde{\mu} - \tilde{E}_z & 2\tilde{k} \\ 2\tilde{k} & \tilde{k}^2 - \tilde{\mu} + \tilde{E}_z \end{pmatrix} \quad (\text{A.34})$$

we obtain the eigenvalues

$$E_{\pm} = \tilde{k}^2 - \tilde{\mu} \pm \sqrt{\tilde{E}_z^2 + 4\tilde{k}^2}. \quad (\text{A.35})$$

The eigenvector equation for eigenvalues E_{\pm} , i.e.

$$0 = \begin{pmatrix} -\tilde{E}_z \mp \sqrt{\tilde{E}_z^2 + 4\tilde{k}^2} & 2\tilde{k} \\ 2\tilde{k} & \tilde{E}_z \mp \sqrt{\tilde{E}_z^2 + 4\tilde{k}^2} \end{pmatrix} \begin{pmatrix} u_{\uparrow} \\ u_{\downarrow} \end{pmatrix}, \quad (\text{A.36})$$

yields the relations for the components of the eigenvectors as

$$u_{\downarrow} = \frac{\tilde{E}_z \pm \sqrt{\tilde{E}_z^2 + 4\tilde{k}^2}}{2\tilde{k}} u_{\uparrow}. \quad (\text{A.37})$$

Thus, for $\tilde{E}_z > 0$ we find the eigenvectors

$$\mathbf{u}_+ = \mathcal{N} \begin{pmatrix} 1 \\ \frac{|\tilde{E}_z + \sqrt{\tilde{E}_z^2 + 4\tilde{k}^2}|}{2\tilde{k}} \end{pmatrix} \quad (\text{A.38})$$

$$\mathbf{u}_- = \mathcal{N} \begin{pmatrix} -2\tilde{k} \\ |\tilde{E}_z - \sqrt{\tilde{E}_z^2 + 4\tilde{k}^2}| \\ 1 \end{pmatrix}. \quad (\text{A.39})$$

For convenience, we define the abbreviation $x = \tilde{E}_z/(2|\tilde{k}|)$ such that the eigenvectors are given by

$$\mathbf{u}_+ = \frac{1}{\sqrt{1 + (x + \sqrt{x^2 + 1})^{-2}}} \begin{pmatrix} \text{sgn}(k)|x + \sqrt{x^2 + 1}|^{-1} \\ 1 \end{pmatrix} \quad (\text{A.40})$$

$$\mathbf{u}_- = \frac{-1}{\sqrt{1 + (x - \sqrt{x^2 + 1})^2}} \begin{pmatrix} 1 \\ -\text{sgn}(k)|x - \sqrt{x^2 + 1}| \end{pmatrix}. \quad (\text{A.41})$$

To transform them to the form in the main text, we consider

$$\frac{1}{\sqrt{1 + (x \pm \sqrt{x^2 + 1})^2}} = \sqrt{\frac{1}{1 + x^2 \pm 2x\sqrt{x^2 + 1} + x^2 + 1}} \quad (\text{A.42})$$

$$= \frac{1}{\sqrt{2}} \sqrt{\frac{1}{1 + x^2 \pm x\sqrt{x^2 + 1}}} \quad (\text{A.43})$$

$$= \frac{1}{\sqrt{2}} \sqrt{\frac{1 + x^2 \pm x\sqrt{x^2 + 1} - x^2 \mp x\sqrt{x^2 + 1}}{1 + x^2 \pm x\sqrt{x^2 + 1}}} \quad (\text{A.44})$$

$$= \frac{1}{\sqrt{2}} \sqrt{1 - \frac{x^2 \pm x\sqrt{x^2 + 1}}{1 + x^2 \pm x\sqrt{x^2 + 1}}} \quad (\text{A.45})$$

$$= \frac{1}{\sqrt{2}} \sqrt{1 - \frac{x}{\sqrt{x^2 + 1}} \frac{x^2 \pm x\sqrt{x^2 + 1}}{\sqrt{x^2 + 1} \pm x}} \quad (\text{A.46})$$

$$= \frac{1}{\sqrt{2}} \sqrt{1 \mp \frac{x}{\sqrt{x^2 + 1}}} \quad (\text{A.47})$$

$$\frac{1}{\sqrt{1 + (x \pm \sqrt{x^2 + 1})^{-2}}} = \sqrt{\frac{(x \pm \sqrt{x^2 + 1})^2}{(x \pm \sqrt{x^2 + 1})^2 + 1}} \quad (\text{A.48})$$

$$= \sqrt{\frac{2x^2 \pm 2x\sqrt{x^2 + 1} + 1}{2 + 2x^2 \pm 2x\sqrt{x^2 + 1}}} \quad (\text{A.49})$$

$$= \frac{1}{\sqrt{2}} \sqrt{\frac{2x^2 \pm 2x\sqrt{x^2 + 1} + 1 + x^2 \pm x\sqrt{x^2 + 1} - x^2 \mp x\sqrt{x^2 + 1}}{1 + x^2 \pm x\sqrt{x^2 + 1}}} \quad (\text{A.50})$$

$$= \frac{1}{\sqrt{2}} \sqrt{1 + \frac{x^2 \pm x\sqrt{x^2 + 1}}{1 + x^2 \pm x\sqrt{x^2 + 1}}} \quad (\text{A.51})$$

$$= \frac{1}{\sqrt{2}} \sqrt{1 \pm \frac{x}{\sqrt{x^2 + 1}} \frac{x \pm \sqrt{x^2 + 1}}{\sqrt{1 + x^2} \pm x}} \quad (\text{A.52})$$

$$= \frac{1}{\sqrt{2}} \sqrt{1 \pm \frac{x}{\sqrt{x^2 + 1}}}. \quad (\text{A.53})$$

Using these relations in the eigenvectors Eqs. (A.40) and (A.41) yields

$$\mathbf{u}_+ = \frac{1}{\sqrt{2}} \begin{pmatrix} \text{sgn}(k) \sqrt{1 - x/\sqrt{x^2 + 1}} \\ \sqrt{1 + x/\sqrt{x^2 + 1}} \end{pmatrix} \quad (\text{A.54})$$

$$\mathbf{u}_- = \frac{1}{\sqrt{2}} \begin{pmatrix} -\sqrt{1 + x/\sqrt{x^2 + 1}} \\ \text{sgn}(k) \sqrt{1 - x/\sqrt{x^2 + 1}} \end{pmatrix}. \quad (\text{A.55})$$

By resubstitution of $x = E_z/(2|k|)$, we obtain the eigenvectors Eq. (2.85).

A.4 Finite Majorana energy

Here we calculate the energy of the Majorana level for a finite Rashba wire of length L based on the analytical solution Eq. (2.136). As we discussed in the main text, the Majorana wave functions from both ends hybridize to

$$\psi_{\pm E_0}(y) = \frac{e^{\pm i\pi/4}}{\sqrt{2}} (\chi_L(y) \mp i\chi_R(y)), \quad (\text{A.56})$$

where we cut the wave functions at the opposite ends using a step function. The energy is then calculated by [368]

$$E_0 = \frac{\int dy \psi_{E_0}^\dagger \mathcal{H} \psi_{E_0}}{\int dy \psi_{E_0}^\dagger \psi_{E_0}} \quad (\text{A.57})$$

$$= \frac{1}{2} \frac{\int_0^L dy [\chi_L^\dagger(y) \mathcal{H} \chi_L(y) + \chi_R^\dagger(y) \mathcal{H} \chi_R(y) - i\chi_L^\dagger(y) \mathcal{H} \chi_R(y) + i\chi_R^\dagger(y) \mathcal{H} \chi_L(y)]}{\int dy \psi_{E_0}^\dagger \psi_{E_0}}. \quad (\text{A.58})$$

By using the unitary $\mathcal{P}\Pi$ symmetry, we show in the main text, that the direct terms vanish and the expression reduces to

$$E_0 = -\frac{i}{2} \frac{\int dy [\chi_L^\dagger \mathcal{H} \chi_R - \chi_R^\dagger \mathcal{H} \chi_L]}{\int dy \psi_{E_0}^\dagger \psi_{E_0}}. \quad (\text{A.59})$$

To evaluate this expression, we use that the analytical solution Eq. (2.136), truncated at the opposite ends are given by

$$\chi_L^\dagger = \theta(L - y) e^{i\pi/4} (\hat{\chi}_{L\uparrow}(y), i\hat{\chi}_{L\downarrow}(y), \hat{\chi}_{L\downarrow}(y), -i\hat{\chi}_{L\uparrow}(y)) \quad (\text{A.60})$$

$$\chi_R^\dagger = \theta(y) e^{-i\pi/4} (\hat{\chi}_{L\uparrow}(L - y), -i\hat{\chi}_{L\downarrow}(L - y), \hat{\chi}_{L\downarrow}(L - y), i\hat{\chi}_{L\uparrow}(L - y)). \quad (\text{A.61})$$

Here the $\hat{\chi}$ are all real functions. We use these to first compute the normalization where the Majorana wave functions themselves are already normalized

$$\int dy \psi_{E_0}^\dagger \psi_{E_0} = \frac{1}{2} \int dy [\chi_L^\dagger \chi_L + \chi_R^\dagger \chi_R - i\chi_L^\dagger \chi_R + i\chi_R^\dagger \chi_L] \quad (\text{A.62})$$

$$= 1 + \frac{i}{2} \int dy [\chi_R^\dagger \chi_L - \chi_L^\dagger \chi_R]. \quad (\text{A.63})$$

From Eq. (A.60) and Eq. (A.61) we find that the second term vanishes and

$$\int dy \psi_{E_0}^\dagger \psi_{E_0} = 1 . \quad (\text{A.64})$$

Therefore, the energy is given by

$$E_0 = -\frac{i}{2} \int dy \left[\chi_L^\dagger \mathcal{H} \chi_R - \chi_R^\dagger \mathcal{H} \chi_L \right] . \quad (\text{A.65})$$

Using the explicit form of the Majorana wave functions yields

$$\begin{aligned} E_0 = \frac{1}{2} \int dy & \left[\theta(L-y) (\hat{\chi}_{L\uparrow}(y), i\hat{\chi}_{L\downarrow}(y), \hat{\chi}_{L\downarrow}(y), -i\hat{\chi}_{L\uparrow}(y)) \times \right. \\ & \begin{pmatrix} -\partial_y^2 - \tilde{\mu} - \tilde{E}_z & -2i\partial_{\tilde{y}} & \tilde{\Delta} & 0 \\ -2i\partial_{\tilde{y}} & -\partial_y^2 - \tilde{\mu} + \tilde{E}_z & 0 & \tilde{\Delta} \\ \tilde{\Delta} & 0 & \partial_y^2 + \tilde{\mu} - \tilde{E}_z & 2i\partial_{\tilde{y}} \\ 0 & \tilde{\Delta} & 2i\partial_{\tilde{y}} & \partial_y^2 + \tilde{\mu} + \tilde{E}_z \end{pmatrix} \theta(y) \begin{pmatrix} \hat{\chi}_{L\uparrow}(L-y) \\ i\hat{\chi}_{L\downarrow}(L-y) \\ \hat{\chi}_{L\downarrow}(L-y) \\ -i\hat{\chi}_{L\uparrow}(L-y) \end{pmatrix} \\ & + \theta(y) (\hat{\chi}_{L\uparrow}(L-y), -i\hat{\chi}_{L\downarrow}(L-y), \hat{\chi}_{L\downarrow}(L-y), i\hat{\chi}_{L\uparrow}(L-y)) \times \\ & \left. \begin{pmatrix} -\partial_y^2 - \tilde{\mu} - \tilde{E}_z & -2i\partial_{\tilde{y}} & \tilde{\Delta} & 0 \\ -2i\partial_{\tilde{y}} & -\partial_y^2 - \tilde{\mu} + \tilde{E}_z & 0 & \tilde{\Delta} \\ \tilde{\Delta} & 0 & \partial_y^2 + \tilde{\mu} - \tilde{E}_z & 2i\partial_{\tilde{y}} \\ 0 & \tilde{\Delta} & 2i\partial_{\tilde{y}} & \partial_y^2 + \tilde{\mu} + \tilde{E}_z \end{pmatrix} \theta(L-y) \begin{pmatrix} \hat{\chi}_{L\uparrow}(y) \\ -i\hat{\chi}_{L\downarrow}(y) \\ \hat{\chi}_{L\downarrow}(y) \\ i\hat{\chi}_{L\uparrow}(y) \end{pmatrix} \right] . \end{aligned} \quad (\text{A.66})$$

Using that $\mathcal{H}\hat{\chi}_L = 0$, $\chi_L(0) = 0$, $\partial_y\theta(y) = \delta(y)$, and $\int dy(-\partial_y^2\theta(y))f(y) = f'(0)$, $\int dy(-\partial_y^2\theta(L-y))f(y) = -f'(L)$, we find

$$\begin{aligned} E_0 = \frac{1}{2} \int_0^L dy & \left[(\hat{\chi}_{L\uparrow}(y), i\hat{\chi}_{L\downarrow}(y), \hat{\chi}_{L\downarrow}(y), -i\hat{\chi}_{L\uparrow}(y)) \begin{pmatrix} -\partial_y^2\theta(y) & 0 & 0 & 0 \\ 0 & -\partial_y^2\theta(y) & 0 & 0 \\ 0 & 0 & \partial_y^2\theta(y) & 0 \\ 0 & 0 & 0 & \partial_y^2\theta(y) \end{pmatrix} \begin{pmatrix} \hat{\chi}_{L\uparrow}(L-y) \\ i\hat{\chi}_{L\downarrow}(L-y) \\ \hat{\chi}_{L\downarrow}(L-y) \\ -i\hat{\chi}_{L\uparrow}(L-y) \end{pmatrix} \right. \\ & + (\hat{\chi}_{L\uparrow}(L-y), -i\hat{\chi}_{L\downarrow}(L-y), \hat{\chi}_{L\downarrow}(L-y), i\hat{\chi}_{L\uparrow}(L-y)) \begin{pmatrix} -\partial_y^2\theta(y) & 0 & 0 & 0 \\ 0 & -\partial_y^2\theta(y) & 0 & 0 \\ 0 & 0 & \partial_y^2\theta(y) & 0 \\ 0 & 0 & 0 & \partial_y^2\theta(y) \end{pmatrix} \begin{pmatrix} \hat{\chi}_{L\uparrow}(y) \\ -i\hat{\chi}_{L\downarrow}(y) \\ \hat{\chi}_{L\downarrow}(y) \\ i\hat{\chi}_{L\uparrow}(y) \end{pmatrix} \left. \right] \\ & = 2 \left(\hat{\chi}'_{L\uparrow}(0)\hat{\chi}_{L\uparrow}(L) - \hat{\chi}'_{L\downarrow}(0)\hat{\chi}_{L\downarrow}(L) \right) . \end{aligned} \quad (\text{A.67})$$

Using the analytic expressions Eq. (2.136) in the topological case where $\xi > \xi_2$ with $e^{-L/\xi_2} \ll e^{-L/\xi}$, and the abbreviation $A = \xi + i k_{\text{eff}}$, the energy is given by

$$\begin{aligned} E_0 = 2\mathcal{N}^2 e^{-L/\xi} & \left\{ \text{Re} \left[-4A^2 a - 2\tilde{\Delta} A a - \xi_2^{-1} (2\xi_2^{-1} + \tilde{\Delta}) \right] \text{Re} \left[4e^{ik_{\text{eff}}L} A a + 2e^{ik_{\text{eff}}L} \tilde{\Delta} a \right] \right. \\ & \left. - \text{Re} \left[-2A^3 a - 2(\tilde{\mu} + \tilde{E}_z) A a - \xi_2^{-1} (\xi_2^{-2} + \tilde{\mu} + \tilde{E}_z) \right] \text{Re} \left[2e^{ik_{\text{eff}}L} A^2 a + 2e^{ik_{\text{eff}}L} (\tilde{\mu} + \tilde{E}_z) a \right] \right\} . \end{aligned} \quad (\text{A.68})$$

Using further that $\text{Re}(e^{ik_{\text{eff}}L}B) = \cos(k_{\text{eff}}L)\text{Re}(B) - \sin(k_{\text{eff}}L)\text{Im}(B)$

$$E_0 = 2\mathcal{N}^2 e^{-L/\xi} \cos(k_{\text{eff}}L) \left\{ \text{Re} \left[-4A^2a - 2\tilde{\Delta}Aa - \xi_2^{-1}(2\xi_2^{-1} + \tilde{\Delta}) \right] \text{Re} \left[4Aa + 2\tilde{\Delta}a \right] \right. \\ \left. - \text{Re} \left[-2A^3a - 2(\tilde{\mu} + \tilde{E}_z)Aa - \xi_2^{-1}(\xi_2^{-2} + \tilde{\mu} + \tilde{E}_z) \right] \text{Re} \left[2A^2a + 2(\tilde{\mu} + \tilde{E}_z)a \right] \right\} \quad (\text{A.69})$$

$$- 2\mathcal{N}^2 e^{-L/\xi} \sin(k_{\text{eff}}L) \left\{ \text{Re} \left[-4A^2a - 2\tilde{\Delta}Aa - \xi_2^{-1}(2\xi_2^{-1} + \tilde{\Delta}) \right] \text{Im} \left[4Aa + 2\tilde{\Delta}a \right] \right. \\ \left. - \text{Re} \left[-2A^3a - 2(\tilde{\mu} + \tilde{E}_z)Aa - \xi_2^{-1}(\xi_2^{-2} + \tilde{\mu} + \tilde{E}_z) \right] \text{Im} \left[2A^2a + 2(\tilde{\mu} + \tilde{E}_z)a \right] \right\} \quad (\text{A.70})$$

$$= R e^{-L/\xi} \cos(k_{\text{eff}}L - \delta) . \quad (\text{A.71})$$

This is the expression given in main text Eq. (2.160). Here, R and δ are given by

$$R = \sqrt{c_{\text{cos}}^2 + c_{\text{sin}}^2} \quad (\text{A.72})$$

and

$$\delta = \arctan \left(\frac{c_{\text{sin}}}{c_{\text{cos}}} \right) , \quad (\text{A.73})$$

with

$$c_{\text{cos}} = 2\mathcal{N}^2 \left\{ \text{Re} \left[-4A^2a - 2\tilde{\Delta}Aa - \xi_2^{-1}(2\xi_2^{-1} + \tilde{\Delta}) \right] \text{Re} \left[4Aa + 2\tilde{\Delta}a \right] \right. \\ \left. - \text{Re} \left[-2A^3a - 2(\tilde{\mu} + \tilde{E}_z)Aa - \xi_2^{-1}(\xi_2^{-2} + \tilde{\mu} + \tilde{E}_z) \right] \text{Re} \left[2A^2a + 2(\tilde{\mu} + \tilde{E}_z)a \right] \right\} \quad (\text{A.74})$$

$$c_{\text{sin}} = -2\mathcal{N}^2 \left\{ \text{Re} \left[-4A^2a - 2\tilde{\Delta}Aa - \xi_2^{-1}(2\xi_2^{-1} + \tilde{\Delta}) \right] \text{Im} \left[4Aa + 2\tilde{\Delta}a \right] \right. \\ \left. - \text{Re} \left[-2A^3a - 2(\tilde{\mu} + \tilde{E}_z)Aa - \xi_2^{-1}(\xi_2^{-2} + \tilde{\mu} + \tilde{E}_z) \right] \text{Im} \left[2A^2a + 2(\tilde{\mu} + \tilde{E}_z)a \right] \right\} . \quad (\text{A.75})$$

A.5 CMA-ES algorithm

In this appendix, we summarize the CMA-ES algorithm closely following the described by Hansen et al. in Ref. [241] and Ref. [253] for optimizing the function $f: \mathbb{R}^N \rightarrow \mathbb{R}$:

1 PARAMETERS

Get starting point $\mathbf{m}_0 \in \mathbb{R}^N$.

Get initial step size $\sigma_0 \in \mathbb{R}$.

Get population size $n_{\text{pop}} \in \mathbb{N}$.

2 INITIALIZATION

Set initial center of the search region $\mathbf{m}^{(0)} = \mathbf{m}_0$.

Set initial step size $\sigma^{(0)} = \sigma_0$.

Set initial covariance matrix as identity $C^{(0)} = \mathbb{1}$.

Set initial evolution path for the correlation matrix $\mathbf{s}_C^{(0)} = \mathbf{0}$.

Set initial evolution path for the step size $\mathbf{s}_\sigma^{(0)} = \mathbf{0}$.

Set selection size $\tilde{n} = \lfloor n_{\text{pop}}/2 \rfloor$ (number of best candidates considered for updates).

Set weights $\tilde{w}_j = \ln[(n_{\text{pop}} + 1)/(2j)]$ and $w_i^+ = \tilde{w}_i / \sum_{j=1}^{\tilde{n}} \tilde{w}_j > 0$.

Set effective number of included best metric values $\tilde{n}_{\text{eff}} = [\sum_{i=1}^{\tilde{n}} \tilde{w}_i]^2 / \sum_{i=1}^{\tilde{n}} \tilde{w}_i^2$.

Set friction parameters $c_\sigma = (\tilde{n}_{\text{eff}} + 2)/(N + \tilde{n}_{\text{eff}} + 5)$, $d_\sigma = 1 + c_\sigma + 2 \max[0, \sqrt{(\tilde{n}_{\text{eff}} - 1)/(N + 1)} - 1]$, and $c_C = (4 + \tilde{n}_{\text{eff}}/N)/(N + 4 + 2\tilde{n}_{\text{eff}}/N)$.

Set weights for the active CMA-ES step $w_i^- = \min \left\{ 1 + c_1/c_\mu, 1 + \tilde{n}_{\text{eff}}^-/(\tilde{n}_{\text{eff}} + 2), (1 - c_1 - c_\mu)/Nc_\mu \right\} \times |\tilde{w}_{i+\tilde{n}}| / \sum_{j=1}^{\tilde{n}} |\tilde{w}_{j+\tilde{n}}|$, with $c_1 = 2/((N + 1.3)^2 + \tilde{n}_{\text{eff}})$, $c_\mu = \min[1 - c_1, 2(\tilde{n}_{\text{eff}} - 2 + 1/\tilde{n}_{\text{eff}})/((N + 2)^2 + \tilde{n}_{\text{eff}})]$, and $\tilde{n}_{\text{eff}}^- = [\sum_{i=1}^{\tilde{n}} \tilde{w}_{i+\tilde{n}}]^2 / \sum_{i=1}^{\tilde{n}} \tilde{w}_{i+\tilde{n}}^2$.

Set $\hat{\chi} = \sqrt{N}(1 - 1/(4N) + 1/(21N^2))$.

Set step counter $t = 1$.

3 OPTIMIZER LOOP

- (i) Draw n_{pop} candidate vectors \mathbf{x}_i form a multivariate normal distribution $\mathcal{N}(\mathbf{m}^{(t)}, (\sigma^{(t)})^2 C^{(t)})$. This is achieved by using the eigendecomposition $C^{(t)} = UD^2U^T$ of the symmetric positive semidefinite correlation matrix $C^{(t)}$, drawing vectors \mathbf{z}_i from $\mathcal{N}(0, \mathbb{1})$ with i.i.d. normal entries with zero mean and unit variance, and computing

$$\mathbf{x}_i = \mathbf{m}^{(t)} + \sigma^{(t)} U D \mathbf{z}_i . \quad (\text{A.76})$$

- (ii) Get function values $f_i = f(\mathbf{x}_i)$ for each candidate and rank order \mathbf{x}_i such that $f_i \leq f_{i+1}$. This is the only information about the function f that the optimizer requires. Store also the current best solution $(\mathbf{x}_{\text{best}}, f_{\text{best}})$ or update it, if this generation produced a function value smaller than any value observed up to this iteration.
- (iii) Update the mean value according to the \tilde{n} best solutions as a weighted average:

$$\mathbf{m}^{(t+1)} = \sum_{i=1}^{\tilde{n}} w_i^+ \mathbf{x}_i . \quad (\text{A.77})$$

- (iv) Compute difference vectors

$$\delta \mathbf{s}_i^{(t+1)} = (\mathbf{x}_i - \mathbf{m}^{(t)}) / \sigma^{(t)} \quad (\text{A.78})$$

$$\delta \mathbf{s}^{(t+1)} = (\mathbf{m}^{(t+1)} - \mathbf{m}^{(t)}) / \sigma^{(t)} . \quad (\text{A.79})$$

- (v) Update the evolution path for the step size and from it the step size:

$$\mathbf{s}_\sigma^{(t+1)} = (1 - c_\sigma) \mathbf{s}_\sigma^{(t)} + \sqrt{c_\sigma(2 - c_\sigma) \tilde{n}_{\text{eff}}} [C^{(t)}]^{-1/2} \delta \mathbf{s}^{(t+1)} \quad (\text{A.80})$$

$$\sigma^{(t+1)} = \sigma^{(t)} \exp \left[\frac{c_\sigma}{d_\sigma} \left(\frac{\|\mathbf{s}_\sigma^{(t+1)}\|}{\hat{\chi}} - 1 \right) \right] . \quad (\text{A.81})$$

The update of the evolution path is chosen such that the scaling with the square root of the eigenvalues D of C is removed, and only the rotation is kept [241]. Here, $\hat{\chi}$ is an estimator for the expectation values of the norm of $\mathbf{s}_\sigma^{(t)}$ [241].

- (vi) Update the evolution path of the correlation matrix and the correlation matrix using the previous one and the update path

$$\mathbf{s}_C^{(t+1)} = (1 - c_C)\mathbf{s}_C^{(t)} + h(\mathbf{s}_\sigma^{(t+1)})\sqrt{c_C(2 - c_C)\tilde{n}_{\text{eff}}}\delta\mathbf{s}^{(t+1)} \quad (\text{A.82})$$

$$C^{(t+1)} = C^{(t)} + c_1\mathbf{s}_C^{(t+1)}[\mathbf{s}_C^{(t+1)}]^T + c_\mu \sum_i^{\tilde{n}} \left(w_i^+ - w_i^- \frac{N}{\|C^{-\frac{1}{2}}\mathbf{x}_i\|^2} \right) \delta\mathbf{s}_i^{(t+1)}[\delta\mathbf{s}_i^{(t+1)}]^T, \quad (\text{A.83})$$

where $h(\mathbf{s}_\sigma^{(t+1)})$ is one if $\|\mathbf{s}_\sigma^{(t+1)}\|/\hat{\chi} < (1.4 + 2)\sqrt{1 - (1 - c_\sigma)^{2(t+1)}/(N + 1)}$ and zero otherwise to avoid that the search space shrinks too fast [253]. The evolution path is updated in the direction of the change of the mean value, keeping some momentum from the last updates by including parts of the previous \mathbf{s} .

- (vii) Update $t \rightarrow t + 1$ and loop until convergence. The convergence condition can for example be based on the length of the path updates, the step size, or on the change in the best candidates for the function value compared to the previous iteration.

4 RESULT

Return the smallest function value f_{best} and corresponding point \mathbf{x}_{best} .

Here, the evolution paths $\mathbf{s}_C^{(t)}$ and $\mathbf{s}_\sigma^{(t)}$ are adjusted in each iteration, which are then used to update the covariance matrix C and step size σ indirectly, retaining information from the previous iteration. This concept is similar to gradient descent with momentum as applied to training neural networks.

B.1 Spectra and singular vectors for more networks

We show additional results for singular value spectra and Kolmogorov-Smirnov tests of singular vector entries against a Porter-Thomas distribution for networks trained with label noise in Fig. B1. We find results similar to those of Sec. 3.3. Averaged p -values for singular vectors corresponding to large singular values are reduced compared to the values in the bulk.

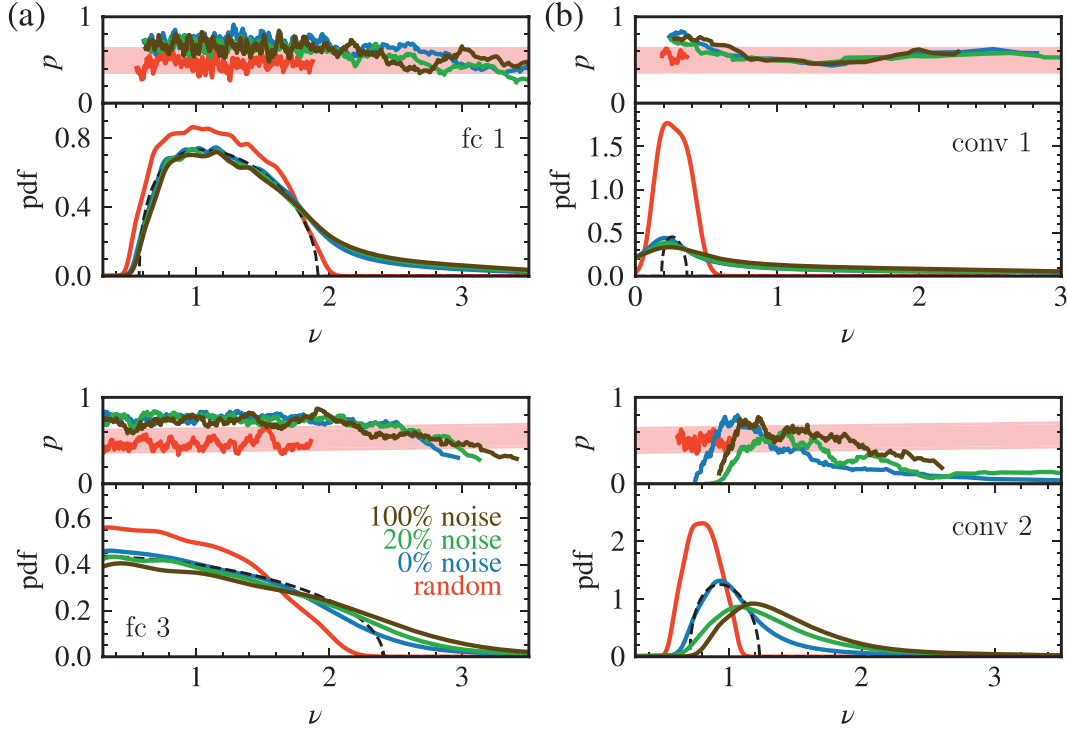


Figure B1: Analysis of singular values ν and vectors V of (a) the first and third hidden layer weight of MLP1024 networks, and (b) the convolutional layers of miniAlexNet networks trained with various amounts of label noise: 0% (blue), 40% (green), and 100% (brown). For reference, we show results for randomly initialized weights in red. The upper panels show the randomness of singular vectors via the p -value of Kolmogorov-Smirnov tests against a Porter-Thomas distribution, averaged over neighboring singular values with a window size of 31. The light red stripe describes the 2σ region around the mean for random vectors. The lower panel depicts the corresponding singular value spectra obtained via Gaussian broadening with a window size of 31 singular values (solid lines). The dashed line shows the fit of a Marcenko-Pastur distribution to the spectrum for 0% label noise.

C.1 Consistency check of the reffermionized field operator

By using $\psi_\alpha(x)$ from Eq. (1.114) and the expression for $\psi_\alpha^\dagger(x')\psi_\alpha(x') = \rho_\alpha(x')$ from Eq. (1.110), we show that $\psi_\alpha(x)$ indeed affects the density at position $x' = x$ according to [162]

$$[\psi_\alpha(x), \psi_\alpha^\dagger(x')\psi_\alpha(x')] = \delta(x - x')\psi_\alpha(x) . \quad (\text{C.84})$$

Using the commutator relation $[e^A, B] = e^A[A, B]$ that is valid if $[[A, B], A] = 0$ [162], together with $[\phi_\alpha(x), b_{\alpha,q}] = +\sqrt{2\pi/(qL)}e^{-i\alpha qx}$, $[\phi_\alpha(x), b_{\alpha,q}^\dagger] = -\sqrt{2\pi/(qL)}e^{i\alpha qx}$, and $e^{i(\varphi_{0,q} + \alpha 2\pi x N_\alpha/L)} = e^{i\varphi_{0,q}}e^{i\alpha 2\pi x N_\alpha/L}e^{-i\alpha \pi x/L}$, one finds

$$\begin{aligned} [\psi_\alpha(x), \rho_\alpha(x')] &= \frac{\chi_\alpha}{\sqrt{2\pi\eta}}e^{i\alpha \frac{2\pi x}{L}N_\alpha}e^{-i\frac{\alpha \pi x}{L}}\frac{1}{L}\left\{\sum_{q>0}\sqrt{\frac{qL}{2\pi}}\left(i\alpha e^{i\alpha qx'}e^{i\varphi_{0,\alpha}}[e^{-i\alpha \phi_\alpha(x)}, b_{\alpha,q}] \right. \right. \\ &\quad \left. \left. - i\alpha e^{-i\alpha qx'}e^{i\varphi_{0,\alpha}}[e^{-i\alpha \phi_\alpha(x)}, b_{\alpha,q}^\dagger] \right) \right. \\ &\quad \left. + e^{-i\alpha \phi_\alpha(x)}[e^{i\varphi_{0,\alpha}}, N_\alpha] \right\} \end{aligned} \quad (\text{C.85})$$

$$\begin{aligned} &= \psi_\alpha(x)\frac{1}{L}\left\{\sum_{q>0}\sqrt{\frac{qL}{2\pi}}\left(i\alpha e^{i\alpha qx'}[-i\alpha \phi_\alpha(x), b_{\alpha,q}] \right. \right. \\ &\quad \left. \left. - i\alpha e^{-i\alpha qx'}[-i\alpha \phi_\alpha(x), b_{\alpha,q}^\dagger] \right) \right. \\ &\quad \left. + [i\varphi_{0,\alpha}, N_\alpha] \right\} \end{aligned} \quad (\text{C.86})$$

$$= \psi_\alpha(x)\frac{1}{L}\left\{\sum_{q>0}\left(e^{i\alpha q(x'-x)} + e^{-i\alpha q(x'-x)}\right) + 1\right\} \quad (\text{C.87})$$

$$= \psi_\alpha(x)\frac{1}{L}\sum_q e^{i\alpha q(x-x')} \quad (\text{C.88})$$

$$= \psi_\alpha(x)\delta(\alpha(x - x')) \quad (\text{C.89})$$

$$= \psi_\alpha(x)\delta(x - x') . \quad (\text{C.90})$$

This shows the importance of the $\varphi_{0,\alpha}$ zero mode, which contributes the $q = 0$ term to the Fourier representation of the Dirac delta distribution.

C.2 Finite size corrections of the entropy in the clustered solid limit

For the clustered solid limit $V/J \rightarrow -\infty$ in a system with N particles on $L = 2N$ lattice sites, we found the eigenvalues of the 1-RDM in Eq. (4.32) as

$$\lambda_n = \frac{1}{2N} \left[1 + \frac{2}{N} \cos(q_n N) \right] \quad (\text{C.91})$$

where due to the (anti-)periodic boundary conditions for (even) odd number of fermions N , the momenta in the Fourier transform are given by

$$q_n = \begin{cases} \frac{2\pi n}{L} & \text{for odd } N \\ \frac{2\pi n}{L} + \frac{\pi}{L} & \text{for even } N \end{cases} . \quad (\text{C.92})$$

Therefore, we find

$$\lambda_n = \begin{cases} \frac{1}{2N} \left[1 + (-1)^n \frac{2}{N} \right] & \text{for odd } N \\ \frac{1}{2N} & \text{for even } N \end{cases} . \quad (\text{C.93})$$

For even particle number, the entropy reaches the thermodynamic limit value $S_\alpha(V/J \rightarrow -\infty) = \ln(2) + \ln(N)$ even for finite lattices, while for odd N , there is a finite size correction

$$S_1(V/J \rightarrow -\infty) - \ln(N) = \ln(2) - \frac{1}{2} \ln \left[1 - \frac{4}{N^2} \right] - \frac{1}{N} \ln \left[\frac{1 + 2/N}{1 - 2/N} \right] \quad (\text{C.94})$$

$$= \ln(2) - 2/N^2 + \mathcal{O}(N^{-4}) \quad (\text{C.95})$$

$$S_\alpha(V/J \rightarrow -\infty) - \ln(N) = \ln(2) + \frac{1}{1-\alpha} \ln \left[\frac{(1 + 2/N)^\alpha + (1 - 2/N)^\alpha}{2} \right] \quad (\text{C.96})$$

$$= \ln(2) - 2\alpha/N^2 + \mathcal{O}(N^{-4}) . \quad (\text{C.97})$$

Bibliographische Beschreibung

Thamm, Matthias

Mesoscopic physics of quantum systems and neural networks

Universität Leipzig, Dissertation

284 S., 497 Lit., 113 Abb.

Abstract

We study three different kinds of mesoscopic systems – in the intermediate region between macroscopic and microscopic scales consisting of many interacting constituents:

We consider particle entanglement in one-dimensional chains of interacting fermions. By employing a field theoretical bosonization calculation, we obtain the one-particle entanglement entropy in the ground state and its time evolution after an interaction quantum quench which causes relaxation towards non-equilibrium steady states. By pushing the boundaries of the numerical exact diagonalization and density matrix renormalization group computations, we are able to accurately scale to the thermodynamic limit where we make contact to the analytic field theory model. This allows to fix an interaction cutoff required in the continuum bosonization calculation to account for the short range interaction of the lattice model, such that the bosonization result provides accurate predictions for the one-body reduced density matrix in the Luttinger liquid phase.

Establishing a better understanding of how to control entanglement in mesoscopic systems is also crucial for building qubits for a quantum computer. We further study a popular scalable qubit architecture that is based on Majorana zero modes in topological superconductors. The two major challenges with realizing Majorana qubits currently lie in trivial pseudo-Majorana states that mimic signatures of the topological bound states and in strong disorder in the proposed topological hybrid systems that destroys the topological phase. We study coherent transport through interferometers with a Majorana wire embedded into one arm. By combining analytical and numerical considerations, we explain the occurrence of an amplitude maximum as a function of the Zeeman field at the onset of the topological phase – a signature unique to MZMs – which has recently been measured experimentally [Whiticar et al., Nature Communications, 11(1):3212, 2020]. By placing an array of gates in proximity to the nanowire, we made a fruitful connection to the field of Machine Learning by using the CMA-ES algorithm to tune the gate voltages in order to maximize the amplitude of coherent transmission. We find that the algorithm is capable of learning disorder profiles and even to restore Majorana modes that were fully destroyed by strong disorder by optimizing a feasible number of gates.

Deep neural networks are another popular machine learning approach which not only has many direct applications to physical systems but which also behaves similarly to physical mesoscopic systems. In order to comprehend the effects of the complex dynamics from the training, we employ Random Matrix Theory (RMT) as a zero-information hypothesis: before training, the weights are randomly initialized and therefore are perfectly described by RMT. After training, we attribute deviations from these predictions to learned information in the weight matrices. Conducting a careful numerical analysis, we verify that the spectra of weight matrices consists of a random bulk and a few important large singular values and corresponding vectors that carry almost all learned information. By further adding label noise to the training data, we find that more singular values in intermediate parts of the spectrum contribute by fitting the randomly labeled images. Based on these observations, we propose a noise filtering algorithm that both removes the singular values storing the noise and reverts the level repulsion of the large singular values due to the random bulk.

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Ich versichere, dass außer den in der Danksagung genannten, keine weiteren Personen bei der geistigen Herstellung der vorliegenden Arbeit beteiligt waren, insbesondere nicht die Hilfe eines Promotionsberaters in Anspruch genommen wurde, und weitere Personen von mir oder in meinem Auftrag weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Weiterhin versichere ich, dass die vorgelegte Arbeit weder im Inland noch im Ausland in gleicher oder in ähnlicher Form einer anderen Prüfungsbehörde zum Zwecke einer Promotion oder eines anderen Prüfungsverfahrens vorgelegt wurde.

Ich versichere außerdem, dass keine früheren, erfolglosen Promotionsversuche stattgefunden haben.

Leipzig, 28.02.2023



Matthias Thamm