LOCALITY IN QUANTUM MANY-BODY SYSTEMS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF PHYSICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Xizhi Han
February 2021

This dissertation is online at: http://purl.stanford.edu/nb059gc4353

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Sean Hartnoll, Primary Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Xiaoliang Qi

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Stephen Shenker

Approved for the Stanford University Committee on Graduate Studies.

Stacey F. Bent, Vice Provost for Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.

Abstract

How can we understand complex quantum many-body systems without finding exact solutions? Recent developments have shown that first principle constraints can provide satisfactory or even accurate estimates for quantum observables in interacting many-body systems. This dissertation summarizes a series of results centered on implications of locality in quantum dynamics and thermodynamics. We first discuss microscopic locality bounds on quantum transport, chaos and ground states in open and closed lattice models. Such locality bounds provide useful and rigorous information in strongly interacting systems. Then we consider the problem of emergent locality in holographic theories, with analytical and numerical examples in holographic tensor networks and matrix quantum mechanics. Emergent locality and causality are detected and analyzed, augmenting our microscopic understanding of holography.

Acknowledgements

I am indebted to my academic advisor, Sean Hartnoll, for his guidance and support during my graduate life. Sean has always been knowledgeable and accessible for my research questions, and considerate and attentive to my personal concerns. I, with others, also greatly admire his infectious enthusiasm and optimism, for holography and quantum physics in general. I feel fortunate that he introduced to me the field of locality bounds and matrix quantum mechanics.

I truly enjoyed my academic life at Stanford, both inside and outside the physics department. In particular I would like to thank Xiaoliang Qi for all the insightful discussions. I benefited a lot from Xiaoliang's group meetings on various topics, and truly admire his broad interests and physics intuitions. I am also grateful to Vedika Khemani for the inviting and soothing group meetings during the most difficult times in the pandemic.

I thank Stanford Graduate Fellowships for the financial support, and thank all my collaborators and physics friends for their company, conversations, collaborations and feedbacks. I have been tremendously fortunate to work with so many intelligent, thoughtful and fun people. I would also like to thank my Chinese friends, at Stanford, around the Bay Area or worldwide, for their company and the cultural comfort when we are far from home.

Lastly and most importantly, I would like to dedicate this dissertation to my parents, for their unconditional love and support. I am deeply indebted to my family for always being my source of resilience.

Contents

List of Tables

List of Figures

Introduction

Interacting quantum many-body systems are complicated in general, and often an exact solution is not necessary. Analysis of some generic principles of quantum mechanics can provide a great deal of information. A famous example would be the effective field theory, where we obtain a useful model for low-energy excitations from symmetry, locality and renormalizability, without resorting to a microscopic formulation. This dissertation summarizes my and my collaborators' explorations of consequences of locality in quantum many-body physics.

In an effective theory, disconnection with microscopic details simplifies the problem, but also introduces free parameters and obscures some quantum constraints. For example, the sound velocity of a spin model is a free parameter in the low-energy theory, while in fact it is constrained by the microscopic couplings via the Lieb-Robinson bound. Such locality bounds can offer a reasonable scale for the transport coefficients, which are otherwise difficult to estimate. In **Part I** we discuss locality constraints, in terms of microscopic parameters, on low-energy dynamics and thermodynamics. In **Chapter 1** we derive a locality upper bound on diffusion in open Markovian quantum dynamics. This type of dynamics has been used to describe cold atoms with strong external dissipation. Conceptually it also describes incoherent thermalization of closed quantum systems. The diffusion bound thus yields rigorous information about transport in these setups. Next, in **Chapter 2**, we discuss a state-dependent chaos bound from locality. As an application of the chaos bound, the temperature dependence of the butterfly velocity and the quantum Lyapunov exponent is connected with the microscopic operator scrambling. This connection sheds light on quantum chaos at low temperatures, reminiscent of the recent work by Maldacena, Shenker and Stanford. In addition to chaos, low-temperature thermodynamics (such as the ground state energy density) of local lattice models can also be effectively constrained by symmetry, locality and positivity. As an example, in **Chapter 3**, I obtain bounds on ground state observables in the Hubbard model, with a bootstrap methodology. This method gives lower bounds on ground state energies, and hence is complementary to various variational results.

In all the discussions in **Part I**, locality is manifest as a microscopic property (for example, only nearest-neighbor interactions are considered). In **Part II** we ponder the possibility that the notion of locality is emergent and approximate. This is the case in various examples of the holographic

principle. For instance in the famous AdS/CFT correspondence, one spatial dimension of the bulk is secretly encoded in the boundary quantum theory without gravity. Matrix quantum mechanics are another class of quantum models of interest, where one or more spatial dimensions emerge. The emergence of locality poses an additional challenge that we must correctly identify and organize the boundary collective degrees of freedom to recover the local dynamics in the bulk.

We start with addressing the issue of probing emergent locality and causality in **Chapter 4**. Dynamically we proposed a unifying framework of quantum causal influence to detect the emergence of causal structures in general quantum many-body systems. We have observed the expected and the exotic causal structures in the examples of unitary evolution, quantum error correction codes, quantum teleportation and holographic tensor networks. Thermodynamically we have proposed and evaluated an entanglement measure for matrix quantum mechanics states in **Chapter 5**. The entanglement entropy scales as the circumference of the region, i.e., is proportional to the measure of the entanglement cut. This provides evidence that we have the ground state of an emergent local Hamiltonian.

In **Chapter 5** and **Chapter 6** we develop two complementary numerical methods to supply ground states of matrix quantum mechanics. The ground state and the low-energy data can then be subject to the locality probes as previously discussed. In **Chapter 5** a variational quantum Monte Carlo algorithm is implemented with neural networks as the variational ansatze. The neural quantum state is favorable in matrix quantum mechanics as the spatial locality, which is often a key assumption of other variational forms, is not manifest. We have shown that the variational method is accurate in a supersymmetric matrix model, by comparing various observables with analytics and previous numerics in different limits. Another bootstrap-type algorithm, proposed in **Chapter 6**, imposes inequalities of ground state observable values, based on first principles such as symmetry and positivity. The bounds are tight in the examples of anharmonic oscillators and bosonic matrix quantum mechanics, and prove that certain bosonic matrix quantum mechanics do not have extended and commutative emergent spacetime. Two methods are complementary as variational and bootstrap methods give upper and lower bounds on the ground state energy, respectively.

The research topics discussed in this dissertation are still very much open and active. Looking forward, we would like to push the transport constraints discussed in **Part I** towards other regimes of experimental interest, such as low temperatures and low frequencies in strongly correlated electronic systems. Also the methods discussed in **Part II** are generally applicable to supersymmetric matrix quantum mechanics, and certain aspects of the dual supergravity should be accessible numerically.

Part I

Locality Constraints

Chapter 1

Locality Bound for Dissipative Quantum Transport

This chapter is essentially the same as

- Han, Xizhi, and Sean A. Hartnoll. "Locality bound for dissipative quantum transport." *Physical Review Letters* 121.17 (2018): 170601.

Abstract

We prove an upper bound on the diffusivity of a dissipative, local and translation invariant quantum Markovian spin system: $D \leq D_0 + (\alpha\, v_{\mathrm{LR}}\tau + \beta\,\xi)\, v_{\mathrm{C}}$. Here v_{LR} is the Lieb-Robinson velocity, v_{C} is a velocity defined by the current operator, τ is the decoherence time, ξ is the range of interactions, D_0 is a decoherence-induced microscopic diffusivity and α and β are precisely defined dimensionless coefficients. The bound constrains quantum transport by quantities that can either be obtained from the microscopic interactions $(D_0, v_{\mathrm{LR}}, v_{\mathrm{C}}, \xi)$ or else determined from independent local non-transport measurements (τ, α, β). We illustrate the general result with the case of a spin half XXZ chain with on-site dephasing. Our result generalizes the Lieb-Robinson bound to constrain the sub-ballistic diffusion of conserved densities in a dissipative setting.

1.1 Introduction

Quantum transport processes are at the heart of experimental studies of unconventional metals [86, 113, 41], ultracold atomic gases [225, 222, 134, 104, 52, 158] and potential spintronic systems [254, 117, 13, 30, 87]. It is crucial to have theoretical tools that connect transport observables to microscopic processes. In quasiparticle systems such as conventional metals, Fermi liquid theory and Boltzmann equations offer an excellent and well-understood handle on transport [262]. For strongly

quantum transport regimes, however, there are many fewer tools available. Controlled theoretical work with realistic interactions has largely been restricted to numerics in one spatial dimension [226, 127, 143, 31, 250].

For general ballistic systems, several important, rigorous bounds on quantum transport have been established. The Mazur-Suzuki inequality bounds the Drude weight in terms of the overlap of currents with conserved charges [165, 238, 204]. The Lieb-Robinson velocity v_{LR} bounds the propagation of linearly dispersing collective modes, such as spin waves [145, 203].

Many important quantum transport processes are diffusive rather than ballistic [98]. A lower bound on the high-temperature diffusivity has been established for certain systems with integrability or additional symmetries [205, 169]. Recently, it was argued that in general local systems, Lieb-Robinson causality requires that the diffusivity be upper bounded as $D \lesssim v_{\mathrm{LR}}^2 \tau_{\mathrm{th}}$ [96]. Here τ_{th} is a 'local thermalization time'. This relation usefully identifies key physical ingredients that constrain diffusive transport. However, it is not totally satisfactory because a numerical prefactor is undetermined and furthermore the timescale τ_{th} was not precisely defined.

In this work we prove a rigorous and precise upper bound on the diffusivity of dissipative quantum Markovian spin systems. The full result is given in (1.23) below. In the limit of long decoherence time τ, the bound takes the form $D \leq \alpha v_{\mathrm{LR}} v_{\mathrm{C}} \tau$. This expression is the dissipative counterpart of the earlier bound [96], and all quantities on the right hand side will now be precisely defined. The velocities v_{LR} and v_{C} are straightforwardly computed given a microscopic Hamiltonian while the dimensionless coefficient α and decoherence time τ can be independently and unambiguously determined from local non-transport observables. Therefore, this bound can be precisely verified in experiments. It generalizes the Lieb-Robinson bound to the diffusive behavior of conserved densities, in the context of dissipative quantum Markovian dynamics.

1.2 Translation invariant Lindbladian dynamics

Non-unitary quantum dynamics describes the quantum evolution of a dissipative system coupled to an external environment. On timescales much longer than the relaxation time of the reservoir, the dynamics can be well approximated as Markovian and hence described by the Lindblad equation [151, 38]. The final state of Lindbladian non-unitary dynamics is expected to be an infinite temperature generalized Gibbs ensemble, so our diffusive dynamics occurs close to this state.

We assume that the external bath couples locally in space to the degrees of freedom of interest, and preserves spatial translation invariance. In this case, the most general Heisenberg equation of motion for an operator $O(t)$ on a lattice takes the Lindblad form

$$\dot{O} = i \sum_{\boldsymbol{x}} [H_{\boldsymbol{x}}, O] + c \sum_{\boldsymbol{x},\alpha} \left(2 L_{\boldsymbol{x}}^{\alpha\dagger} O L_{\boldsymbol{x}}^{\alpha} - \left\{ L_{\boldsymbol{x}}^{\alpha\dagger} L_{\boldsymbol{x}}^{\alpha}, O \right\} \right), \tag{1.1}$$

where $\boldsymbol{x}$ is the lattice index and $H_{\boldsymbol{x}}$ is a term in the Hamiltonian localized near lattice site $\boldsymbol{x}$. The anticommutator $\{A, B\} = AB + BA$. The $L^{\alpha}_{\boldsymbol{x}}$ are decoherence operators localized near site $\boldsymbol{x}$ and $c \geq 0$ is the decoherence strength. It will be important that a Lieb-Robinson velocity exists for such local Lindbladian dynamics [203, 181].

Throughout, we illustrate our general formalism and results with the example of an infinite, spin-half antiferromagnetic XXZ chain with on-site dephasing:

$$H_x = X_x X_{x+1} + Y_x Y_{x+1} + \Delta Z_x Z_{x+1}\,, \quad L_x = Z_x\,. \tag{1.2}$$

Here X_x, Y_x, Z_x are Pauli matrices acting on spin $x \in \mathbb{Z}$ and $\Delta > 0$ is the anisotropy. The dephasing Lindbladian is a common phenomenological description of decoherence due to coupling to a photon or phonon bath [200]. Diffusion in this model was studied numerically in [114, 115, 171, 170, 154], and we will compare with those results. Our approach, however, is not limited to one dimensional models.

The model (1.2) conserves spin: $\sum_x \dot{Z}_x = 0$. More generally, we require a local charge operator C such that

$$\sum_{\boldsymbol{x}} \dot{C}_{\boldsymbol{x}} = 0, \tag{1.3}$$

where $C_{\boldsymbol{x}}$ is the operator C translated to site $\boldsymbol{x}$. A conserved operator in the sense of (1.3) has important consequences for the dynamics on the longest timescales, after all non-conserved operators have decayed. A single, scalar conserved operator is expected to lead to a diffusive mode with long wavelength dispersion $\omega(\boldsymbol{k}) = -iDk^2 + \ldots$, see e.g. [50]. Here D is the diffusivity and $\ldots$ denotes terms of higher order in the wavevector $\boldsymbol{k}$. Our objective in the remainder is to connect the microscopic Lindbladian dynamics (1.1) to the long wavelength hydrodynamic mode, and in this way bound the diffusivity D in terms of microscopic quantities.

To exploit the translation invariance of the dynamics, we introduce the linear space of operators $\mathcal{O}_{\boldsymbol{k}}$ with wavevector $\boldsymbol{k}$, defined to be the space of all operators O on an infinite lattice Λ such that

$$\mathcal{T}_{\boldsymbol{x}}[O] = O\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}}, \tag{1.4}$$

where $\mathcal{T}_{\boldsymbol{x}}$ translates operators by a vector $\boldsymbol{x}$. It will be useful to take the following basis of operators in $\mathcal{O}_{\boldsymbol{k}}$. Fix an origin of the lattice and a direction $\hat{\boldsymbol{k}}$ of the wavevector. We can then write the basis elements of $\mathcal{O}_{\boldsymbol{k}}$ as

$$|O_a) \equiv (O_a)_{\boldsymbol{k}} \equiv \sum_{\boldsymbol{x}} \mathcal{T}_{\boldsymbol{x}}[O_a] e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}, \tag{1.5}$$

where $\{O_a\}$ is the set of product operators that are localized in the region $\{\boldsymbol{x} \in \Lambda \,|\, \boldsymbol{x}\cdot\hat{\boldsymbol{k}} \geq 0\}$ and are not the identity at the origin [1]. We drop the $\boldsymbol{k}$ label on the $|O_a)$ to avoid clutter, this basis gives

[1] Also, the identity operator I itself doesn't contribute to $\mathcal{O}_{\boldsymbol{k}}$ for $\boldsymbol{k} \neq 0$ because then $(I)_{\boldsymbol{k}} = 0$ in (1.5)

a canonical isomorphism between the different $\mathcal{O}_{\boldsymbol{k}}$. For the example of the XXZ chain, the $\{O_a\}$ are strings of Pauli operators starting at the origin: $X_0, Y_0, Z_0, X_0X_1, X_0Y_1, \ldots, X_0I_1Y_2, \ldots$, where subscripts are lattice indices $x \in \mathbb{Z}$. The corresponding basis elements in (1.5) are then operators such as $(X_0Y_1)_k = \ldots + X_{-1}Y_0e^{ik} + X_0Y_1 + X_1Y_2e^{-ik} + X_2Y_3e^{-2ik} + \ldots$, from which it is clear that $(X_0Y_1)_k$ and $(X_1Y_2)_k$ only differ by a phase prefactor. This is why the operators must be taken to start at $x = 0$.

Translational symmetry implies that the $\mathcal{O}_{\boldsymbol{k}}$ are preserved by time evolution. Therefore, it is possible to diagonalize ∂_t in each $\boldsymbol{k}$-sector. An eigenoperator $O_{\boldsymbol{k}} \in \mathcal{O}_{\boldsymbol{k}}$ satisfies

$$\dot{O}_{\boldsymbol{k}} = -i\omega(O, \boldsymbol{k})O_{\boldsymbol{k}}, \tag{1.6}$$

for some $\omega(O, \boldsymbol{k}) \in \mathbb{C}$ and with $\operatorname{Im}\omega \leq 0$. Note that $i\partial_t$ is not Hermitian but the negative imaginary part of its eigenvalues means that time evolution is stable. Diffusion is then described by a coarse-grained charge operator $\widetilde{C}_{\boldsymbol{k}}$ that is an eigenoperator of ∂_t with

$$\omega(\widetilde{C}, \boldsymbol{k}) \equiv \Omega_{\boldsymbol{k}} = -iDk^2 + o(k^2)\,, \tag{1.7}$$

which defines the diffusivity D of the conserved charge. More generally D may depend on the direction of $\boldsymbol{k}$, and this definition works for any fixed direction of $\boldsymbol{k}$. We will obtain the operator $\widetilde{C}_{\boldsymbol{k}}$ explicitly below.

We are able to discuss diffusion as an operator equation, as in (1.6) and (1.7) above, because decoherence causes operator norms to decay. This is a significant technical simplification relative to the case of unitary evolution at finite temperature, where diffusion only occurs within thermal expectation values. In the following section we compute $\Omega_{\boldsymbol{k}}$ in small k perturbation theory. This will give an explicit expression for D.

1.3 Perturbation theory at small wavevector

At small k, we can expand $\partial_t|_{\mathcal{O}_{\boldsymbol{k}}}$ in k. Fixing a direction of $\boldsymbol{k}$:

$$\partial_t|_{\mathcal{O}_{\boldsymbol{k}}} = \mathcal{L} \equiv \sum_{n\geq 0} k^n \mathcal{L}^n, \tag{1.8}$$

which defines superoperators $\mathcal{L}^n$. For example, in the XXZ chain, the operator $(Z_0)_k \in \mathcal{O}_k$ obeys

$$(\dot{Z}_0)_k = 2(e^{-ik} - 1)(X_0Y_1)_k - 2(e^{-ik} - 1)(Y_0X_1)_k\,. \tag{1.9}$$

With respect to the Pauli string basis, $\partial_t|_{\mathcal{O}_k}$ is represented as a k-dependent matrix. Expanding the coefficients of the basis elements in (1.9) at small k we obtain components of the superoperators in

(1.8). As expected from conservation of Z, $\mathcal{L}^0|Z_0) = 0$, while $\mathcal{L}^1|Z_0) = 2i|Y_0X_1) - 2i|X_0Y_1)$ and $\mathcal{L}^2|Z_0) = |Y_0X_1) - |X_0Y_1)$.

The eigenvalue $-i\Omega_{\boldsymbol{k}}$ of $\partial_t|_{\mathcal{O}_{\boldsymbol{k}}}$ can be found using standard second-order perturbation theory in small $\boldsymbol{k}$ (cf. the memory matrix formalism [77]). At $\boldsymbol{k} = 0$ we know that the eigenvector is the conserved charge $|C)$, with vanishing eigenvalue. Therefore, up to order k^2:

$$-i\Omega_{\boldsymbol{k}} = k(C|\mathcal{L}^1|C) + k^2 \left[(C|\mathcal{L}^2|C) - \sum_{E_a^0 \neq 0} (C|\mathcal{L}^1|E_a^0) \frac{1}{E_a^0} (E_a^0|\mathcal{L}^1|C) \right] . \tag{1.10}$$

The basis vectors $|E_a^0)$ are given by linear combinations of the $|O_a)$ in (1.5) with $\mathcal{L}^0|E_a^0) = E_a^0|E_a^0)$. The corresponding eigenoperator $|\widetilde{C}) = |C) - k\sum_{E_a^0\neq 0}|E_a^0)\frac{1}{E_a^0}(E_a^0|\mathcal{L}^1|C)$ is the dressed charge operator to this order. We are assuming that the only operator with $E_a^0 = 0$ is the single conserved charge C. It is straightforward to extend our analysis to a finite number of conserved charges. We will be more precise about the absence of additional slow operators in the following section. The superoperator ∂_t is not antihermitian in general and the eigenoperators $|E_a^0)$ are not necessarily orthogonal. The above perturbation theory formulae retain their standard form, but $(E_a^0|O)$ is defined to be the coefficient in front of $|E_a^0)$ in the expansion of $|O)$ in the basis $\{|E_a^0)\}$. The $(E_a^0|$ are elements of a dual vector space to that spanned by the $|E_a^0)$, and hence have opposite dimensionality. In the case that the operators E_a^0 are orthogonal, i.e. $\mathrm{tr}(E_a^{0\dagger}E_b^0) = \delta_{ab}\mathrm{tr}(E_a^{0\dagger}E_a^0)$ for some given $\boldsymbol{k}$, then we can write $(E_a^0|O) = \mathrm{tr}(E_a^{0\dagger}O)/\mathrm{tr}(E_a^{0\dagger}E_a^0)$, as usual.

Our main objective is to use the expression (1.10) to bound the diffusivity (1.7) in generality. However, in simple models such as the dissipative XXZ chain (1.2) it is possible to compute the diffusivity by evaluating (1.10). The on-site dephasing in that model suppresses Pauli strings with X and Y terms. For example: $\mathcal{L}_{\mathrm{dis}}[X_0] = -4X_0$ and $\mathcal{L}_{\mathrm{dis}}[X_0Y_1] = -8X_0Y_1$, where $\mathcal{L}_{\mathrm{dis}}[O]$ is the second sum in (1.1). The explicit computation is easiest in the limit $c \gg 1$, where the Hamiltonian term in (1.1) is negligible compared to the dephasing term. In this limit (1.10) becomes (to leading order in k and c^{-1})

$$-i\Omega_{\boldsymbol{k}} = \frac{k^2}{8c}\sum_A (Z_0|\mathcal{L}^1|A)(A|\mathcal{L}^1|Z_0) = -\frac{k^2}{c}\,, \tag{1.11}$$

where in the sum $A = X_0Y_1$ and Y_0X_1. The system is diffusive with $D = c^{-1} + O(c^{-2})$ for strong decoherence $c \gg 1$, cf. [116]. This asymptotic behavior is verified numerically in Fig. 1.1, showing numerical results for finite c.

The numerical results are obtained using a truncated space of operators in (1.5) to evaluate (1.10). This is a different method compared to previous work, and is relatively straightforward to implement. It works best for larger values of c where long operators are strongly suppressed by dissipation. These results on the XXZ example agree with those in the literature [114, 115, 171, 170, 154]. In particular,

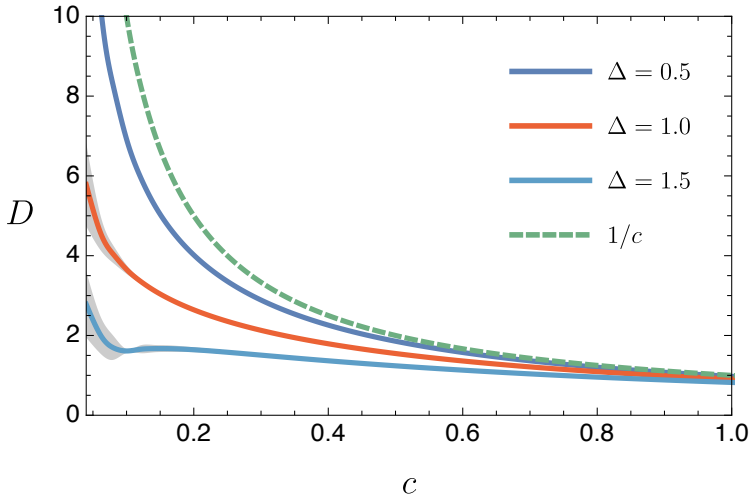


Figure 1.1: Diffusivity D of the dissipative XXZ model versus dephasing strength c, with anisotropies $\Delta = 0.5, 1.0, 1.5$. The asymptotic behavior $D \sim 1/c$ is also shown. Operator spaces are truncated in numerics so that only Pauli string operators of length at most $n = 7$ are kept. Finite-size effects are strong for small c and indicated by the shaded region, which is estimated from truncations with $n = 6, 8$.

for $0 < \Delta < 1$ the system is known to show ballistic spin transport in the absence of dephasing ($c = 0$). Therefore, while transport is diffusive at nonzero dephasing, the diffusivity diverges as $c \to 0$.

1.4 Constraints from the Lieb-Robinson bound

For diffusive rather than ballistic transport, $(C|\mathcal{L}^1|C)$ must vanish in (1.10). Indeed, $|J) = \mathcal{L}^1|C)$ is the current operator, and it is known from the Mazur-Suzuki bound [165, 238] that if $(C|J) \neq 0$ at $k = 0$, transport is ballistic. We restrict attention to non-ballistic systems [2]. Then the diffusivity can be rewritten as

$$D = -\left[(C|\mathcal{L}^2|C) + \int_0^\infty dt\, (C|\mathcal{L}^1 e^{\mathcal{L}^0 t}\mathcal{L}^1|C)\right], \tag{1.12}$$

where $\frac{1}{E_a^0}|E_a^0)(E_a^0|$ in (1.10) has been replaced by an integral of $\exp\left(\mathcal{L}^0 t\right)$. D is manifestly real in (1.12) because in a basis of hermitian operators $\mathcal{L}^0, \mathcal{L}^2$ are real matrices and $\mathcal{L}^1$ is imaginary, and furthermore $(C|$ is a real vector.

To isolate the dynamics of the single conserved density we make a physical assumption about the spectrum of the Lindbladian decoherence operators: all local operators other than the charge density decay exponentially, at least as fast as $e^{-t/\tau}$. Here τ defines the 'local decoherence time'. The local

[2] In the ballistic case, our argument gives a bound on the rate of attenuation of the linearly-dispersing mode

difference in behavior between conserved densities and all other operators will be important for our argument. Technically, we will require *single-mode ansatz:* There exist $A, \tau > 0$ such that any local operator O can be decomposed into local operators $O = \gamma I + O_1 + O_2$, where I is the identity operator and γ a coefficient, O_1 is a sum of C's and $\|O_2(t)\| \leq A\|O\|e^{-t/\tau}$, $\|\dot{O}_2(t)\| \leq A\tau^{-1}\|O\|e^{-t/\tau}$ for $t > 0$. We will bound the diffusivity by the Lieb-Robinson velocity and the decoherence time τ.

Let $\|\cdot\|$ be any operator norm contracted by the time evolution (1.1) [3]. This induces a seminorm (with $\||I)\| = 0$) for $|O) \in \mathcal{O}_0$:

$$\||O)\| \equiv \lim_{\boldsymbol{k}\to 0} \lim_{N\to\infty} N^{-1} \| \sum_{\boldsymbol{x}} \mathcal{T}_{\boldsymbol{x}}[O] e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \|, \tag{1.13}$$

where $N = \sum_{\boldsymbol{x}} 1$ is the number of lattice sites. For example,

$$\||Z)\| = \lim_{k\to 0} \lim_{N\to\infty} N^{-1} \| \sum_{x} Z_x e^{-ikx} \| = \|Z\|, \tag{1.14}$$

and generally $\||O)\| \leq \|O\|$, by the triangle inequality. From the definition (1.13), this seminorm is also contracted by time evolution. As a result of contraction in time combined with the single-mode ansatz:

$$\||O)\| \geq \lim_{t\to\infty} \||O(t))\| = |(C|O)|\||C)\|, \tag{1.15}$$

bounding the norm of the $\boldsymbol{k} = 0$ state by its projection onto the conserved charge.

We use (1.15) to bound the two terms in the diffusivity (1.12). For the first term, let $|O) = \mathcal{L}^2|C) \in \mathcal{O}_0$. Then

$$|(C|\mathcal{L}^2|C)| \leq \|\mathcal{L}^2|C)\| / \||C)\| \,. \tag{1.16}$$

Given the operator equation of motion, the right-hand side of (1.16) is easily calculable.

To bound the second term in (1.12), take a local operator O such that $|O) = \mathcal{L}^1|C) + \alpha|C) \in \mathcal{O}_0$, with $\alpha \in \mathbb{C}$. Then:

$$(C|\mathcal{L}^1 e^{\mathcal{L}^0 t}\mathcal{L}^1|C) = (C|\mathcal{L}^1 e^{\mathcal{L}^0 t}|O) = (C|\lim_{k\to 0} \partial_k(\mathcal{L}e^{\mathcal{L}t})|O) = (C|\lim_{k\to 0} \partial_k|\dot{O}(t)). \tag{1.17}$$

The first equality uses $\mathcal{L}^0|C) = 0$ and $(C|\mathcal{L}^1|C) = 0$; the second equality uses $(C|\mathcal{L}^0 = 0$. In (1.17), ∂_k is defined to be the k-derivative of the components of matrices such as $\mathcal{L}$ or vectors such as $|O)$ in the prescribed basis $|O_a)$ in (1.5). Explicitly, for any local operator P we can uniquely write $P = cI + \sum c^a_{\boldsymbol{x}} \mathcal{T}_{\boldsymbol{x}}[O_a]$ so that in $\mathcal{O}_{\boldsymbol{k}}$, $|P) = \sum c^a_{\boldsymbol{x}} e^{i\boldsymbol{k}\cdot\boldsymbol{x}}|O_a)$ and $-i\partial_k|P) = \sum c^a_{\boldsymbol{x}}(\hat{\boldsymbol{k}}\cdot\boldsymbol{x})|\mathcal{T}_{\boldsymbol{x}}[O_a])$, which is seen to be the first moment of the operator P. Using (1.17) in (1.15) gives the bound

$$\left| \int_0^\infty dt\, (C|\mathcal{L}^1 e^{\mathcal{L}^0 t}\mathcal{L}^1|C) \right| \leq \int_0^\infty dt\, \frac{\|\partial_k|\dot{O}(t))\|}{\||C)\|}. \tag{1.18}$$

[3] Completely positive unital maps — such as Lindbladian time evolution — between C*-algebras contract C*-norms, see Chap 8 of [183] and Chap 3 of [197].

Take an operator J (the current) localized near the origin such that $|J) = \mathcal{L}^1|C)$. According to the single-mode ansatz we can write $J = O + \sum_{\boldsymbol{x}} c_{\boldsymbol{x}} \mathcal{T}_{\boldsymbol{x}}[C]$, where O is also localized near the origin and for $t > 0$

$$\|\dot{O}(t)\| \leq A\tau^{-1}\|J\|e^{-t/\tau}. \tag{1.19}$$

We can choose this O as the operator in (1.17). From the bound (1.19) on $\|\dot{O}(t)\|$ we must now obtain a bound on $\|\partial_k|\dot{O}(t))\|$, that appears in (1.18).

Let $\mathcal{P}_l$ for $l \in \mathbb{R}$ be the projection onto the operator subspace spanned by all product operators supported on the half-space $\hat{\boldsymbol{k}} \cdot \boldsymbol{x} \geq l$ and let $\mathcal{Q}_l = \mathrm{Id} - \mathcal{P}_l$. By an adaption [4] of the Lieb-Robinson bound [203, 181] for (1.19), there exist $A' \geq 1$ and $v, \xi > 0$ such that for all $l \in \mathbb{R}, t > 0$,

$$\|\mathcal{P}_l[\dot{O}(t)]\| \leq A\|J\|\tau^{-1}\min\{e^{-t/\tau}, A'e^{(vt-l)/\xi}\}, \tag{1.20}$$

$$\|\mathcal{Q}_l[\dot{O}(t)]\| \leq 2A\|J\|\tau^{-1}\min\{e^{-t/\tau}, A'e^{(vt+l)/\xi}\}. \tag{1.21}$$

The length ξ is the range of microscopic interactions. We saw that ∂_k corresponds to taking the first moment. Therefore $\partial_k|\dot{O}(t)) = i|O'(t))$, with

$$O'(t) = \int_0^\infty dl\, \mathcal{P}_l[\dot{O}(t)] - \int_{-\infty}^0 dl\, \mathcal{Q}_l[\dot{O}(t)]. \tag{1.22}$$

Indeed, from $\mathcal{P}_l[\mathcal{T}_{\boldsymbol{x}}[O_a]] = \mathcal{T}_{\boldsymbol{x}}[O_a]$ for $\hat{\boldsymbol{k}} \cdot \boldsymbol{x} \geq l$, and vanishing otherwise, we have $\int_0^\infty dl\, \mathcal{P}_l[\mathcal{T}_{\boldsymbol{x}}[O_a]] = \hat{\boldsymbol{k}} \cdot \boldsymbol{x}\, \mathcal{T}_{\boldsymbol{x}}[O_a]$ if $\hat{\boldsymbol{k}} \cdot \boldsymbol{x} \geq 0$, which is precisely the first moment. The second integral of $\mathcal{Q}_l$ similarly takes care of the $\hat{\boldsymbol{k}} \cdot \boldsymbol{x} \leq 0$ terms. Now, $\|\partial_k|\dot{O}(t))\| \leq \|O'(t)\|$ and, using (1.20) and (1.21), $\|O'(t)\| \leq 3A\|J\|e^{-t/\tau}\tau^{-1}[vt + \xi(1 + t/\tau + \ln A')]$. Hence, substituting into (1.18),

$$\left|\int_0^\infty dt\, (C|\mathcal{L}^1 e^{\mathcal{L}^0 t}\mathcal{L}^1|C)\right| \leq 3A[v\tau + \xi(2 + \ln A')]\frac{\|J\|}{\||C)\|}.$$

Putting the results together gives the diffusivity bound

$$D \leq D_0 + (\alpha\, v_{\mathrm{LR}}\tau + \beta\, \xi)\, v_{\mathrm{C}}\,. \tag{1.23}$$

Here $D_0 = \|\mathcal{L}^2|C)\|/\||C)\|$ is a 'microscopic' diffusivity from the dissipative equation of motion. The Lieb-Robinson velocity $v_{\mathrm{LR}} = v$ and $v_{\mathrm{C}} = \|J\|/\||C)\|$ is a velocity obtained by dividing the current by the charge. As above τ is the decoherence time and ξ is the range of microscopic interactions. The dimensionless coefficients $\alpha = 3A$ and $\beta = 3A(2 + \ln A')$. Equation (1.23) establishes that the diffusivity is bounded by microscopic velocities and time and lengthscales in the system.

[4] By [178], for any $l \in \mathbb{R}$ and $t > 0$ there exists $\tilde{O}$ localized in $\hat{\boldsymbol{k}} \cdot \boldsymbol{x} < l$ such that $\|\dot{O}(t) - \tilde{O}\| \leq A'\|\dot{O}\|e^{(vt-l)/\xi}$, hence $\|\mathcal{P}_l[\dot{O}(t)]\| = \|\mathcal{P}_l[\dot{O}(t) - \tilde{O}]\| \leq A'\|\dot{O}\|e^{(vt-l)/\xi}$. Note $\|\mathcal{P}_l[\dot{O}(t)]\| \leq \|\dot{O}(t)\|$ as well, hence $\|\mathcal{P}_l[\dot{O}(t)]\| \leq \min\{A'\|\dot{O}\|e^{(vt-l)/\xi}, \|\dot{O}(t)\|\}$ and plug in (1.19) to obtain (1.20). Similarly for (1.21) except that $\|\mathcal{Q}_l[\dot{O}(t)]\| = \|\dot{O}(t) - \mathcal{P}_l[\dot{O}(t)]\| \leq 2\|\dot{O}(t)\|$.

The quantities D_0, v_{LR}, v_{C} and ξ can be obtained from the equations of motion. The quantities A and τ are instead best determined experimentally or numerically from the decay of local non-conserved operators. From equation (1.19), τ determines the late time decay rate of the non-conserved part of the local current and $A = \max_{t>0} \|\dot{O}(t)\|/(\tau^{-1}\|J\|e^{-t/\tau})$. A' does not have a strong effect as it appears in a logarithm in our bound.

1.5 Final comments

The bound (1.23) has nontrivial consequences for the dephasing XXZ chain. For $0 < \Delta < 1$ the diffusivity diverges in Fig. 1.1 as $c \to 0$. The bound states that D cannot diverge faster than τ. In the XXZ model $v_{\mathrm{C}} = 4$ and, from [181], $v_{\mathrm{LR}} \leq 2 + \Delta$ are independent of c. Now $\tau = \max_k 1/(-\mathrm{Re}\, E_k^1)$, where $E_k^1 \in \mathbb{C}$ is the first eigenvalue of $\partial_t|_{\mathcal{O}_k}$ above the slow mode. We evaluated this eigenvalue numerically by truncating the operator space as described around Fig. 1.1. At $\Delta = 0.5$ the ratio $D/\tau = 3.8(2)$ indeed remains finite as $c \to 0$.

We end with some broader comments. Firstly, (exponential) locality of interactions and a finite decoherence time are essential, as otherwise there can be superdiffusive transport [222, 104, 154], where the perturbation theory (1.10) is no longer valid due to degeneracies or divergences.

The decoherence-induced decay of operators such as long Pauli strings is phenomenologically similar to the decay of the thermal expectation values of those operators. To obtain a rigorous bound on diffusion in unitary quantum dynamics in a thermal state, however, there will be several challenges to overcome. The diffusivity must be discussed in terms of expectation values rather than operators, and projections with respect to thermal inner products are difficult to evaluate (e.g. §5.6 of [97]). The butterfly velocity may causally constrain finite temperature transport [96, 157, 230, 217, 218], but a temperature-dependent bound on this velocity has not been established. These interesting problems are left for future work.

If a rigorous bound of the form $D \lesssim v^2\tau + v\xi$ can indeed be established for diffusion in finite temperature states, it may shed light on the phenomenon of resistivity saturation [86, 113]. As temperature is increased τ will typically descrease, but ξ is a microscopic and temperature-independent lengthscale. Therefore, the resistivity $\rho \propto 1/D \gtrsim 1/(v^2\tau + \xi v)$ is able to saturate at high temperatures where $v\tau < \xi$.

Acknowledgements

The work of SAH is partially supported by DOE award DE-SC0018134. XH is supported by a Stanford Graduate Fellowship.

Chapter 2

State Dependence of the Butterfly Velocity

This chapter is essentially the same as

- Han, Xizhi, and Sean A. Hartnoll. "Quantum scrambling and state dependence of the butterfly velocity." *SciPost Phys* 7 (2019): 045.

Abstract

Operator growth in spatially local quantum many-body systems defines a scrambling velocity. We prove that this scrambling velocity bounds the state dependence of the out-of-time-ordered correlator in local lattice models. We verify this bound in simulations of the thermal mixed-field Ising spin chain. For scrambling operators, the butterfly velocity shows a crossover from a microscopic high temperature value to a distinct value at temperatures below the energy gap.

2.1 Introduction

Strongly quantum many-body systems have been important in condensed matter [236, 141] and nuclear physics [36, 232] for some time and are likely to become increasingly important with the ongoing development of quantum information processing technology [120, 16, 29]. It is essential to understand the spatio-temporal dynamics of these systems in highly quantum regimes where semiclassical methods such as the Boltzmann equation are inapplicable.

Significant progress has been made recently by considering quantum scrambling in many-body systems [228, 103, 229, 231, 109, 220]. Quantum scrambling arises when operator growth under Heisenberg time evolution redistributes local information to non-local degrees of freedom. It has been found that scrambling in spatially local systems is characterized by both a rate and a velocity,

e.g. [217, 162, 219, 173]. These universal properties are manifested in the so-called out-of-time-ordered correlator (OTOC):

$$\mathcal{C}(\boldsymbol{x}, t; \rho) \equiv \mathrm{tr}\left(\rho\,[O_1(0,t), O_2(\boldsymbol{x},0)]^\dagger [O_1(0,t), O_2(\boldsymbol{x},0)]\right), \tag{2.1}$$

defined for local operators O_1, O_2 in state ρ. The OTOC has been found to reveal a 'light cone' spread of quantum information, with two state-dependent characteristics: the quantum Lyapunov exponent λ and the butterfly velocity v_B. Just outside the light cone (or 'butterfly cone') $|\boldsymbol{x}| \gtrsim v_B t$ for $t > 0$, the OTOC grows as the front is approached according to [128, 256]:

$$\mathcal{C}(\boldsymbol{x}, t; \rho) \sim e^{-\lambda(|\boldsymbol{x}-\boldsymbol{x}_0|/v_B - t)^{1+p}/t^p}. \tag{2.2}$$

In systems with many local degrees of freedom (e.g. large N systems) the exponent $p = 0$ and the growth is exponential. This case is reminiscent of the classical butterfly effect. In spin lattice systems, generally $p > 0$, so that the front broadens as it spreads.

The butterfly velocity is a state-dependent speed of information propagation that is universally present in local systems, plausibly controlling important physical processes such as transport in strongly quantum regimes [28, 85, 61, 196, 96, 157, 32]. The state dependence means that the butterfly velocity is a more powerful probe of dynamics than the widely employed microscopic Lieb-Robinson velocity [146]. In this work we will show that this state dependence (e.g. temperature dependence) is tied to the underlying quantum scrambling of operators.

In quantum field theories that describe a nontrivial (quantum critical) continuum limit of lattice systems, the scaling of the butterfly velocity with temperature is $v_B \sim T^{1-1/z}$ in the simplest cases [219, 28]. The dynamical critical exponent z describes the relative scaling of space and time. In this work we will characterize the butterfly velocity in general lattice models, away from critical points and without a large N limit. We will obtain the temperature dependence of the butterfly velocity in quantum spin systems, extending previous infinite temperature results [142, 256]. The temperature dependence of scrambling in classical spin systems has been recently discussed in [27].

In a spatially local system the growth of operators determines a 'scrambling velocity' v_S, defined in (2.8) below. Our first result (2.9) states that the change of the velocity-dependent Lyapunov exponent — defined shortly in (2.6) — with temperature is bounded by the scrambling velocity. This result is rigorous for one-dimensional systems and plausibly true more generally. We verify the bound in numerical simulations of the mixed-field Ising model, focusing on the temperature dependence of the butterfly velocity. In Fig. 2.2 below we see that the non-interacting transverse field model has a temperature-independent butterfly velocity whereas the velocity is temperature-dependent for the interacting mixed field models. In these curves, the butterfly velocity crosses over from a microscopic infinite-temperature value to a low-temperature value. The temperature scale of the crossover is set by the energy gap.

2.2 Three velocities from locality

It will be crucial to understand three different velocities that characterize spatially local quantum systems. Our results will tie these velocities together. The velocities emerge in any lattice Λ of spins (or fermions) with a local Hamiltonian

$$H = \sum_{\boldsymbol{x}\in\Lambda} h_{\boldsymbol{x}}, \tag{2.3}$$

where $h_{\boldsymbol{x}}$ are operators localized near lattice site $\boldsymbol{x}$. Translation symmetry is not required.

2.2.1 Lieb-Robinson velocity

The Lieb-Robinson velocity defines an emergent 'light-cone' causality from local dynamics on a lattice [146]. It is a state-independent, microscopic velocity set by the magnitude of couplings in the Hamiltonian, and is insensitive to operator growth or lack thereof.

A convenient and powerful definition of v_{LR} is in terms of space-time rays. That is, consider an operator O_2 located along the ray $\boldsymbol{x} = vt\boldsymbol{n}$ (here $\boldsymbol{n}$ is a unit vector). At large times we can introduce a velocity-dependent exponent $\lambda(v)$ that determines the growth or decay of the norm of the commutator along the ray, $\|[O_1(0,t), O_2(vt\boldsymbol{n},0)]\| \sim e^{\lambda(v)t}$. Here $O(\boldsymbol{x},t)$ denotes O translated by a lattice vector $\boldsymbol{x}$ in space and a time t with Heisenberg evolution, and $\|\cdot\|$ is the operator norm. The causal light cone defined by v_{LR} is such that for all $v > v_{\mathrm{LR}}$ the norm decays exponentially at late times, so that $\lambda(v) < 0$. Therefore we can define v_{LR} as the largest velocity such that the norm does not decay along a ray:

$$v_{\mathrm{LR}} \equiv \sup\left\{ v : \lim_{t\to\infty} \frac{1}{t} \ln \|[O_1(0,t), O_2(vt\boldsymbol{n},0)]\| \geq 0 \right\}. \tag{2.4}$$

We shall not keep the dependence on direction $\boldsymbol{n}$ and operators O_1, O_2 explicit.

For any $v > v_{\mathrm{LR}}$ there are (v-dependent) constants $\xi_{\mathrm{LR}}, C_{\mathrm{LR}} > 0$ such that for all $t, x > 0$,

$$\|[O_1(0,t), O_2(x\boldsymbol{n},0)]\| \leq C_{\mathrm{LR}} \|O_1\| \|O_2\| e^{(vt-x)/\xi_{\mathrm{LR}}}. \tag{2.5}$$

Intuitively, inequality (2.5) states that for $v > v_{\mathrm{LR}}$, the norm $\|[O_1(0,t), O_2(x\boldsymbol{n},0)]\|$ is exponentially small outside the ray $x = vt$, with a tail of length $\xi_{\mathrm{LR}}(v)$.

2.2.2 Butterfly velocity

The butterfly velocity is defined analogously to the Lieb-Robinson velocity, but using the OTOC instead of the operator norm of the commutator [229, 217]. It therefore depends on the quantum state ρ.

The 'velocity-dependent Lyapunov exponent' is defined by the late time growth or decay of the

OTOC along a ray [128]:

$$\lambda(\boldsymbol{v};\rho) \equiv \lim_{t\to\infty} \frac{1}{t} \ln \mathcal{C}(\boldsymbol{v}t, t; \rho)\,. \tag{2.6}$$

Analogously to the Lieb-Robinson case, the butterfly velocity can now be defined as

$$v_B(\rho) \equiv \sup\left\{v : \lambda(v\boldsymbol{n};\rho) \geq 0\right\}\,, \tag{2.7}$$

which is state-dependent. The operator norm bounds the OTOC and hence $0 \leq v_B(\rho) \leq v_{\mathrm{LR}}$.

2.2.3 Scrambling velocity

The Lieb-Robinson bound (2.5) implies that the size of an operator can grow at most polynomially in time (as t^d in a d-dimensional system). In contrast, the growth can be exponential without spatial locality, such as in SYK models [163, 215, 207]. Operator growth under Heisenberg evolution in quantum systems with a local Hamiltonian will therefore define another velocity. We will call this the 'scrambling velocity' v_S. For example, in strongly scrambling models, such as random unitary circuits [249, 182, 129, 211], generic operators quickly grow into a superposition of product operators with radius $\sim v_{\mathrm{LR}}t$. In this case $v_S = v_{\mathrm{LR}}$.

More precisely, we define the scrambling velocity as follows. Given local operators O_1 and O_2, the commutator $[O_1(0,t), O_2(\boldsymbol{x},0)]$ will grow along the ray $\boldsymbol{x} = \boldsymbol{v}t$. We are interested in the growth of the operator itself rather than its norm or OTOC. Let $R(\boldsymbol{x},t)$ be the radius of support of the commutator[1] and define

$$v_S(\boldsymbol{v}) \equiv \lim_{t\to\infty} \frac{R(\boldsymbol{v}t, t)}{t}. \tag{2.8}$$

This is a velocity-dependent velocity because the growth of the operator can depend on the ray that we follow, just like the exponents in (2.4) and (2.6) above. This operator growth is illustrated in Fig. 2.1.

In the random circuit, let O_1 and O_2 be two single-site operators. Inside the Lieb-Robinson cone, i.e. for $|\boldsymbol{x}| \leq v_{\mathrm{LR}}t$, the commutator $[O_1(0,t), O_2(\boldsymbol{x},0)]$ has the same support as $O_1(0,t)$ so $R(\boldsymbol{x},t) = v_{\mathrm{LR}}t$ and $v_S(\boldsymbol{v}) = v_{\mathrm{LR}}$ for $|\boldsymbol{v}| \leq v_{\mathrm{LR}}$. For general systems and for $|\boldsymbol{v}| \leq v_{\mathrm{LR}}$ we expect that $0 \leq v_S(\boldsymbol{v}) \leq v_{\mathrm{LR}}$. A proof of this statement, along with more precise definitions and technical details, is collected in the appendices.

The definition (2.8) also captures the absence of scrambling in non-interacting theories. A non-interacting field obeys $\phi(\boldsymbol{x},t) = \int d\boldsymbol{y}\, f(\boldsymbol{y},\boldsymbol{x};t)\phi(\boldsymbol{y},0)$, for some function $f(\boldsymbol{y},\boldsymbol{x};t)$. Although the support of the operator $\phi(\boldsymbol{x},t)$ spreads out as t increases, it remains a superposition of local operators. Consider the conjugate pair (ϕ,π). It follows that $[\phi(0,t), \pi(\boldsymbol{x},0)] = if(\boldsymbol{x},0;t)$. This is a c-number and its support has radius $R(\boldsymbol{x},t) = 0$. Hence $v_S(\boldsymbol{v}) = 0$ for any $\boldsymbol{v}$.

[1] The radius of an operator O is the minimal distance R such that O is supported in a ball (centered at an arbitrary site) of radius R. Throughout the main text 'support' should be understood as up to an exponentially decaying tail. Exponential tails are discussed in detail in the appendices.

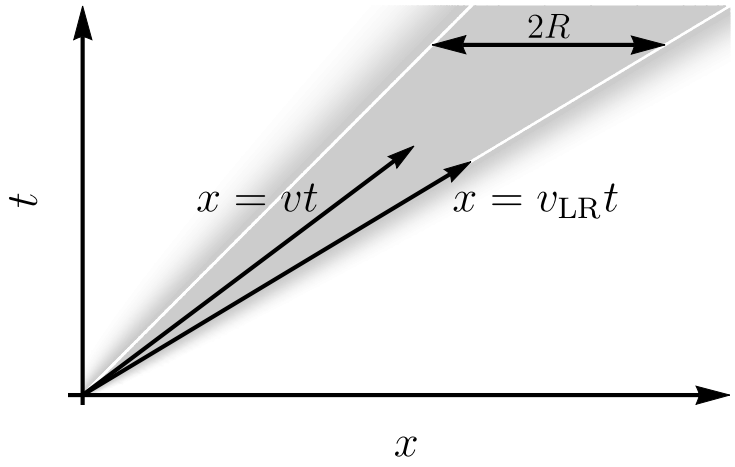


Figure 2.1: **Operator growth along a ray:** Schematic plot showing the definition of $R(\boldsymbol{v}t, t)$. The shaded region shows the radius of support of $O \equiv [O_1(0,t), O_2(\boldsymbol{x}, 0)]$ along the ray $\boldsymbol{x} = \boldsymbol{v}t$. R is the radius of the support up to an exponential tail. Because of the Lieb-Robinson bound for $O_1(0,t)$ and that $O_2(\boldsymbol{x}, 0)$ sits on the line $\boldsymbol{x} = \boldsymbol{v}t$, the support contains the ray $\boldsymbol{x} = \boldsymbol{v}t$ and is within the Lieb-Robinson cone.

Even in non-interacting theories, however, more general operators — such as a pair of entangled quasiparticles moving in opposite directions — can have a nonzero scrambling velocity according to the definition (2.8). Relatedly, simple operators in weakly interacting theories need not have a small scrambling velocity. In this work we will mostly be interested in strongly scrambling systems. The bound we obtain will not, in general, usefully constrain weakly scrambling dynamics.

2.3 Scrambling bounds the state dependence of the OTOC

In the following subsections we prove a bound on the temperature dependence of the velocity-dependent Lyapunov exponent (2.6), in one spatial dimension. We also make an argument that an analogous result holds in higher dimensions. Namely:

$$|\partial_\beta \lambda(\boldsymbol{v}; \rho)| \leq \frac{2h}{a}\Big(v_S(\boldsymbol{v}) - (\xi + \xi_{\mathrm{LR}})\lambda(\boldsymbol{v}; \rho)\Big)\,, \tag{2.9}$$

where β is the inverse temperature, a the lattice spacing, ξ the correlation length, ξ_{LR} the microscopic lengthscale in (2.5), essentially the interaction range, and $h \equiv 2\sup_{\boldsymbol{x}\in\Lambda} \|h_{\boldsymbol{x}}\|$ for the Hamiltonian in (2.3). The content of (2.9) is that the change with temperature of the Lyapunov exponent along a ray is bounded by the rate of growth of the commutator along the ray. Zooming in on the butterfly light cone $v \sim v_B$, this bound implies that the growth of the commutator at the butterfly light cone bounds the change of characteristics such as the butterfly velocity. As (for example) the temperature

is increased, these growing operators are 'activated' and contribute to scrambling.

A generalization, with full proof in the appendices, is as follows: For any Gibbs state $\rho = e^{-\sum_i \mu_i C^i}/\operatorname{tr} e^{-\sum_i \mu_i C^i}$ with mutually commuting conserved charges C^i, where $\mu_i \in \mathbb{R}$ and $C^i = \sum_{\boldsymbol{x}\in\Lambda} c^i_{\boldsymbol{x}}$ is a sum of local operators, then

$$\left|\frac{\partial\lambda(\boldsymbol{v};\rho)}{\partial\mu_i}\right| \leq \frac{2c^i}{a}\Big(v_S(\boldsymbol{v}) - (\xi+\xi_{\mathrm{LR}})\lambda(\boldsymbol{v};\rho)\Big). \tag{2.10}$$

The definition of $c^i > 0$ is similar to h above: $c^i \equiv 2\sup_{\boldsymbol{x}\in\Lambda}\|c^i_{\boldsymbol{x}}\|$.

2.3.1 Outline of proof in one dimension

The following gives an outline of the proof of (2.9). The logic is straightforward, but technical complications arise, for example, due to the fact that time evolution generates exponentially decaying tails in space for local operators, so one cannot assume that local operators have strictly finite support. These technical points are addressed in the appendices.

Let $\rho = e^{-\beta H}/\operatorname{tr} e^{-\beta H}$ be a thermal state with inverse temperature β and correlation length ξ. The steps will be as follows: *(i)* Differentiate the OTOC with respect to the inverse temperature, *(ii)* show that the main contribution to this derivative is from operators inside the support of the commutator, and *(iii)* balance the growth of this contribution, due to the growing size of the commutator along a ray, with the growth or decay of the OTOC. We now outline these steps.

(i) *Temperature derivative of the OTOC.* Taking the derivative of the OTOC (2.1) with respect to the inverse temperature gives

$$\partial_\beta \mathcal{C}(\boldsymbol{x},t;\rho) = -\operatorname{tr}(\rho\,\widetilde{H}O^\dagger O) = -\operatorname{tr}(\widetilde{H}\sqrt{\rho}O^\dagger O\sqrt{\rho})\,, \tag{2.11}$$

where $O \equiv i[O_1(0,t), O_2(\boldsymbol{x},0)]$ and $\widetilde{H} \equiv H - \operatorname{tr}(\rho H)$ is the Hamiltonian with thermal expectation value subtracted out.

The Hamiltonian H in (2.3) is written as a sum of local terms. We can split this sum up into terms that are inside and outside the support of the commutator O (for some location $\boldsymbol{x}$ and time t). As in the definition of v_S, let O be roughly supported in a ball of center $\boldsymbol{y}_0$ and radius R. Then

$$\widetilde{H} = \sum_{|\boldsymbol{y}-\boldsymbol{y}_0|\leq R+\delta} \widetilde{h}_{\boldsymbol{y}} + \sum_{|\boldsymbol{y}-\boldsymbol{y}_0|>R+\delta} \widetilde{h}_{\boldsymbol{y}}, \tag{2.12}$$

where $\delta > 0$ can take any value. As for $\widetilde{H}$, $\widetilde{h}_{\boldsymbol{y}} \equiv h_{\boldsymbol{y}} - \operatorname{tr}(\rho h_{\boldsymbol{y}})$. This decomposition can now be inserted into the derivative (2.11).

(ii) *Dominance by operators inside the commutator.* We first bound the contribution from outside of the support of the commutator, with $|\boldsymbol{y}-\boldsymbol{y}_0| > R+\delta$ in (2.12). Due to the thermal correlation

length ξ, the connected correlation function of $\widetilde{h}_{\boldsymbol{y}}$ with $O^\dagger O$ will decay exponentially in the distance $|\boldsymbol{y}-\boldsymbol{y}_0|$. Thus, for some constant $C>0$ and all $\boldsymbol{y}\in\Lambda$ such that $|\boldsymbol{y}-\boldsymbol{y}_0|>R$: $|\mathrm{tr}(\widetilde{h}_{\boldsymbol{y}}\sqrt{\rho}\,O^\dagger O\sqrt{\rho})| \le C\|\widetilde{h}_{\boldsymbol{y}}\|\|O\|^2 e^{(R-|\boldsymbol{y}-\boldsymbol{y}_0|)/\xi}$. Summing over $|\boldsymbol{y}-\boldsymbol{y}_0|>R+\delta$, the contribution to (2.11) from operators outside of the commutator is bounded by

$$\sum_{|\boldsymbol{y}-\boldsymbol{y}_0|>R+\delta}\left|\mathrm{tr}(\widetilde{h}_{\boldsymbol{y}}\sqrt{\rho}O^\dagger O\sqrt{\rho})\right| \le C'\sup_{\boldsymbol{y}\in\Lambda}\|\widetilde{h}_{\boldsymbol{y}}\|\|O\|^2 e^{-\delta/\xi}\,. \tag{2.13}$$

In d spatial dimensions and for $R+\delta\gg\xi$, $C'\sim C\xi(R+\delta)^{d-1}/a^d$ from doing the sum over $|\boldsymbol{y}-\boldsymbol{y}_0|>R+\delta$ (a is the lattice spacing). There is a technical subtlety in obtaining (2.13) due to the need to commute factors of $\sqrt{\rho}$ through $\widetilde{h}_{\boldsymbol{y}}$; we deal with this in the appendices.

We can similarly bound the contribution to (2.11) from operators inside the support of the commutator, with $|\boldsymbol{y}-\boldsymbol{y}_0|\le R+\delta$. As in the main text, define the maximal local coupling in the Hamiltonian as

$$h\equiv 2\sup_{\boldsymbol{y}\in\Lambda}\|h_{\boldsymbol{y}}\|\,. \tag{2.14}$$

Note that $\|\widetilde{h}_{\boldsymbol{y}}\|\le 2\|h_{\boldsymbol{y}}\|$, so that

$$|\mathrm{tr}(\widetilde{h}_{\boldsymbol{y}}\sqrt{\rho}\,O^\dagger O\sqrt{\rho})| \le \|\widetilde{h}_{\boldsymbol{y}}\|\,\mathrm{tr}(\rho\,O^\dagger O) \le h\,\mathcal{C}(\boldsymbol{x},t;\rho). \tag{2.15}$$

Notice that the inequality still goes through if we take

$$h=\sup_{\boldsymbol{y}\in\Lambda}\frac{|\mathrm{tr}(\widetilde{h}_{\boldsymbol{y}}\sqrt{\rho}\,O^\dagger O\sqrt{\rho})|}{\mathrm{tr}(\rho\,O^\dagger O)}\,. \tag{2.16}$$

Now, the number of terms in the first sum of (2.12) is $V_{R+\delta}$, the number of lattice points in a ball of radius $R+\delta$. Therefore, putting together (2.13) and (2.15), we can bound the derivative (2.11) by:

$$|\partial_\beta\mathcal{C}(\boldsymbol{x},t;\rho)| \le V_{R+\delta}\,h\,\mathcal{C}(\boldsymbol{x},t;\rho)+C'h\,\|O\|^2 e^{-\delta/\xi}. \tag{2.17}$$

We will see that in a certain kinematic limit, the final term in (2.17), from outside of the support of the commutator, is small compared to the other terms.

(iii) *Bounding the derivative by the growth of the commutator.* The inequality (2.17) simplifies at late times along a ray $\boldsymbol{x}=\boldsymbol{v}t$. From the definition (2.6) of the velocity-dependent Lyapunov exponent, $\mathcal{C}(\boldsymbol{v}t,t;\rho)\sim e^{\lambda(\boldsymbol{v};\rho)t}$ as $t\to\infty$. We furthermore set $\delta=(-\xi\lambda(\boldsymbol{v};\rho)+\epsilon)t>0$, with $\epsilon>0$ a small number. This choice is such that the final term in (2.17) decays exponentially faster than the others as $t\to\infty$. This final term is therefore negligible in this limit. In this

way, as $t \to \infty$ the following inequality is obtained:

$$|\partial_\beta \lambda(\boldsymbol{v};\rho)| \leq h \lim_{t\to\infty} \frac{V_{R-\xi\lambda(\boldsymbol{v};\rho)t}}{t}. \tag{2.18}$$

This expression bounds the temperature dependence of the Lyapunov exponent in terms of the late time growth of the commutator along a ray. The late time limit in (2.18) is manifestly finite in one spatial dimension, $d = 1$. In one dimension at large radii $V_r \approx 2r/a$, where a is the lattice spacing. In this case, the operator growth in (2.18) is precisely given by the scrambling velocity defined in (2.8). Thus, in terms of the scrambling velocity we obtain (A more rigorous treatment in the appendices, allowing for exponential tails in the support, shows that $\xi \to \xi + \xi_{\mathrm{LR}}$. We include this shift in the following statement of the bound.)

$$|\partial_\beta \lambda(\boldsymbol{v};\rho)| \leq \frac{2h}{a}\Big(v_S(\boldsymbol{v}) - (\xi + \xi_{\mathrm{LR}})\lambda(\boldsymbol{v};\rho)\Big). \tag{2.19}$$

2.3.2 Generalization to higher dimensions

In higher dimensions, V_r will scale as r^d for $d > 1$ and hence the late time bound (2.18) is always trivially true. However, we conjecture that the bound stated in (2.9) holds for arbitrary dimensions, based on a Lieb-Robinson type argument. One way of understanding the Lieb-Robinson bound is to expand

$$O_1(t) = \sum_{n=0}^{\infty} \frac{(it[H,\,\cdot\,])^n}{n!} O_1 = O_1 + it[H, O_1] - \frac{t^2}{2}[H,[H,O_1]] + \ldots, \tag{2.20}$$

and observe that in the expansion, for $[O_1(0,t), O_2(\boldsymbol{x},0)]$ to be nonzero, a commutator sequence of local terms in H connecting O_1 and O_2 is necessary, which starts at order $n \approx |\boldsymbol{x}|/R_H$ where R_H is the range of local terms in H. For such a high order term to be significant, t has to be later than $|\boldsymbol{x}|/(R_H h)$ and this gives an estimate of $v_{\mathrm{LR}} \approx R_H h$.

In a proof along these lines it is intuitively clear that outside the Lieb-Robinson cone $|\boldsymbol{x}| = v_{\mathrm{LR}} t$, the leading contributions to the commutator $[O_1(0,t), O_2(\boldsymbol{x},0)]$ come from O_1 taking commutators with local terms in H (as shown in (2.20)), via the shortest path from the origin to $\boldsymbol{x}$. Hence it is plausible that the operator $[O_1(0,t), O_2(\boldsymbol{x},0)]$, for $|\boldsymbol{x}| \gg v_{\mathrm{LR}} t$, is approximately one-dimensional, along the line connecting 0 and $\boldsymbol{x}$. Then the bound (2.9) is still expected to be true, although possibly with a larger 'renormalized' h.

2.4 Temperature dependence of the butterfly velocity

2.4.1 Numerical results on the mixed field Ising chain

To motivate the general discussion of butterfly velocities, it will be useful to have some explicit numerical results for the temperature dependence of the butterfly velocity at hand. To this end we have studied the mixed field Ising chain with Hamiltonian

$$H = -J \sum_{i=1}^{N-1} Z_i Z_{i+1} + h_X \sum_{i=1}^{N} X_i + h_Z \sum_{i=1}^{N} Z_i, \tag{2.21}$$

where X_i, Y_i and Z_i are Pauli matrices at site i. Numerics is done with a straightforward generalization of the Matrix Product Operator (MPO) method discussed in [142, 256] to finite temperatures. Some analytic results on OTOCs in the transverse field model ($h_Z = 0$) can be found in [148]. In numerics we will have $N = 25$. More details can be found in the appendices. Results for the temperature dependence of the butterfly velocity for Pauli Z operators are shown in Fig. 2.2.

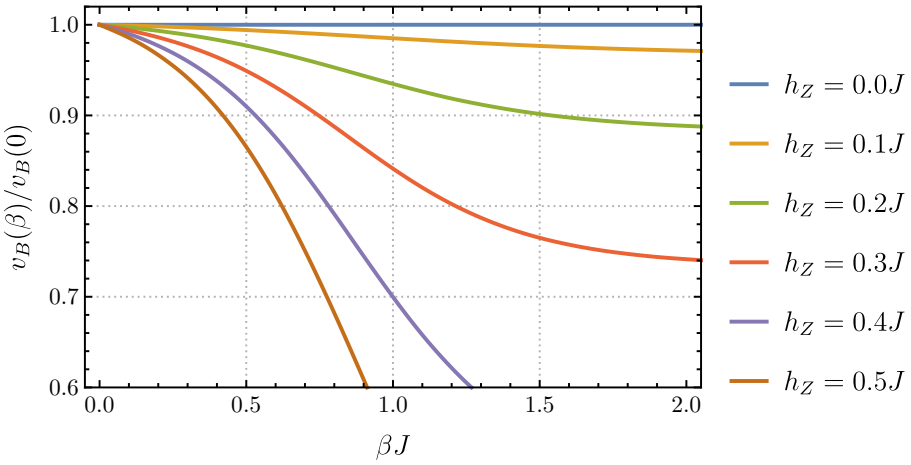


Figure 2.2: **Temperature-dependent butterfly velocity in the mixed field Ising chain** (2.21) with $h_X = 1.05J$ and different h_Z. The inverse temperature is denoted as β. The model with $h_Z = 0$ is dual to free fermions and has a temperature-independent butterfly velocity. The appendices contain more details about numerics and error estimates.

The numerical results in Fig. 2.2 exhibit the behavior advertised in the introduction, and which we will understand in detail below. The transverse field Ising model ($h_Z = 0$) is dual to free fermions via a Jordan-Wigner transformation. The longitudinal field h_Z introduces interactions. We expect interactions to induce scrambling dynamics and hence a nontrivial temperature dependence of the butterfly velocity, and this is what the figure shows.

The temperature-independent butterfly velocity of the transverse field model deserves some

elaboration. There are two points to make. Firstly, the transverse field model is special in its duality to a *non-interacting* integrable system, where $v_S = 0$ for the commutator of fermion creation and annihilation operators, for example. For interacting integrable systems, typically $v_S > 0$ and the butterfly velocity is state-dependent [82]. Indeed, we have verified numerically that the butterfly velocity is temperature-dependent in such models. Interacting integrable systems are scrambling, even while they are not chaotic.

Secondly, in the transverse field model, Pauli Z's in the spin frame are dual to nonlocal fermion chains by the Jordan-Wigner transformation. Due to this nonlocality, our inequality doesn't apply in the fermion frame. In fact, even local operators describing small numbers of quasiparticles in a non-interacting theory can have $v_S > 0$ by our definition because entangled pairs of quasiparticles moving in opposite directions technically lead to a linearly growing radius of support for the operator. We believe that it may be possible to overcome this technical complication in the future with an improved definition of the scrambling velocity, such that $v_S = 0$ for spatially separated but entangled non-scrambling operators. Indeed, we shall now argue that the butterfly velocity is temperature independent for all local operators in a non-interacting system.

In a non-interacting theory the propagation of quasiparticles is independent of the state they are propagating in, due to the absence of interactions between them. While the quasiparticles may have a nontrivial dispersion and hence temperature-dependent average velocity, any local operator includes modes of all wavevectors and, in particular, maximal velocity modes. Thus we expect v_B is independent of the state. Therefore, the temperature-independence of the butterfly velocity observed in our numerics is indeed symptomatic of the non-interacting integrability of the system.

2.4.2 Bounding the butterfly velocity

The temperature dependence shown in Fig. 2.2 can be understood from the connections between the OTOC and scrambling velocity that we have described. The 'light front' form (2.2) for the OTOC implies that the velocity-dependent Lyapunov exponent is

$$\lambda(\boldsymbol{v};\rho) = -\lambda(v/v_B - 1)^{1+p} \qquad \text{for} \qquad v \geq v_B\,. \tag{2.22}$$

This precise form for $\lambda(\boldsymbol{v};\rho)$ is conveniently explicit, but the only qualitatively essential aspect for our results is the presence of a 'butterfly cone'. As we explained above, in general λ, v_B and $p \geq 0$ are state-dependent. Therefore, the ∂_{μ_i} derivative in (2.10) will act on each of these quantities. Substituting the specific form (2.22) for $\lambda(\boldsymbol{v};\rho)$ into (2.10), for $v > v_B$, leads to the following slightly complicated expression:

$$\begin{aligned} a\lambda(\Delta v)^{1+p}\Big|\partial_{\mu_i}\ln\lambda + \ln(\Delta v)\partial_{\mu_i}p - (1+p)\frac{v/v_B}{\Delta v}\,\partial_{\mu_i}\ln v_B\Big| \\ \leq 2c^i\left[v_S(v) + (\xi+\xi_{\mathrm{LR}})\lambda(\Delta v)^{1+p}\right], \end{aligned} \tag{2.23}$$

where $\Delta v \equiv v/v_B - 1 > 0$ is a dimensionless measure of how far the velocity is outside the butterfly cone. A simple consequence of (2.23) follows, when there is no scrambling. Suppose that $v_S(v) = 0$. In that case, taking $\Delta v \to 0^+$, the leading term on the left side of (2.23) is the last one. It follows that

$$v_S = 0 \;\; \Rightarrow \;\; \partial_{\mu_i} v_B = 0\,. \tag{2.24}$$

Hence v_B is constant for operators that do not scramble. We noted above, however, that this result is not directly applicable to the transverse field Ising chain.

Increasing variation of v_B with temperature is observed in Fig. 2.2 as integrability is increasingly broken by turning on h_Z in the mixed field Ising model. The crossover temperature in Fig. 2.2 is set by the energy gap Δ (of order J for $h_Z = 0.1 \sim 0.5J$), as we now explain. Intuitively, one might expect v_B to cease varying at temperatures $T \ll \Delta$. This is what is seen in the numerical data. We can argue for this by improving an aspect of the proof outlined previously. As we note there, the proof still goes through if we take h in (2.9) to be instead given by

$$h = \sup_{t>0,\, \boldsymbol{y}\in\Lambda} \frac{|\mathrm{tr}(\widetilde{h}_{\boldsymbol{y}}\sqrt{\rho}\, O^\dagger O\sqrt{\rho})|}{\mathrm{tr}(\rho\, O^\dagger O)}\,, \tag{2.25}$$

where $O \equiv i[O_1(0,t), O_2(\boldsymbol{v}t, 0)]$ and $\widetilde{h}_{\boldsymbol{y}} \equiv h_{\boldsymbol{y}} - \mathrm{tr}(\rho h_{\boldsymbol{y}})$. This is not an especially tractable expression in general, but it can be evaluated for a gapped system at zero temperature, where $\rho \equiv |0\rangle\langle 0|$. In that case $h = \sup_{\boldsymbol{y}\in\Lambda} \langle 0|\widetilde{h}_{\boldsymbol{y}}|0\rangle = 0$, where now $\widetilde{h}_{\boldsymbol{y}} \equiv h_{\boldsymbol{y}} - \langle 0|h_{\boldsymbol{y}}|0\rangle$. Hence in gapped systems at low temperatures, we may set $h \approx 0$ in the bound (2.9). It follows that $\partial_\beta v_B \to 0$ when $T \to 0$ in a gapped system, consistent with the finite low temperature butterfly velocities seen in Fig. 2.2.

The numerical results in Fig. 2.3 substantiate the above argument, suggesting that $\partial_\beta v_B$ decays exponentially as $\beta\Delta \to \infty$. In Fig. 2.3 the bound has furthermore been written as a bound on the derivative of the butterfly velocity, and is found to be most constraining at intermediate temperatures and with strong scrambling, where it is within an order of magnitude of the true value.

Our bound combined together with numerics leads to a consistent picture of the temperature dependence of the butterfly velocity in chaotic spin systems with a gap Δ. Stronger scrambling allows for stronger temperature dependence of v_B, which furthermore approaches a constant at $T \ll \Delta$. These facts explain the crossover features of the curves in Fig. 2.2. More quantitatively, the overall variation $v_B(\beta = 0)/v_B(\beta = \infty)$ can be bounded by integrating our bound from $\beta = 0$ to $\beta\Delta \sim 1$ (assuming that there are no intervening thermal phase transitions). For small $v_S(v)$, this integration can be done explicitly, leading to a bound on the change in the butterfly velocity from infinite to zero temperature. For notational convenience let $v_S^B \equiv v_S(v_B)$. At small v_S^B one may take $\Delta v \sim (v_S^B/v_B)^{1/(1+p)}$ in (2.23) and the leading term on the left hand side is again the final one,

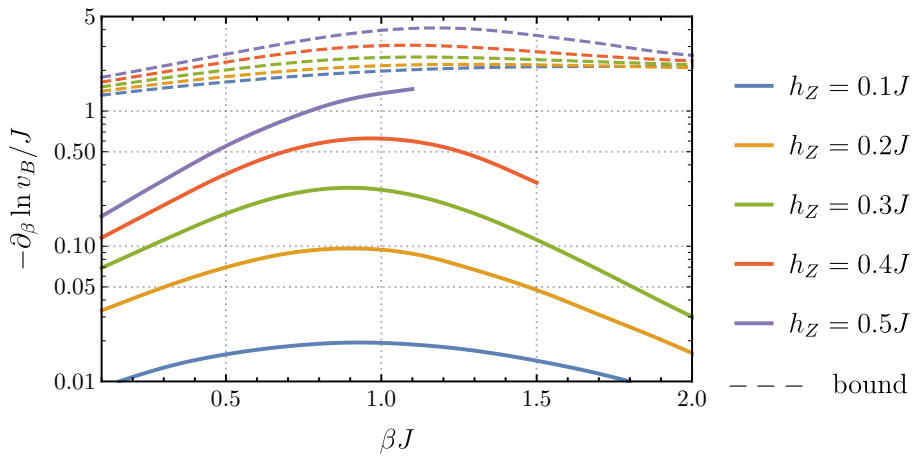


Figure 2.3: **Bounding the temperature derivative of the butterfly velocity:** Temperature derivative of the butterfly velocity in mixed field Ising chains, with $h_X = 1.05J$ and different h_Z in (2.21). The inverse temperature is denoted as β. The bound (2.23) is shown as the dashed curves. In the bound v_S is replaced by $3Ja$ ($a = 1$ is the lattice spacing), using the fact that $v_S \leq v$ for $v = 3Ja$ and $\xi_{\rm LR} = a$ in the Lieb-Robinson inequality (2.5), in the spin duality frame. Curves are cut off when estimated error is significant (see the appendices for more details).

which integrates to

$$\left| \ln \frac{v_B(\beta = \infty)}{v_B(\beta = 0)} \right| \lesssim \int_0^{1/\Delta} d\beta \, \frac{2h v_B^{p/(1+p)} [1 + (\xi + \xi_{\rm LR})\lambda / v_B]}{a\lambda(1+p)} \left(v_S^B\right)^{1/(1+p)} , \tag{2.26}$$

to leading order in $v_S^B \to 0$. Typically $v_B(\beta = 0) \sim v_{\rm LR}$. Schematically we can therefore write

$$v_B(T = 0) \gtrsim v_{\rm LR} \, e^{-\alpha v_S^\gamma / \Delta} . \tag{2.27}$$

Here α is a dimensionful constant, γ a dimensionless constant and we have singled out the v_S and Δ dependences. It follows that *(i)* as $v_S \to 0$, $\ln v_B$ can vary as a power v_S^γ of the scrambling velocity, and *(ii)* if the gap $\Delta \to 0$, v_B may approach zero at $T = 0$. Indeed, power law butterfly velocities $v_B \sim T^{1-1/z}$, with z the dynamical critical exponent, are found in strongly chaotic gapless holographic models [219, 28].

2.5 Final comments

In summary, we have shown how locality of quantum dynamics ties operator growth to the butterfly velocity. This connection arises because the growth of the spatial support of the commutator right outside the butterfly cone bounds the change of the butterfly velocity with e.g. temperature. The butterfly velocity is state-dependent and therefore gives a richer characterization of the finite

temperature dynamics than is possible from the microscopic Lieb-Robinson velocity alone. We have demonstrated these ideas explicitly in numerical studies of quantum chaotic lattice models at finite temperature. Looking forward, we hope that the methods we have developed can be used to bound other important quantities that underpin quantum many-body systems, in particular the thermalization length and time, as well as transport observables such as the thermal diffusivity.

Acknowledgements

It is a pleasure to acknowledge Jacob Marks for helping with numerics and Daren Chen for reading the proofs. We are grateful to Vedika Khemani and Xiao-Liang Qi for insightful comments on an earlier version. SAH is partially funded by DOE award de-sc0018134. XH is supported by a Stanford Graduate Fellowship. Computational work was performed on the Sherlock cluster at Stanford University, with the ITensor library for implementing tensor network calculations.

Appendices

This appendix contains six sections: section 2.6 sets up notations and backgrounds for discussions that follow. In section 2.7 we review the Lieb-Robinson, Araki and correlation length bounds used in our proof. Precise definitions for Lieb-Robinson, butterfly and scrambling velocities are given in section 2.8 and we prove several inequalities regarding them. Section 2.9 collects technical lemmas for exponentially local operators and section 2.10 gives a rigorous proof of the general results. Details of numerical implementations and data analysis are presented in section 2.11.

2.6 Appendix A: Notation

In this section we introduce notations and concepts necessary for a rigorous proof of our result. The bound will be formulated for a lattice[2] Λ of spins in d spatial dimensions, and rigorously proved for $d = 1$. There are isomorphic finite-dimensional Hilbert spaces $\mathbb{H}_{\boldsymbol{x}}$ associated to each lattice site $\boldsymbol{x} \in \Lambda$ and denote $\mathcal{B}_{\boldsymbol{x}}$ as the space of linear operators acting on $\mathbb{H}_{\boldsymbol{x}}$. An operator O is said to be supported on a subset $S \subset \Lambda$ if $O \in \bigotimes_{\boldsymbol{x}\notin S} \mathbb{C}I \otimes \bigotimes_{\boldsymbol{x}\in S} \mathcal{B}_{\boldsymbol{x}}$, i.e. O is a sum of product operators that are identity outside S. The minimal set that O is supported on is called the support of O, denoted as $\operatorname{supp} O$.[3]

To better characterize the spatial distribution of operators, define superoperators $\mathcal{P}_S$ and $\mathcal{Q}_S \equiv \mathrm{Id} - \mathcal{P}_S$ such that $\mathcal{P}_S$ is the projection onto the subspace $\bigotimes_{\boldsymbol{x}\notin S} \mathbb{C}I \otimes \bigotimes_{\boldsymbol{x}\in S} \mathcal{B}_{\boldsymbol{x}}$. That is, $\mathcal{P}_S$ projects onto operators supported on S (so $\mathcal{P}_S[O] = O$ if O is supported on S). More explicitly

$$\mathcal{P}_S[O] \equiv \int_{\operatorname{supp} U \cap S = \emptyset} dU\, UOU^\dagger, \tag{2.28}$$

where the integral is Haar averaging over unitaries outside S. However, note $\mathcal{Q}_S$ is *not* the projection onto operators supported on $\Lambda - S$. Consider an example of two sites $\Lambda = \{1, 2\}$ and an operator $O = O_1 \otimes O_2$, where neither O_1 nor O_2 is the identity. By definition, $0 = \mathcal{P}_1[O] = \mathcal{P}_2[O] \neq \mathcal{Q}_1[O] = \mathcal{Q}_2[O] = O$.

Henceforth if the subscript $S = \{\boldsymbol{x}\}$ is a single-element set, $\mathcal{P}_{\{\boldsymbol{x}\}}$ and $\mathcal{Q}_{\{\boldsymbol{x}\}}$ are written as $\mathcal{P}_{\boldsymbol{x}}$ and $\mathcal{Q}_{\boldsymbol{x}}$ for short. Also define the superoperator $\mathcal{P}_T^r$ with a superscript $r > 0$ as $\mathcal{P}_S$ for $S = \{\boldsymbol{y} \in \Lambda : \exists\, \boldsymbol{x} \in T, |\boldsymbol{x} - \boldsymbol{y}| < r\}$, i.e. projection onto operators supported within a distance r from the set T, and $\mathcal{Q}_T^r \equiv \mathrm{Id} - \mathcal{P}_T^r$.

From (2.28) we have the following inequalities:

$$\|\mathcal{P}_S[O]\| \leq \|O\|, \quad \|\mathcal{Q}_S[O]\| = \|O - \mathcal{P}_S[O]\| \leq \|O\| + \|\mathcal{P}_S[O]\| \leq 2\|O\|, \tag{2.29}$$

[2] Technically the infinite lattice should be thought as the limit of a sequence of increasing finite subsystems. We will not delve into subtleties related to this point.

[3] Note $\operatorname{supp} O = \emptyset$ if and only if $O = cI$ for some $c \in \mathbb{C}$.

as $\|U\| = \|U^\dagger\| = 1$. Also $\mathcal{P}_S[I] = I$, $\mathcal{Q}_S[I] = 0$ for any $S \subset \Lambda$. Unless otherwise specified, $\|O\|$ will always denote the operator norm, i.e. the maximal singular value of O.

We will be interested primarily in operators that are "exponentially local", denoted as $\mathcal{B}(\boldsymbol{x}, R; \xi, C)$. We say $O \in \mathcal{B}(\boldsymbol{x}, R; \xi, C)$ with $\boldsymbol{x} \in \Lambda$, $R, C \geq 0$ and $\xi > 0$, if for any $r \geq R$,

$$\|\mathcal{Q}^r_{\boldsymbol{x}}[O]\| \leq C\|O\|e^{-(r-R)/\xi}. \tag{2.30}$$

Intuitively, this means O is supported on the ball of radius R and centered at $\boldsymbol{x}$, up to an exponential tail of lengthscale ξ. Operators supported on a finite number of sites (called "finitely supported") are of course exponentially local as well. We shall assume the Hamiltonian is a sum of finitely supported hermitian terms:

$$H = \sum_\alpha J_\alpha H^\alpha, \quad H^\alpha \equiv \sum_{\boldsymbol{x}\in\Lambda} h^\alpha_{\boldsymbol{x}}, \quad h^\alpha_{\boldsymbol{x}} \in \mathcal{B}(\boldsymbol{x}, R_H; 0^+, 0), \tag{2.31}$$

which also defines $R_H > 0$ and α labels different couplings in the Hamiltonian. Translational invariance is not necessary but $\|h^\alpha\| \equiv \sup_{\boldsymbol{x}\in\Lambda} \|h^\alpha_{\boldsymbol{x}}\|$ should be bounded.

A Gibbs state is a density matrix of the form

$$\rho = e^{-\sum_i \mu_i C^i} / \operatorname{tr} e^{-\sum_i \mu_i C^i}, \tag{2.32}$$

for some $\mu_i \in \mathbb{R}$ and

$$C^i \equiv \sum_{\boldsymbol{x}\in\Lambda} c^i_{\boldsymbol{x}}, \quad c^i_{\boldsymbol{x}} \in \mathcal{B}(\boldsymbol{x}, R_H; 0^+, 0). \tag{2.33}$$

In the proof it is *not* required that $[C^i, C^j] = 0$. With only one i, with μ the inverse temperature and with $C = H$, ρ is the thermal density matrix.

2.7 Appendix B: Review of locality bounds

In this section we review some established locality bounds. First is the Lieb-Robinson bound in local lattice systems [146, 37, 179, 177]. This both bounds the spread of support of a local operator by the distance $v|t|$, where t is the real time of Heisenberg evolution, and also implies an emergent causality with v acting as the "speed of light". For a discussion of the relation between (i) and (ii) in the following theorem, see section 3 of [180].

Theorem 1 (Lieb-Robinson)**.** *There exist $v, \xi_{\mathrm{LR}}, C_{\mathrm{LR}} > 0$, dependent on lattice geometry and Hamiltonian, such that*

(i) for any $t \in \mathbb{R}$, $r > 0$ and operator O,

$$\|\mathcal{Q}^r_{\operatorname{supp} O}[O(t)]\| \leq C_{\mathrm{LR}}|\partial \operatorname{supp} O|\|O\| \min\{1, e^{(v|t|-r)/\xi_{\mathrm{LR}}}\}, \tag{2.34}$$

where $|\partial S|$ is the number of lattice links (say, between x and y) such that $x \in S$ but $y \notin S$;

(ii) for any $t \in \mathbb{R}$, operators O_1 and O_2,

$$\|[O_1(t), O_2]\| \leq C_{\mathrm{LR}} \min\{|\partial \operatorname{supp} O_1|, |\partial \operatorname{supp} O_2|\} \|O_1\| \|O_2\| \min\{1, e^{(v|t|-d)/\xi_{\mathrm{LR}}}\}, \tag{2.35}$$

where $d = \min\{|\boldsymbol{x} - \boldsymbol{y}| : \boldsymbol{x} \in \operatorname{supp} O_1, \boldsymbol{y} \in \operatorname{supp} O_2\}$ is the distance between the support of O_1 and O_2.

In this bound $v \sim \sum_\alpha |J_\alpha| \|h^\alpha\| R_H$, recall (2.31), i.e. coupling times range of local terms in the Hamiltonian, and $\xi_{\mathrm{LR}} \sim R_H$. So quantities in the Lieb-Robinson bound are set by microscopic scales, to be differentiated from the butterfly velocity, which is an analog of a "renormalized" Lieb-Robinson velocity in thermal states [219].

Next is the Araki bound [180, 11, 34] extending the Lieb-Robinson bound to *complex* times. Note the theorem is specific to one dimension [34] and $l_{\mathrm{A}}(\mu_i)$ may be exponential in $|\mu_i|$; in this sense the restriction is weaker for complex time evolution:

Theorem 2 (Araki)**.** *In one dimension, for any Gibbs state ρ as defined in (2.32) but with $\mu_i \in \mathbb{C}$, there exist $l_{\mathrm{A}}(\mu_i), C_{\mathrm{A}}(\mu_i), \xi_{\mathrm{A}} > 0$, dependent on lattice geometry and charges C^i, such that for any finitely supported operator O and $r \geq l_{\mathrm{A}}(\mu_i)$,*

$$\|\rho O \rho^{-1}\| \leq C_{\mathrm{A}}(\mu_i) |\operatorname{supp} O| \|O\|, \tag{2.36}$$

$$\|\mathcal{Q}^r_{\operatorname{supp} O}[\rho O \rho^{-1}]\| \leq C_{\mathrm{A}}(\mu_i) |\operatorname{supp} O| \|O\| e^{(l_{\mathrm{A}}(\mu_i) - r)/\xi_{\mathrm{A}}}, \tag{2.37}$$

where $|\operatorname{supp} O|$ is the number of sites in $\operatorname{supp} O$.

Note, however, from the proof of the Araki bound (e.g., Theorem 3.1 of [34]) one can see that there are Araki inequalities as stated in Theorem 2 for *arbitrarily* small ξ_{A}, at the expense of a possibly large l_{A}. Later in the proof of our bound only ξ_{A} enters the final expression; hence at that time one can take $\xi_{\mathrm{A}} \to 0$ as a large l_{A} doesn't affect the result.

Originally the Araki bound is only stated for finitely supported operators but it is straightforward to generalize it to exponentially local ones. Such generalization will be useful in proving our bound, so a proof is given in section 2.9.

Finally we would like to introduce some exponential clustering theorems: for particular kinds of states, equal-time connected correlations decay exponentially in space. More precisely for a state (density matrix) ρ, the correlation length of ρ is the $\xi > 0$ that is optimal with respect to the following property: there exists $C > 0$ and a function $l_0(\cdot) > 0$ such that for any operators O_1 and O_2 (supported on sets S and T) sufficiently far apart, i.e., $d \geq l_0(\delta)$,

$$|\mathrm{tr}(\rho O_1 O_2) - \mathrm{tr}(\rho O_1)\mathrm{tr}(\rho O_2)| \leq C\delta \|O_1\| \|O_2\| e^{-d/\xi}, \tag{2.38}$$

where $\delta \equiv \min\{|\partial S|, |\partial T|\}$ is the number of lattice links crossing the boundary of S or T, and $d \equiv \min\{|\boldsymbol{x}-\boldsymbol{y}| : \boldsymbol{x} \in S, \boldsymbol{y} \in T\}$ is the distance between two sets. Note that in this statement, O_1 and O_2 could be *any*, not necessarily local, operators.

Existence of a finite $\xi > 0$ with the property stated around (2.38) has been proved for (i) one-dimensional Gibbs states [11] (restricted to local operators O_1 and O_2), (ii) $\rho = |0\rangle\langle 0|$ where $|0\rangle$ is the unique ground state of a gapped Hamiltonian [179, 101], and (iii) thermal states $\rho \propto \exp(-\beta H)$ in general dimensions at sufficiently high temperatures [132] (clearly $\xi \to 0$ when $\beta \to 0$). Of course the Hamiltonians associated with these states must be local, as in (2.31) above. It is plausible that the correlation length ξ as defined around (2.38) is finite for Gibbs states ρ in general systems with local dynamics and away from phase transitions.

2.8 Appendix C: Definitions of velocities

In this section we define precisely the (possibly anisotropic) Lieb-Robinson, butterfly and scrambling velocities introduced in the main text and prove the bound $v_B, v_S \le v_{\mathrm{LR}}$. For definiteness fix a class of local operators, denoted as $\mathcal{O}$; for example, $\mathcal{O}$ could be all single-site operators with unit norm, localized at origin. The Lieb-Robinson bound Theorem 1 (ii) can be stated for such operators along any particular direction $\boldsymbol{n}$:

Theorem 3 (Operator-dependent anisotropic Lieb-Robinson)**.** *For any direction $\boldsymbol{n}$ and operator $O_1, O_2 \in \mathcal{O}$, there exist v, ξ_{LR}, $C_{\mathrm{LR}} > 0$, dependent on $\boldsymbol{n}$, O_1, O_2, lattice geometry and Hamiltonian, such that for any $t > 0$, $x > 0$,*

$$\|[O_1(0,t), O_2(x\boldsymbol{n},0)]\| \le C_{\mathrm{LR}} \|O_1\| \|O_2\| \min\{1, e^{(vt-x)/\xi_{\mathrm{LR}}}\}. \tag{2.39}$$

From Theorem 3 one immediate candidate for defining the Lieb-Robinson velocity is

$$v_{\mathrm{LR}}^{(1)}(\boldsymbol{n}; O_1, O_2) \equiv \inf\{v > 0 : \exists\, \xi_{\mathrm{LR}}, C_{\mathrm{LR}} > 0 \text{ with the property stated in Theorem 3}\}, \tag{2.40}$$

that is, the smallest velocity with a Lieb-Robinson inequality. However such a definition shows some disadvantages in numerical or experimental applications: it is inaccurate to fit data to exponential tails because the theorem only states an inequality (not an equality), and in fact in many lattice systems of interest the tail is observed to be sub-exponential (e.g., Gaussian) [128, 256]; also it is impractical, if not impossible, to decide whether such ξ_{LR} and C_{LR} exist for all times, from only a finite number of data points.

A more convenient definition is found in the original Lieb-Robinson paper [146]

$$v_{\mathrm{LR}}^{(2)}(\boldsymbol{n}; O_1, O_2) \equiv \sup\left\{ v : \lim_{t\to\infty} \frac{1}{t} \ln \|[O_1(0,t), O_2(vt\boldsymbol{n},0)]\| \ge 0 \right\}. \tag{2.41}$$

We will assume that the limit exists and is a continuous function of v. By definition $v_{\mathrm{LR}}^{(2)}$ gives a causality "lightcone" outside which (for $x/t > v$) the commutator vanishes exponentially at late times.

It is relatively easy to see that $v_{\mathrm{LR}}^{(1)} \geq v_{\mathrm{LR}}^{(2)}$:

Proposition 1. *For any direction $\boldsymbol{n}$ and operators $O_1, O_2 \in \mathcal{O}$, we have*

$$v_{\mathrm{LR}}^{(1)}(\boldsymbol{n}; O_1, O_2) \geq v_{\mathrm{LR}}^{(2)}(\boldsymbol{n}; O_1, O_2). \tag{2.42}$$

Proof. Let $v > 0$ belong to the set in (2.40), i.e., there exist $\xi, C > 0$ such that for all $x, t > 0$, $\|[O_1(0,t), O_2(x\boldsymbol{n}, 0)]\| \leq C\|O_1\|\|O_2\| \min\{1, e^{(vt-x)/\xi}\}$. Then, for any $v' > v$,

$$\lim_{t\to\infty} t^{-1} \ln \|[O_1(0,t), O_2(v't\boldsymbol{n}, 0)]\| \leq \lim_{t\to\infty} t^{-1} \ln(C\|O_1\|\|O_2\| e^{(v-v')t/\xi}) = (v-v')/\xi < 0, \tag{2.43}$$

and hence any $v' > v$ is not contained in the set in (2.41). Therefore the supremum $v_{\mathrm{LR}}^{(2)}$ is at most v. This is true for any $v > 0$ in the set of (2.40), hence $v_{\mathrm{LR}}^{(2)} \leq v_{\mathrm{LR}}^{(1)}$. □

Conversely to show that $v_{\mathrm{LR}}^{(1)} \leq v_{\mathrm{LR}}^{(2)}$, we need the following lemma:

Lemma 1. *For any positive functions $f(x,t)$ and $g(x,t)$, if limits*

$$\lim_{t\to\infty} \frac{1}{t} \ln f(vt, t) = \lambda_f(v), \quad \lim_{t\to\infty} \frac{1}{t} \ln g(vt, t) = \lambda_g(v), \tag{2.44}$$

exist, are uniform for $v \in [v_0, \infty)$, and $\lambda_f(v) + a < \lambda_g(v)$ for some $a > 0$ and all $v \geq v_0$, then there is $t_0 > 0$ that

$$f(x,t) < g(x,t) \quad \forall\, x \geq v_0 t,\ t \geq t_0. \tag{2.45}$$

Proof. Because the limits (2.44) are uniform, for any $\varepsilon > 0$ there is $T(\varepsilon) > 0$ such that for any $t \geq T(\varepsilon)$ and $v \geq v_0$, $\ln f(vt,t)/t < \lambda_f(v) + \varepsilon$, $\ln g(vt,t)/t > \lambda_g(v) - \varepsilon$. Now choose $\varepsilon = a/2$ and $t_0 = T(a/2)$, we have $\ln f(vt,t)/t < \lambda_f(v) + a/2 < \lambda_g(v) - a/2 < \ln g(vt,t)/t$ hence $f(vt,t) < g(vt,t)$, for all $t \geq t_0$, $v \geq v_0$. □

Proposition 2. *$v_{\mathrm{LR}}^{(1)}(\boldsymbol{n}; O_1, O_2) \leq v_{\mathrm{LR}}^{(2)}(\boldsymbol{n}; O_1, O_2)$, given the limit in (2.41) is uniform for all $v > v_{\mathrm{LR}}^{(2)}(\boldsymbol{n}; O_1, O_2)$.*

Proof. We would like to prove the proposition in the following two steps:

Step one: For any $v > v_{\mathrm{LR}}^{(2)}$, we show that (i) implies (ii), and (ii) implies (iii), where

(i) $\lim_{t\to\infty} t^{-1} \ln \|[O_1(0,t), O_2(v't\boldsymbol{n}, 0)]\| < 0$ for any $v' \geq v$;

(ii) $\exists\, \varepsilon, \xi > 0$ that $\lim_{t\to\infty} t^{-1} \ln \|[O_1(0,t), O_2(v't\boldsymbol{n}, 0)]\| \leq (v-v')/\xi - \varepsilon$ for any $v' \geq v$;

(iii) $\exists\, C, \xi > 0$ that $\|[O_1(0,t), O_2(x\boldsymbol{n}, 0)]\| \leq C\|O_1\|\|O_2\| \min\{1, e^{(vt-x)/\xi}\}$ for $x, t > 0$.

Step two: By definition (2.41) we have for any $v > v_{\mathrm{LR}}^{(2)}$, (i) holds for v; so (iii) is true for v as well, and v should be in the set on the right-hand side of (2.40) hence $v_{\mathrm{LR}}^{(1)} \leq v$. This shows that $v_{\mathrm{LR}}^{(1)} \leq v_{\mathrm{LR}}^{(2)}$.

So now it remains to prove that (i) $\Rightarrow$ (ii) and (ii) $\Rightarrow$ (iii):

(i) $\Rightarrow$ (ii): For clarity let's denote $\lambda(v) \equiv \lim_{t\to\infty} t^{-1} \ln \|[O_1(0,t), O_2(vt\boldsymbol{n}, 0)]\|$, then (i) says that $\lambda(v') < 0$ for any $v' \geq v$ and to arrive at (ii) we hope to find $\varepsilon, \xi > 0$ such that $\lambda(v') \leq (v - v')/\xi - \varepsilon$ for all $v' \geq v$.

Before construction of ε and ξ, it is remarkable that there is a restriction on $\lambda(v')$ from Theorem 3: the Lieb-Robinson bound states that there are some C_0, v_0, $\xi_0 > 0$ such that

$$\lambda(v') \leq \lim_{t\to\infty} t^{-1} \ln(C_0 \|O_1\| \|O_2\| e^{(v_0 - v')t/\xi_0}) = (v_0 - v')/\xi_0 \tag{2.46}$$

for all $v' > 0$.

We shall construct $\varepsilon > 0$ first. Note $(v - v')/\xi \leq 0$ for $v' \geq v$, hence it is required that $\lambda(v') \leq -\varepsilon$ for all $v' \geq v$. So we may choose $\varepsilon = \inf_{v'\geq v}(-\lambda(v')/2) \geq 0$. To show that $\varepsilon > 0$, we have to check that $-\lambda(v') > 0$ is bounded from zero on $[v, \infty)$. The only concern is $\lambda(v')$ may be arbitrarily close to zero when $v' \to \infty$; but this is not possible because from the previous paragraph $-\lambda(v') \geq (v' - v_0)/\xi_0 \to \infty$ as $v' \to \infty$. Hence $\varepsilon > 0$ is well-defined in this way.

Then to satisfy $\lambda(v') \leq (v - v')/\xi - \varepsilon$ for all $v' \geq v$, choose (ξ_0 is there for future convenience) $\xi \equiv \max\{\xi_0, \sup_{v'\geq v}(v - v')/(\lambda(v') + \varepsilon)\}$ (as constructed in the last paragraph the denominator is always negative). The task is then to show that $\xi < \infty$; similarly the only place things could go wrong is when $v' \to \infty$, but in that limit $|\lambda(v') + \varepsilon| \geq |\lambda(v')|/2 \geq (v' - v_0)/2\xi_0$ hence $\lim_{v'\to\infty}(v - v')/(\lambda(v') + \varepsilon) \leq 2\xi_0$ is bounded. So $\xi > 0$ is well-defined as well and (ii) is proved.

(ii) $\Rightarrow$ (iii): We would like to apply the Lemma 1 for $f(x,t) = \|[O_1(0,t), O_2(x\boldsymbol{n}, 0)]\|$ and $g(x,t) = \|O_1\|\|O_2\|e^{(vt-x)/\xi}$. Note in this case $\lambda_f(v') = \lambda(v') \leq (v - v')/\xi - \varepsilon = \lambda_g(v') - \varepsilon$ for any $v' \geq v$. Then by the lemma there is $t_0 > 0$ such that $\|[O_1(0,t), O_2(x\boldsymbol{n}, 0)]\| \leq \|O_1\|\|O_2\|e^{(vt-x)/\xi}$ for all $x \geq vt$ and $t \geq t_0$. Hence for (iii) to hold it suffices to choose that $C \equiv \max\{2, \sup_{0<x<vt \text{ or } 0<t<t_0} f(x,t)/g(x,t)\}$. As before we have to check that the supremum is not infinite. We will discuss the three cases (a) $0 < x < vt$, (b) $0 < t < t_0$ with $x \geq v_0 t$, and (c) $0 < t < t_0$ with $0 < x < v_0 t$ separately.

For $0 < x < vt$, $f(x,t)/g(x,t) = \|[O_1(0,t), O_2(x\boldsymbol{n}, 0)]\|/\|O_1\|\|O_2\|e^{(vt-x)/\xi}$ is less than $\|[O_1(0,t), O_2(x\boldsymbol{n}, 0)]\|/\|O_1\|\|O_2\| \leq 2$. So indeed $f(x,t)/g(x,t)$ is bounded in this region.

For $0 < t < t_0$ and $x \geq v_0 t$, $f(x,t)/g(x,t) = \|[O_1(0,t), O_2(x\boldsymbol{n}, 0)]\|/\|O_1\|\|O_2\|e^{(vt-x)/\xi}$ can be bounded using the Lieb-Robinson Theorem 3: there is some $C_0, v_0, \xi_0 > 0$ such that $\|[O_1(0,t), O_2(x\boldsymbol{n}, 0)]\| \leq C_0\|O_1\|\|O_2\|e^{(v_0 t - x)/\xi_0} \leq C_0\|O_1\|\|O_2\|e^{(v_0 t - x)/\xi}$ (by construction $\xi \geq \xi_0$) so $f(x,t)/g(x,t) \leq C_0 e^{(v_0 - v)t/\xi}$ which is a bounded function for $0 < t < t_0$.

Finally for $0 < t < t_0$ and $0 < x < v_0 t$, $f(x,t)/g(x,t)$ is bounded because it is continuous and the region is bounded. Hence we've shown that $C > 0$ is well-defined and with ξ appearing in (ii), (iii) is true. □

Henceforth the Lieb-Robinson velocity will be defined as $v_{\mathrm{LR}} \equiv v_{\mathrm{LR}}^{(1)} = v_{\mathrm{LR}}^{(2)}$. The technical uniformity condition is true for known examples. The same proof shows the equivalence of two definitions of the butterfly velocity. For future use only the definition corresponding to $v_{\mathrm{LR}}^{(2)}$ is recorded:

$$v_B(\boldsymbol{n}; O_1, O_2, \rho) \equiv \sup \left\{ v : \lim_{t\to\infty} \frac{1}{t} \ln \mathcal{C}_{O_1O_2}(vt\boldsymbol{n}, t; \rho) \geq 0 \right\}, \tag{2.47}$$

where the OTOC $\mathcal{C}_{O_1O_2}(\boldsymbol{x}, t; \rho)$ is defined in (2.1). As the velocity-dependent quantum Lyapunov exponent is defined as in (2.6), an equivalent definition of v_B reads:

$$v_B(\boldsymbol{n}; O_1, O_2, \rho) \equiv \sup\{v : \lambda_{O_1O_2}(v\boldsymbol{n}; \rho) \geq 0\}. \tag{2.48}$$

As expected, the butterfly velocity in any state is bounded by the Lieb-Robinson velocity:

Proposition 3. *$v_B(\boldsymbol{n}; O_1, O_2, \rho) \leq v_{\mathrm{LR}}(\boldsymbol{n}; O_1, O_2)$ for any $O_1, O_2 \in \mathcal{O}$, density matrix ρ and direction $\boldsymbol{n}$.*

Proof. This follows from definition (2.41) and (2.47), and $\mathcal{C}_{O_1O_2}(\boldsymbol{x}, t; \rho) \leq \|[O_1(0,t), O_2(\boldsymbol{x}, 0)]\|^2$. □

Finally the scrambling velocity can be precisely defined in the language of exponentially local operators, defined around (2.30). Let $O \equiv i[O_1(0,t), O_2(\boldsymbol{v}t, 0)]$, then[4]

$$v_S(\boldsymbol{v}; O_1, O_2, \xi) \equiv \inf_{C>0} \overline{\lim_{t\to\infty}} \frac{1}{t} \inf \{R \geq 0 : \exists\, \boldsymbol{x} \in \Lambda,\, O \in \mathcal{B}(\boldsymbol{x}, R; \xi, C)\}, \tag{2.49}$$

where the smallest ball, with radius R and centered at $\boldsymbol{x}$, is understood as roughly the "support" of the commutator O. The quantities ξ and C characterize the exponential tail that we neglected in the main text. Clearly $v_S \geq 0$ and decreases with increasing ξ.

For any triple $(v, \xi \equiv \xi_{\mathrm{LR}}, C_{\mathrm{LR}})$ from Theorem 1, we now show that $v_S(\boldsymbol{v}; \xi) \leq v$. Thus we have an upper bound of v_S by velocities with a Lieb-Robinson inequality. Note the ξ-dependence of v_S was omitted in the main text. More precisely, if $O_2(\boldsymbol{v}t, 0)$ is within the "support" of $O_1(0,t)$, for scrambling systems at late times we would expect $\|[O_1(0,t), O_2(\boldsymbol{v}t, 0)]\|$ to equilibrate to a nonzero constant value; if so, $v_S \leq v$:

Proposition 4. *Given $\boldsymbol{v}$, $\xi > 0$, $O_1, O_2 \in \mathcal{O}$, if for any $t > 0$, $O_1(0,t) \in \mathcal{B}(0, vt; \xi, C)$ for some $v > |\boldsymbol{v}|$, $C > 0$ and $\underline{\lim}_{t\to\infty} \|[O_1(0,t), O_2(\boldsymbol{v}t, 0)]\| > 0$, then $v_S(\boldsymbol{v}; O_1, O_2, \xi) \leq v$.*

Proof. Let $O(t) \equiv [O_1(0,t), O_2(\boldsymbol{v}t, 0)]$, $c \equiv \underline{\lim}_{t\to\infty} \|O(t)\| > 0$. As $|\boldsymbol{v}| < v$, $\mathcal{Q}_0^r[O_2(\boldsymbol{v}t, 0)] = 0$ for $r \geq vt$ at late times. Then $O(t) = [\mathcal{P}_0^r[O_1(0,t)], \mathcal{P}_0^r[O_2(\boldsymbol{v}t, 0)]] + [\mathcal{Q}_0^r[O_1(0,t)], \mathcal{P}_0^r[O_2(\boldsymbol{v}t, 0)]]$. But

[4] To make sure the limit exists, we have used the limit superior $\overline{\lim}$ and the limit inferior $\underline{\lim}$.

the first term is supported in the ball of radius r centered at origin, so

$$\begin{aligned}\|\mathcal{Q}_0^r[O(t)]\| &= \|\mathcal{Q}_0^r[\mathcal{Q}_0^r[O_1(0,t)], \mathcal{P}_0^r[O_2(\boldsymbol{v}t,0)]]\| \\ &\leq 4\|\mathcal{Q}_0^r[O_1(0,t)]\|\|\mathcal{P}_0^r[O_2(\boldsymbol{v}t,0)]\| \leq 4C\|O_1\|\|O_2\|e^{(vt-r)/\xi},\end{aligned} \tag{2.50}$$

where we have used the definition (2.30) that for all $t > 0$ and $r \geq vt$, $\|\mathcal{Q}_0^r[O_1(0,t)]\| \leq C\|O_1\|e^{(vt-r)/\xi}$ with the inequalities (2.29).

So there is a time $t_0 > 0$ that for all $t > t_0$, $\|O(t)\| \geq c/2$ as well as $\|\mathcal{Q}_0^r[O(t)]\| \leq 4C\|O_1\|\|O_2\|e^{(vt-r)/\xi}$ for all $r \geq vt$. Hence $\|\mathcal{Q}_0^r[O(t)]\| \leq C'\|O(t)\|e^{(vt-r)/\xi}$, for all $t > t_0$ and $r \geq vt$, if we choose $C' = 8C\|O_1\|\|O_2\|/c$. That is, $O(t) \in \mathcal{B}(0, vt; \xi, C')$ for $t > t_0$ hence by definition (2.49), $v_S(\boldsymbol{v}; O_1, O_2, \xi) \leq v$. □

All velocities can be maximized over direction $\boldsymbol{n}$ to recover their isotropic definitions, or over $O_1, O_2 \in \mathcal{O}$ to remove the operator dependence.

2.9 Appendix D: Bounds for exponentially local operators

In this section we collect some lemmas and generalize Theorem 2 and the exponential clustering condition (2.38) to exponentially local operators. Readers are encouraged to review sections 2.6 and 2.7. The following inequality will be useful: for any $A, B \geq 0$ and $k, \gamma > 0$,

$$\sum_{n=\lceil k \rceil}^{\infty} (An + B)e^{-\gamma n} \leq (Ak + A + B)e^{-\gamma k}(1 - e^{-\gamma})^{-2}, \tag{2.51}$$

where $\lceil x \rceil$ denotes the least integer greater than or equal to x. To show this, by doing the summation exactly it is easy to check that for any $A, B \geq 0$, $\gamma > 0$ and integer $m \geq 1$,

$$\sum_{n=m}^{\infty} (An + B)e^{-\gamma n} \leq (Am + B)e^{-\gamma m}(1 - e^{-\gamma})^{-2}, \tag{2.52}$$

and the inequality (2.51) follows because if $m = \lceil k \rceil$, $m \leq k + 1$ in the linear factor and $k \leq m$ implies that $e^{-\gamma m} \leq e^{-\gamma k}$ as well.

The following lemma bounds the product of two exponentially local operators:

Lemma 2. *Let $O_1 \in \mathcal{B}(\boldsymbol{x}, R; \xi_1, C_1)$ and $O_2 \in \mathcal{B}(\boldsymbol{x}, R; \xi_2, C_2)$, then for any $r \geq R$,*

$$\|\mathcal{Q}_{\boldsymbol{x}}^r[O_1 O_2]\| \leq 2(C_1 + C_2)\|O_1\|\|O_2\|e^{(R-r)/\max\{\xi_1, \xi_2\}}. \tag{2.53}$$

Proof. Note that for any $r > 0$, $O_1 O_2 = \mathcal{P}_{\boldsymbol{x}}^r[O_1]\mathcal{P}_{\boldsymbol{x}}^r[O_2] + O_1\mathcal{Q}_{\boldsymbol{x}}^r[O_2] + \mathcal{Q}_{\boldsymbol{x}}^r[O_1]\mathcal{P}_{\boldsymbol{x}}^r[O_2]$, and

$\mathcal{Q}^r_{\boldsymbol{x}}[\mathcal{P}^r_{\boldsymbol{x}}[O_1]\mathcal{P}^r_{\boldsymbol{x}}[O_2]] = 0$. So by (2.29) and (2.30), for $r \geq R$,

$$\begin{aligned}\|\mathcal{Q}^r_{\boldsymbol{x}}[O_1O_2]\| &\leq 2\|O_1\|\|\mathcal{Q}^r_{\boldsymbol{x}}[O_2]\| + 2\|\mathcal{Q}^r_{\boldsymbol{x}}[O_1]\|\|O_2\| \\ &\leq 2C_2\|O_1\|\|O_2\|e^{(R-r)/\xi_2} + 2C_1\|O_1\|\|O_2\|e^{(R-r)/\xi_1}. \end{aligned} \tag{2.54}$$

□

Next is the Araki bound (cf. Theorem 2) for exponentially local operators:

Theorem 4. *For any one-dimensional Gibbs state ρ as defined in (2.32) with $\mu_i \in \mathbb{C}$ and operator $O \in \mathcal{B}(\boldsymbol{x}, R; \xi, C)$, there exists $C'(\mu_i, \xi, C) > 0$ (dependent on lattice geometry and C^i as well) such that for all $r \geq R + l_{\mathrm{A}}(\mu_i) + a$,*

$$\|\rho O\rho^{-1}\| \leq C'(\mu_i, \xi, C)\|O\|(1 + 2R/a), \tag{2.55}$$

$$\|\mathcal{Q}^r_{\boldsymbol{x}}[\rho O\rho^{-1}]\| \leq C'(\mu_i, \xi, C)\|O\|[1 + 2(r - l_{\mathrm{A}}(\mu_i))/a]e^{(R+l_{\mathrm{A}}(\mu_i)+a-r)/(\xi_{\mathrm{A}}+\xi)}. \tag{2.56}$$

Here $l_{\mathrm{A}}(\mu_i)$ and ξ_{A} are those appearing in the Araki bound, and a is the lattice spacing.

Proof. For the first inequality, let $m \equiv \lceil (R + a)/a \rceil$. Decompose $O = \mathcal{P}^{(m-1)a}_{\boldsymbol{x}}[O] + \sum_{n \geq m} O_n$, where $O_n \equiv \mathcal{P}^{na}_{\boldsymbol{x}}\mathcal{Q}^{(n-1)a}_{\boldsymbol{x}}[O] = \mathcal{P}^{na}_{\boldsymbol{x}}[O] - \mathcal{P}^{(n-1)a}_{\boldsymbol{x}}[O]$. Then by Theorem 2 with (2.29) and (2.30), for $n \geq m$,

$$\begin{aligned}\|\rho O_n\rho^{-1}\| = \|\rho\mathcal{P}^{na}_{\boldsymbol{x}}\mathcal{Q}^{(n-1)a}_{\boldsymbol{x}}[O]\rho^{-1}\| &\leq C_{\mathrm{A}}(\mu_i)(2n+1)\|\mathcal{P}^{na}_{\boldsymbol{x}}\mathcal{Q}^{(n-1)a}_{\boldsymbol{x}}[O]\| \\ &\leq C_{\mathrm{A}}(\mu_i)(2n+1)\|\mathcal{Q}^{(n-1)a}_{\boldsymbol{x}}[O]\| \leq C_{\mathrm{A}}(\mu_i)(2n+1)C\|O\|e^{(R-na+a)/\xi}. \end{aligned} \tag{2.57}$$

Also by Theorem 2, $\|\mathcal{P}[O]\| \leq \|O\|$ and $m \leq (R + a)/a + 1 = R/a + 2$,

$$\|\rho\mathcal{P}^{(m-1)a}_{\boldsymbol{x}}[O]\rho^{-1}\| \leq C_{\mathrm{A}}(\mu_i)(2m-1)\|\mathcal{P}^{(m-1)a}_{\boldsymbol{x}}[O]\| \leq C_{\mathrm{A}}(\mu_i)(2R/a+3)\|O\|. \tag{2.58}$$

Sum (2.57) with (2.51) (where $A = 2$, $B = 1$, $k = (R + a)/a$ and $\gamma = a/\xi$) to get the bound

$$\|\rho O\rho^{-1}\| \leq C_{\mathrm{A}}(\mu_i)(2R/a+3)\|O\| + C_{\mathrm{A}}(\mu_i)C(2R/a+5)\|O\|(1 - e^{-a/\xi})^{-2}. \tag{2.59}$$

Denote $C_1(\mu_i, \xi, C) = 3C_{\mathrm{A}}(\mu_i) + 5C_{\mathrm{A}}(\mu_i)C(1 - e^{-a/\xi})^{-2}$, so that

$$\|\rho O\rho^{-1}\| \leq C_1(\mu_i, \xi, C)\|O\|(1 + 2R/a). \tag{2.60}$$

For the second inequality, expand $O = \mathcal{P}_\emptyset[O] + \sum_{n \geq 0} O_n$, where $\mathcal{P}_\emptyset[O]$ is proportional to identity

and $O_n \equiv \mathcal{P}_{\boldsymbol{x}}^{na}\mathcal{Q}_{\boldsymbol{x}}^{(n-1)a}[O] = \mathcal{P}_{\boldsymbol{x}}^{na}[O] - \mathcal{P}_{\boldsymbol{x}}^{(n-1)a}[O]$. Because $\mathcal{Q}_{\boldsymbol{x}}^{r}[I] = 0$,

$$\mathcal{Q}_{\boldsymbol{x}}^{r}[\rho O \rho^{-1}] = \sum_{n=0}^{\infty} \mathcal{Q}_{\boldsymbol{x}}^{r}[\rho O_n \rho^{-1}]. \tag{2.61}$$

Let $\delta \equiv \alpha(r - l_{\mathrm{A}}(\mu_i) - R - a) \geq 0$ for any $0 < \alpha < 1$ and split the sum (2.61) into two parts: $0 \leq na < R + \delta + a$ and $na \geq R + \delta + a$. Apply Theorem 2 for the first part (also note $\|O_n\| \leq 2\|O\|$ by (2.29)):

$$\|\mathcal{Q}_{\boldsymbol{x}}^{r}[\rho O_n \rho^{-1}]\| \leq 2C_{\mathrm{A}}(\mu_i)(2n+1)\|O\| e^{(l_{\mathrm{A}}(\mu_i)+na-r)/\xi_{\mathrm{A}}}, \tag{2.62}$$

and further with inequalities (2.29) and definition (2.30) for the second part:

$$\begin{aligned}\|\mathcal{Q}_{\boldsymbol{x}}^{r}[\rho O_n \rho^{-1}]\| &\leq 2\|\rho O_n \rho^{-1}\| \leq 2C_{\mathrm{A}}(\mu_i)(2n+1)\|O_n\| \\ &\leq 2C_{\mathrm{A}}(\mu_i)(2n+1)\|\mathcal{Q}_{\boldsymbol{x}}^{(n-1)a}[O]\| \leq 2CC_{\mathrm{A}}(\mu_i)(2n+1)\|O\| e^{(R-na+a)/\xi}.\end{aligned} \tag{2.63}$$

Overall, sum (2.62) as geometric series after applying $n \leq k$ and sum (2.63) with (2.51) (where $A = 2$, $B = 1$, $k = (R + \delta + a)/a$ and $\gamma = a/\xi$):

$$\begin{aligned}\|\mathcal{Q}_{\boldsymbol{x}}^{r}[\rho O \rho^{-1}]\| &\leq 2C_{\mathrm{A}}(\mu_i)(2k+1)\|O\| e^{(l_{\mathrm{A}}(\mu_i)+ka-r)/\xi_{\mathrm{A}}}(1 - e^{-a/\xi_{\mathrm{A}}})^{-1} \\ &\quad + 2CC_{\mathrm{A}}(\mu_i)(2k+3)\|O\| e^{(R-ka+a)/\xi}(1 - e^{-a/\xi})^{-2} \\ &\leq 2C_{\mathrm{A}}(\mu_i)[1 + 2(r - l_{\mathrm{A}}(\mu_i))/a]\|O\| e^{-(1-\alpha)\delta/\alpha\xi_{\mathrm{A}}}(1 - e^{-a/\xi_{\mathrm{A}}})^{-1} \\ &\quad + 2CC_{\mathrm{A}}(\mu_i)[3 + 2(r - l_{\mathrm{A}}(\mu_i))/a]\|O\| e^{-\delta/\xi}(1 - e^{-a/\xi})^{-2},\end{aligned} \tag{2.64}$$

where in the second inequality we have replaced $ka = R + \delta + a$ in the exponents and applied the bound $k \leq (r - l_{\mathrm{A}}(\mu_i))/a$ (because $\alpha \leq 1$) in the prefactors. Now

$$\|\mathcal{Q}_{\boldsymbol{x}}^{r}[\rho O \rho^{-1}]\| \leq C_2(\mu_i, \xi, C)\|O\|[1 + 2(r - l_{\mathrm{A}}(\mu_i))/a] e^{(R + l_{\mathrm{A}}(\mu_i) + a - r)/(\xi_{\mathrm{A}} + \xi)}, \tag{2.65}$$

if one chooses $\alpha = \xi/(\xi_{\mathrm{A}} + \xi)$ to equate the exponents and $C_2(\mu_i, \xi, C) = 2C_{\mathrm{A}}(\mu_i)(1 - e^{-a/\xi_{\mathrm{A}}})^{-1} + 6CC_{\mathrm{A}}(\mu_i)(1 - e^{-a/\xi})^{-2}$.

Finally it suffices to choose $C'(\mu_i, \xi, C) \equiv \max\{C_1(\mu_i, \xi, C), C_2(\mu_i, \xi, C)\}$. □

Observe that the operator $\rho O \rho^{-1}$ as stated in (2.56), is not exponentially local explicitly (due to the prefactor that is linear in r). To work around this the following corollary of Theorem 4 is particularly useful:

Corollary 1. *For any $\varepsilon > 0$, there is a $\widetilde{C}'(\mu_i, \xi, C, \varepsilon)$ such that*

$$\rho O \rho^{-1} \in \mathcal{B}\left(\boldsymbol{x}, R + l_{\mathrm{A}}(\mu_i) + a; \xi_{\mathrm{A}} + \xi + \varepsilon, \widetilde{C}' e^{\varepsilon R/(\xi_{\mathrm{A}} + \xi)^2}\|O\|/\|\rho O \rho^{-1}\|\right). \tag{2.66}$$

Proof. First note that for $\zeta(\xi) \equiv \xi_{\mathrm{A}} + \xi$,

$$e^{R/\zeta} = e^{R/(\zeta+\varepsilon)} e^{\varepsilon R/\zeta(\zeta+\varepsilon)} \leq e^{R/(\zeta+\varepsilon)} e^{\varepsilon R/\zeta^2}, \tag{2.67}$$

so it suffices to find $\widetilde{C}'(\mu_i, \xi, C, \varepsilon)$ such that for all $x \equiv r - l_{\mathrm{A}}(\mu_i) - a \geq 0$,

$$C'(\mu_i, \xi, C)[1 + 2(x+a)/a] e^{-x/\zeta(\xi)} \leq \widetilde{C}' e^{-x/(\zeta(\xi)+\varepsilon)}, \tag{2.68}$$

which clearly exists. □

Finally we generalize inequality (2.38) to exponentially local operators as well; for future use we will work in one dimension only:

Theorem 5. *Let ρ be a one-dimensional state with ξ, C and $l_0(\cdot) > 0$ as stated around (2.38). If $O_1 \in \mathcal{B}(\boldsymbol{x}, R_1; \xi_1, C_1)$, $O_2 \in \mathcal{B}(\boldsymbol{y}, R_2; \xi_2, C_2)$ and $|\boldsymbol{x} - \boldsymbol{y}| \geq l_0(2) + R_1 + R_2$,*

$$\begin{aligned} &|\mathrm{tr}(\rho\, O_1 O_2) - \mathrm{tr}(\rho\, O_1)\mathrm{tr}(\rho\, O_2)| \\ &\qquad \leq 2(C + C_1 + C_2)\|O_1\|\|O_2\| e^{(R_1+R_2+l_0(2)-|\boldsymbol{x}-\boldsymbol{y}|)/(\xi+\xi_1+\xi_2)}. \end{aligned} \tag{2.69}$$

Proof. Let $\Delta \equiv |\boldsymbol{x} - \boldsymbol{y}| - l_0(2) - R_1 - R_2 \geq 0$, and define $r \equiv R_1 + \alpha_1 \Delta$ and $s \equiv R_2 + \alpha_2 \Delta$ for $\alpha_1, \alpha_2 > 0$ and $\alpha_1 + \alpha_2 < 1$. Denote $c(O_1, O_2) \equiv \mathrm{tr}(\rho\, O_1 O_2) - \mathrm{tr}(\rho\, O_1)\mathrm{tr}(\rho\, O_2)$ for convenience and observe $|c(O_1, O_2)| \leq 2\|O_1\|\|O_2\|$. Then

$$c(O_1, O_2) = c(\mathcal{P}^r_{\boldsymbol{x}}[O_1], \mathcal{P}^s_{\boldsymbol{y}}[O_2]) + c(\mathcal{Q}^r_{\boldsymbol{x}}[O_1], \mathcal{P}^s_{\boldsymbol{y}}[O_2]) + c(O_1, \mathcal{Q}^s_{\boldsymbol{y}}[O_2]). \tag{2.70}$$

By inequality (2.38), (note $\delta = 2$ if S and T are intervals in (2.38) and $\|\mathcal{P}[O]\| \leq \|O\|$)

$$|c(\mathcal{P}^r_{\boldsymbol{x}}[O_1], \mathcal{P}^s_{\boldsymbol{y}}[O_2])| \leq 2C\|O_1\|\|O_2\| e^{-l_0(2)/\xi} e^{-(1-\alpha_1-\alpha_2)\Delta/\xi}, \tag{2.71}$$

and by definition (2.30),

$$|c(\mathcal{Q}^r_{\boldsymbol{x}}[O_1], \mathcal{P}^s_{\boldsymbol{y}}[O_2])| \leq 2\|\mathcal{Q}^r_{\boldsymbol{x}}[O_1]\|\|O_2\| \leq 2C_1\|O_1\|\|O_2\| e^{-\alpha_1\Delta/\xi_1}, \tag{2.72}$$

$$|c(O_1, \mathcal{Q}^s_{\boldsymbol{y}}[O_2])| \leq 2\|O_1\|\|\mathcal{Q}^s_{\boldsymbol{y}}[O_2]\| \leq 2C_2\|O_1\|\|O_2\| e^{-\alpha_2\Delta/\xi_2}. \tag{2.73}$$

Now choose $\alpha_1 = \xi_1/(\xi + \xi_1 + \xi_2)$ and $\alpha_2 = \xi_2/(\xi + \xi_1 + \xi_2)$ so that the exponents with Δ are all equal. Sum them up to get (2.69). □

2.10 Appendix E: Proof of the bound

In this section we give a proof of the bounds stated in the main text. To avoid clutter of notations, all quantities in this section may depend on lattice geometry, Hamiltonian H (2.31) and charges C^i (2.33) implicitly.

Theorem 6. *For any one-dimensional Gibbs state ρ as defined in (2.32) with correlation length $\xi_{\rm cor}$ (read around (2.38) for a definition), $\varepsilon, \delta > 0$, any operators O_1, O_2 and $\boldsymbol{x} \in \Lambda$, $t > 0$, there exist $A(\mu_i, \xi, C, \varepsilon)$, $B(\mu_i) > 0$ such that*

$$\left|\frac{\partial \mathcal{C}_{O_1O_2}(\boldsymbol{x}, t; \rho)}{\partial \mu_i}\right| \leq A \sup_{\boldsymbol{y}\in\Lambda} \|c^i_{\boldsymbol{y}}\| \|O_1\|^2 \|O_2\|^2 (1 + 2R/a) e^{\varepsilon R/(\xi+\xi_{\rm A})^2} e^{-\delta/(\xi_{\rm cor}+\xi_{\rm A}+\xi+\varepsilon)} + 2c^i \left(R + \delta + B\right) \mathcal{C}_{O_1O_2}(\boldsymbol{x}, t; \rho)/a, \tag{2.74}$$

and

$$\left|\frac{\partial \mathcal{C}_{O_1O_2}(\boldsymbol{x}, t; \rho)}{\partial J_\alpha}\right| \leq A\beta \sup_{\boldsymbol{y}\in\Lambda} \|h^\alpha_{\boldsymbol{y}}\| \|O_1\|^2 \|O_2\|^2 (1 + 2R/a) e^{\varepsilon R/(\xi+\xi_{\rm A})^2} e^{-\delta/(\xi_{\rm cor}+\xi_{\rm A}+\xi+\varepsilon)} + 2\beta h^\alpha \left(R + \delta + B\right) \mathcal{C}_{O_1O_2}(\boldsymbol{x}, t; \rho)/a + 2\int_0^t ds \sqrt{\mathcal{C}_{O_1O_2}(\boldsymbol{x}, t; \rho) \mathcal{C}_{[H^\alpha(-s), O_1]O_2}(\boldsymbol{x}, t; \rho)}, \tag{2.75}$$

where a is the lattice spacing and $\xi_{\rm A}$ is defined in Theorem 2. The inverse temperature is denoted as β and J_α labels couplings in the Hamiltonian (2.31). Denote $O \equiv i[O_1(0,t), O_2(\boldsymbol{x}, 0)]$; R, ξ and C are such that $O \in \mathcal{B}(\boldsymbol{y}_0, R; \xi, C)$ for some $\boldsymbol{y}_0 \in \Lambda$. Finally

$$c^i \equiv \int_0^1 ds\, c^i(s) \equiv \int_0^1 ds \sup_{\boldsymbol{y}\in\Lambda} \frac{|\mathrm{tr}(\rho^s \widetilde{c}^i_{\boldsymbol{y}} \rho^{1-s} O^\dagger O)|}{\mathrm{tr}(\rho\, O^\dagger O)}, \tag{2.76}$$

where $\widetilde{c}^i_{\boldsymbol{y}} \equiv c^i_{\boldsymbol{y}} - \mathrm{tr}(\rho\, c^i_{\boldsymbol{y}})$, and same for h^α with $c^i_{\boldsymbol{y}}$ replaced by $h^\alpha_{\boldsymbol{y}}$. And if C^i commute with each other, c^i can be chosen as

$$c^i \equiv \sup_{\boldsymbol{y}\in\Lambda} \frac{|\mathrm{tr}(\sqrt{\rho}\, \widetilde{c}^i_{\boldsymbol{y}} \sqrt{\rho} O^\dagger O)|}{\mathrm{tr}(\rho\, O^\dagger O)} \leq 2 \sup_{\boldsymbol{y}\in\Lambda} \|c^i_{\boldsymbol{y}}\|. \tag{2.77}$$

Proof. We start with proving (2.74). By definition (2.1) and (2.32),

$$\frac{\partial \mathcal{C}_{O_1O_2}(\boldsymbol{x}, t; \rho)}{\partial \mu_i} = -\int_0^1 ds\, \mathrm{tr}(\rho^s \widetilde{C}^i \rho^{1-s} O^\dagger O), \tag{2.78}$$

where for any operator C, $\widetilde{C} \equiv C - \mathrm{tr}(\rho\, C)$. Now recall C^i is a sum of local terms (2.33):

$$C^i = \sum_{\boldsymbol{y}\in S(r)} c^i_{\boldsymbol{y}} + \sum_{\boldsymbol{y}\in\Lambda - S(r)} c^i_{\boldsymbol{y}}, \tag{2.79}$$

for any $S(r) \equiv \{\boldsymbol{y} \in \Lambda : |\boldsymbol{y} - \boldsymbol{y}_0| \leq r\}$. For any $\boldsymbol{y} \in \Lambda$, by definition of $c^i(s)$,

$$|\mathrm{tr}(\rho^s \widetilde{c}^i_{\boldsymbol{y}} \rho^{1-s} O^\dagger O)| \leq c^i(s)\, \mathrm{tr}(\rho\, O^\dagger O). \tag{2.80}$$

The inequality (2.80) is good enough for the terms in $S(r)$. For the remaining terms with $\boldsymbol{y}$ away from $\boldsymbol{y}_0$ we have a better estimate because connected correlation decays when operators are far apart. There is a technical complication due to the fact that the factors of ρ are separated by – and do not necessarily commute with – the $\widetilde{c}^i_{\boldsymbol{y}}$. For this reason we need to use the Araki bound to show that operators remain sufficiently local under conjugation by the density matrix. Indeed by Lemma 2, $O^\dagger O \in \mathcal{B}(\boldsymbol{y}_0, R; \xi, 4C)$ and from Theorem 4 and Corollary 1, there is $C_1(\mu_i, \xi, C, \varepsilon) > 0$ and $l(\mu_i) > 0$ such that for any $0 \leq s \leq 1$,

$$\|\rho^{-s} O^\dagger O \rho^s\| \leq C_1 \|O^\dagger O\|(1 + 2R/a), \tag{2.81}$$

$$\rho^{-s} O^\dagger O \rho^s \in \mathcal{B}\left(\boldsymbol{y}_0, R + l(\mu_i) + a; \xi_{\mathrm{A}} + \xi + \varepsilon, C_1 e^{\varepsilon R/(\xi_{\mathrm{A}}+\xi)^2} \|O^\dagger O\| / \|\rho^{-s} O^\dagger O \rho^s\|\right). \tag{2.82}$$

Hence by Theorem 5, because $\mathrm{tr}(\rho\, \widetilde{c}^i_{\boldsymbol{y}}) = 0$, for any $0 \leq s \leq 1$,

$$\begin{aligned} |\mathrm{tr}(\rho^s \widetilde{c}^i_{\boldsymbol{y}} \rho^{1-s} O^\dagger O)| &= |\mathrm{tr}(\rho \rho^{-s} O^\dagger O \rho^s \widetilde{c}^i_{\boldsymbol{y}})| \\ &\leq 2C_2 e^{R + l(\mu_i) + a + R_H + l_0(2) - |\boldsymbol{y} - \boldsymbol{y}_0|)/(\xi_{\mathrm{cor}} + \xi_{\mathrm{A}} + \xi + \varepsilon)}, \end{aligned} \tag{2.83}$$

where C_2 is defined in terms of the prefactor $C_{\mathrm{cor}}(\mu_i)$ in (2.38) as, using (2.81),

$$\begin{aligned} C_2 &\equiv C_{\mathrm{cor}} \sup_{\boldsymbol{y} \in \Lambda} \|\widetilde{c}^i_{\boldsymbol{y}}\| \|\rho^{-s} O^\dagger O \rho^s\| + C_1 e^{\varepsilon R/(\xi_{\mathrm{A}}+\xi)^2} \sup_{\boldsymbol{y} \in \Lambda} \|\widetilde{c}^i_{\boldsymbol{y}}\| \|O^\dagger O\| \\ &\leq C_{\mathrm{cor}} C_1 \sup_{\boldsymbol{y} \in \Lambda} \|\widetilde{c}^i_{\boldsymbol{y}}\| \|O\|^2 (1 + 2R/a) + C_1 e^{\varepsilon R/(\xi_{\mathrm{A}}+\xi)^2} \sup_{\boldsymbol{y} \in \Lambda} \|\widetilde{c}^i_{\boldsymbol{y}}\| \|O\|^2. \end{aligned} \tag{2.84}$$

Now bound the sum (2.79) by choosing $r = R + l(\mu_i) + a + R_H + l_0(2) + \delta$ and apply (2.80) for $\boldsymbol{y} \in S(r)$ and (2.83) for $\boldsymbol{y} \notin S(r)$, (denote $\zeta \equiv \xi_{\mathrm{cor}} + \xi_{\mathrm{A}} + \xi + \varepsilon$)

$$|\mathrm{tr}(\rho^s \widetilde{C}^i \rho^{1-s} O^\dagger O)| \leq c^i(s)(1 + 2r/a)\mathrm{tr}(\rho\, O^\dagger O) + 4C_2 e^{-\delta/\zeta}(1 - e^{-a/\zeta})^{-1}, \tag{2.85}$$

and use the inequality (2.84) and $\|\widetilde{c}^i_{\boldsymbol{y}}\| \leq 2\|c^i_{\boldsymbol{y}}\|$ to reduce to the form (2.74).

Proving (2.75) is essentially the same except in the first step:

$$\frac{\partial \mathcal{C}_{O_1 O_2}(\boldsymbol{x}, t; \rho)}{\partial J_\alpha} = -\beta \int_0^1 ds\, \mathrm{tr}(\rho^s \widetilde{H}^\alpha \rho^{1-s} O^\dagger O) + 2\,\mathrm{Re}\,\mathrm{tr}\left(\rho\, O^\dagger \frac{\partial O}{\partial J_\alpha}\right), \tag{2.86}$$

there is an additional term due to coupling dependence of $O_1(0,t)$. By definition,

$$\frac{\partial O}{\partial J_\alpha} = -\int_0^t ds\ [[H^\alpha(s), O_1(0,t)], O_2(\boldsymbol{x},0)]\,, \tag{2.87}$$

and (2.75) follows from the Cauchy-Schwartz inequality for the inner product $\langle O_1, O_2\rangle \equiv \mathrm{tr}(\rho O_1^\dagger O_2)$.

Finally if C^i commute with each other, the first step (2.78) can be replaced with

$$\frac{\partial \mathcal{C}_{O_1O_2}(\boldsymbol{x},t;\rho)}{\partial \mu_i} = -\mathrm{tr}(\sqrt{\rho}\widetilde{C}^i\sqrt{\rho}O^\dagger O), \tag{2.88}$$

and the same proof goes through with c^i as in (2.77). It is bounded by $2\sup\|c^i_{\boldsymbol{y}}\|$ because $\sqrt{\rho}\,O^\dagger O\sqrt{\rho}$ is a positive operator and for any operator S and positive operator T, $|\mathrm{tr}ST| \le \|S\|\mathrm{tr}\,T$. □

The theorem, as stated, seems complicated; but the physics is much clearer in terms of the velocity-dependent Lyapunov exponent (2.6):

Corollary 2. *For $v_S(\boldsymbol{v}; O_1, O_2, \xi)$ defined in (2.49),*

$$\left|\frac{\partial \lambda_{O_1O_2}(\boldsymbol{v};\rho)}{\partial \mu_i}\right| \le \frac{2c^i}{a}\Big(v_S(\boldsymbol{v}; O_1, O_2, \xi) - \lambda_{O_1O_2}(\boldsymbol{v};\rho)(\xi_{\mathrm{cor}} + \xi)\Big). \tag{2.89}$$

Proof. Divide both sides of (2.74) by $t\,\mathcal{C}_{O_1O_2}(\boldsymbol{x},t;\rho)$, choose

$$\delta(t) = (\xi_{\mathrm{cor}} + \xi_{\mathrm{A}} + \xi + \varepsilon)\left[-\lambda_{O_1O_2}(\boldsymbol{v};\rho)t + \varepsilon R/(\xi+\xi_{\mathrm{A}})^2 + \varepsilon t\right] > 0, \tag{2.90}$$

$\boldsymbol{x} = \boldsymbol{v}t$ and take the limit $t\to\infty$ (assuming the limit and derivative commute):

$$\left|\frac{\partial \lambda_{O_1O_2}}{\partial \mu_i}\right| \le 2c^i\left\{v_S + \left[\varepsilon + \varepsilon v_S/(\xi+\xi_{\mathrm{A}})^2 - \lambda_{O_1O_2}\right](\xi_{\mathrm{cor}} + \xi_{\mathrm{A}} + \xi + \varepsilon)\right\}/a. \tag{2.91}$$

Finally let $\varepsilon, \xi_{\mathrm{A}} \to 0$ to conclude[5]. □

The operator O must decay at large distances at least as quickly as the rate set by ξ_{LR} (appearing in any triple $(v, \xi_{\mathrm{LR}}, C_{\mathrm{LR}})$ with a Lieb-Robinson bound Theorem 1). Therefore we take $\xi = \xi_{\mathrm{LR}}$ in the main text. We have already noted in section 2.8 that this then defines a $v_S(\boldsymbol{v};\xi) \le v$.

The coupling dependence of $\lambda_{O_1O_2}(\boldsymbol{v};\rho)$ can be bounded in the same way:

Corollary 3. *If $\mathcal{C}_{O_1O_2}(\boldsymbol{v}t, t;\rho) \sim \kappa_1^2 e^{\lambda_{O_1O_2}(\boldsymbol{v};\rho)t}$ and*

$$\mathcal{C}_{[H^\alpha(-s),O_1]O_2}(\boldsymbol{v}t, t;\rho) \sim \kappa_2^2\|h^\alpha\|^2 e^{\lambda_{O_1O_2}(\boldsymbol{v};\rho)t} \tag{2.92}$$

[5]Regarding the limit $\xi_{\mathrm{A}} \to 0$ we refer readers to the discussions following Theorem 2.

for $\kappa_1, \kappa_2 > 0$ *and* $\|h^\alpha\| \equiv \sup_{\boldsymbol{y}\in\Lambda} \|h^\alpha_{\boldsymbol{y}}\|$ *at* $t \to \infty$,

$$\left|\frac{\partial \lambda_{O_1O_2}(\boldsymbol{v};\rho)}{\partial J_\alpha}\right| \leq \frac{2\beta h^\alpha}{a}\Big(v_S(\boldsymbol{v};O_1,O_2,\xi) - \lambda_{O_1O_2}(\boldsymbol{v};\rho)(\xi_{\text{cor}}+\xi)\Big) + 2\|h^\alpha\|\kappa_2/\kappa_1. \tag{2.93}$$

If we assume that the growth rate of the OTOC does not depend on choices of operators, i.e., the growth rate in (2.92) is $\lambda_{O_1O_2}(\boldsymbol{v};\rho)$, the same as that of $\mathcal{C}_{O_1O_2}(\boldsymbol{v}t,t;\rho)$, this corollary shows that divergence of $\partial_J v_B$ at zero temperature pinpoints quantum phase transitions at which the system becomes gapless. Indeed, if to the contrary the system is gapped, as observed in Fig. 2.3 and discussed in the main text, the first term on the right side of (2.93) is expected to vanish at zero temperature so the right-hand side of (2.93) should be finite, contradicting the divergence of $\partial_J v_B$ via an inequality similar to (2.23). Cusps of scrambling characteristics are indeed observed at quantum critical points in e.g. [152, 223].

2.11 Appendix F: Numerical details

Our method is a generalization of the Matrix Product Operator (MPO) approach to calculating the butterfly velocity, presented in [256], to finite temperature states. The algorithm is implemented with the ITensor library, with operators $O_1(0,t)$, $O_2(\boldsymbol{x},0)$ and thermal density matrix ρ represented as MPOs and evolved with a Time-Evolving Block Decimation (TEBD) method (for MPOs). For general quantum systems the thermal entanglement entropy is expected to be extensive. We find in practice that the MPO representation of thermal states works at sufficiently high but finite temperatures (in our case, $0 \leq \beta J \leq 3$). Numerical truncation ε in the MPO is set to $\varepsilon = 10^{-14}$ and maximal bond dimension is denoted as $\chi = 256$. We will only investigate the mixed field Ising model with hopping J and external fields h_X and h_Z as defined in (2.21), and probe the OTOC with Pauli Z operators ($O_1 = O_2 = Z$ in (2.1)). Scrambling characteristics are then determined by least-squares fitting of $\ln \mathcal{C}$ at the wavefront to the expression (2.2).

The wavefront is determined as follows. First, due to numerical truncation with $\varepsilon = 10^{-14}$ only data with $\ln \mathcal{C} > -22$ will be used. This delimits the right end r of the wavefront; the default left end l_0 is then defined as the position where $\partial_x \ln \mathcal{C}$ is half the value at r. To eliminate the arbitrariness of l_0 a hyperparameter $\delta > 0$ is introduced and the left end $l \equiv r - (r - l_0)\delta$. When $\delta = 1$, $l = l_0$ and when $\delta = 0$, $l = r$; hence δ tunes the range of the wavefront, ending at r.

As a sanity check our implementation is verified against Exact Diagonalization (ED), which may be regarded as the MPO approximation with no bond dimension restrictions ($\chi = \infty$). The result is shown in Fig. 2.4. From the figure the MPO algorithm matches with ED perfectly at times before maximal bond dimension restriction is reached and starts to deviate afterwards. However, as shown in the figure, the wavefront dynamics is well captured by the MPO approximation, even after the bond dimension is saturated inside the butterfly cone. Such effectiveness of MPO (at least

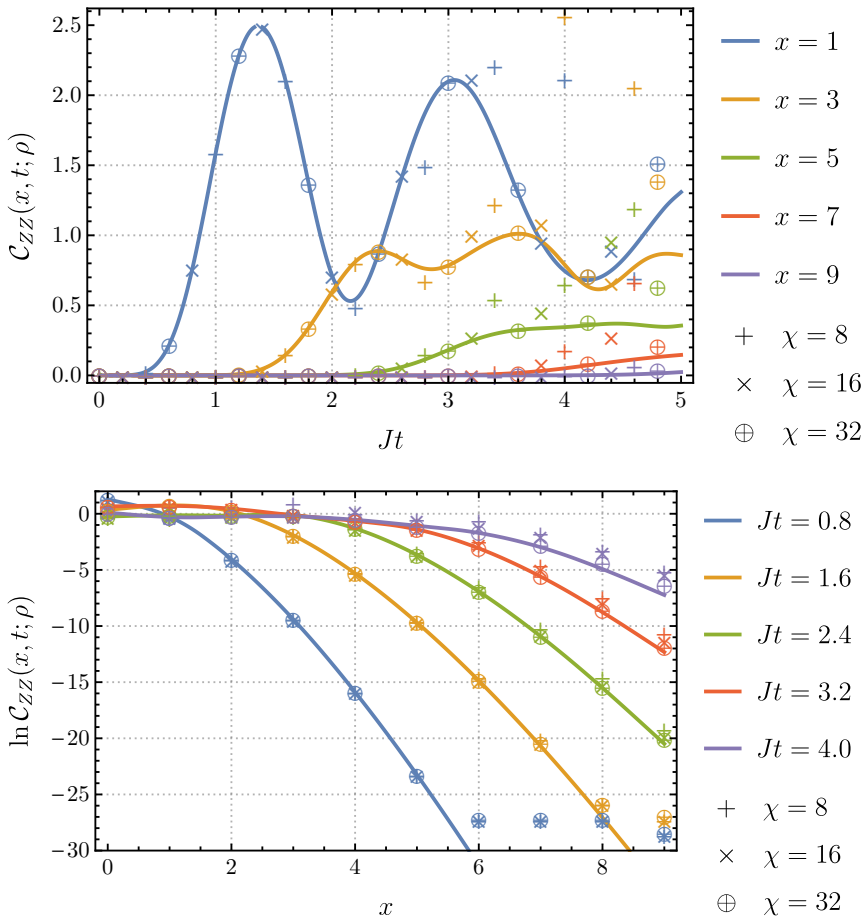


Figure 2.4: Comparison with Exact Diagonalization. The solid curves are from ED numerics in a mixed field Ising chain with $N = 10$, $h_X = 1.05J$ and $h_Z = 0.5J$ (see (2.21) for the Hamiltonian) and ρ is the thermal state with $T = J$. In the first panel, each curve shows the time dependence of the OTOC at a fixed distance ($O_1 = Z_1$ and $O_2 = Z_{x+1}$). For finite bond dimension truncations $\chi = 8, 16$ and 32, the MPO result agrees with ED at early times, and starts to deviate when the truncation is reached, which is near $Jt = 2, 3$ and 4 respectively. In the second panel, each curve is a spatial profile of the OTOC at a fixed time. Propagation of a butterfly wavefront is clearly observed. For all χ the agreement with ED is remarkable until the MPO truncation $\varepsilon = 10^{-14}$ kicks in after $\ln \mathcal{C}$ drops to approximately -25.

at infinite temperature) is observed in [256] and explained by the fact that at the wavefront the operator $O_1(0,t)$ is less complex, so only a smaller bond dimension is necessary.

A careful error analysis is necessary to extract reliable information from the nonlinear fit to the five parameters $(C, \lambda, x_0, v_B, p)$, appearing in (2.2). Here C is the prefactor. Three major causes of systematic errors are identified: finite bond dimension χ, a finite time range $[t_0, t_1]$ of data and inaccuracy of the functional form (2.2). The convergence with respect to bond dimensions is verified:

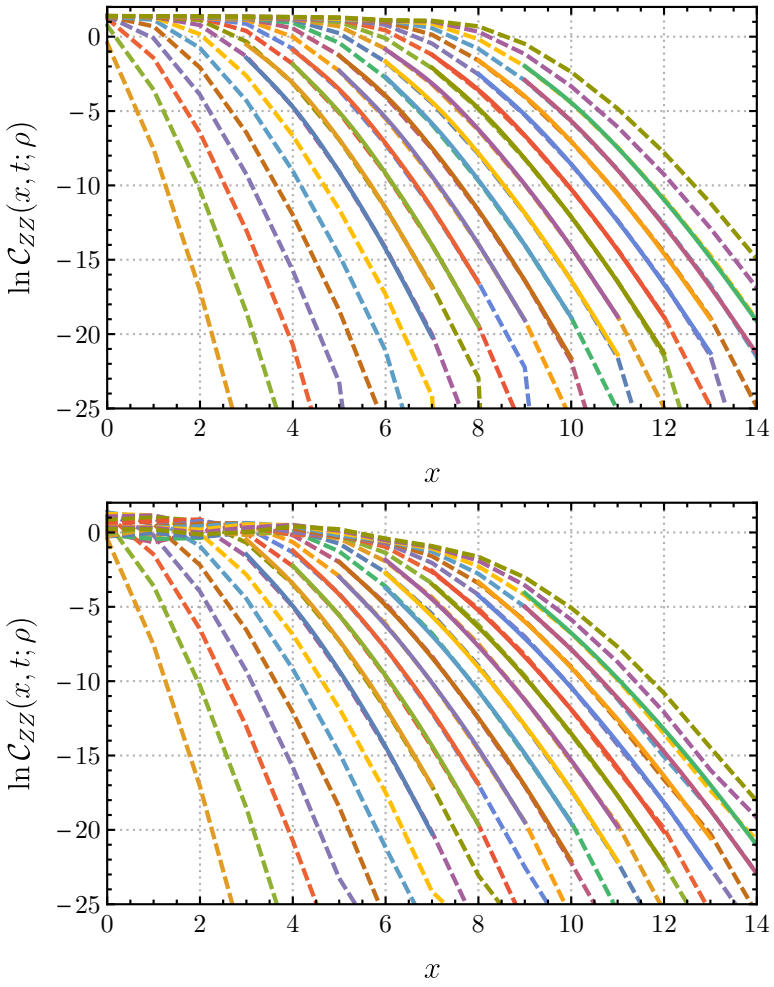


Figure 2.5: Examples of fitting. Dashed curves are from MPO numerics and fitting of (2.2) to wavefront is marked as solid. Each curve is $\ln \mathcal{C}$ for a fixed $Jt = 0.2, 0.4, \ldots, 4.8$. The first plot is for $\beta = 0$ and $h_X = 1.05J$, $h_Z = 0$ with a fitting $v_B = 1.95Ja$, $p = 0.46$ to be compared with exact values $v_B = 2Ja$ and $p = 0.5$ ($a = 1$ is the lattice spacing); the second plot is for $\beta J = 3$, $h_X = 1.05J$, $h_Z = 0.3J$ and the best fitting is $v_B = 1.39Ja$ with $p = 0.65$.

for all data used the difference in $\ln \mathcal{C}$ between $\chi = 256$ and $\chi = 512$ is less than 0.05 and our main results do not depend on such a small difference. Also the fitting as presented in Fig. 2.5 is visually reasonably good, even for the chaotic Hamiltonian $h_Z = 0.3J$ at low temperature $\beta J = 3$.

The effect of a finite range of data and inaccuracy of the functional form is quantitatively manifested as dependence on the hyperparameters δ and t_0. Since the butterfly velocity is defined in the late time limit, t_0 should not be too small; but because only data up to time t_1 are available, t_0 cannot be arbitrarily large either. Moreover, larger t_0 means less data and more significant numerical instability. In Fig. 2.6, dependence on δ and t_0 of the fitted butterfly velocity for $\beta J = 3$

and $h_Z = 0.4J$ is shown. We will work with the values $\delta = 1.0$, $Jt_0 = 1.5$ and $Jt_1 = 4.4$.

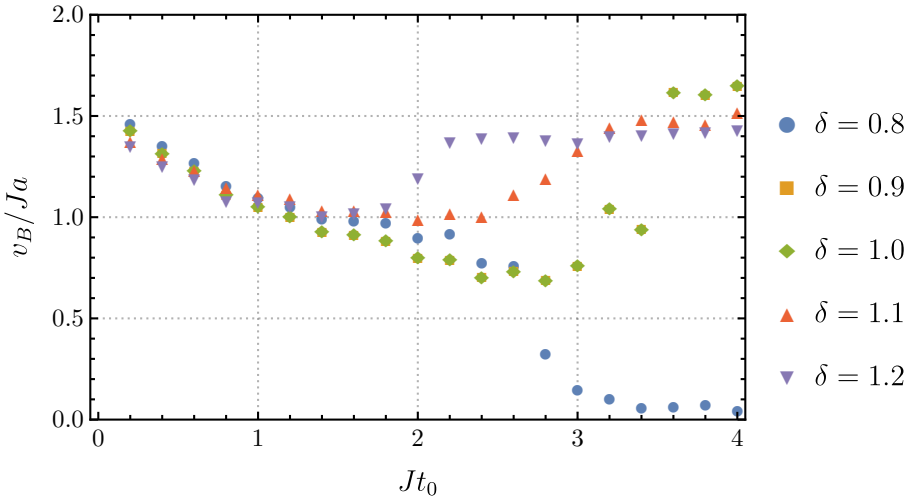


Figure 2.6: Fitted butterfly velocity at $h_X = 1.05J$, $h_Z = 0.4J$ and $\beta J = 3$ for different hyperparameters δ and t_0 ($Jt_1 = 4.4$ and $a = 1$). For small t_0, fluctuation with respect to δ is insignificant due to a larger amount of data. However, at these early times there is a systematic error leading to a dependence on t_0. When $Jt_0 > 2$ the fitting is not stable. The optimal choice of hyperparameters, from the figure, would be $Jt_0 \approx 1.5$ with $\delta \approx 1.0$.

With this choice of hyperparameters, we produce the figures in the main text. Errors are estimated via slightly tuning hyperparameters. Details are summarized in Fig. 2.7, with fitted values of p and λ given as well. From the plot errors are estimated to be within a scale of 0.05, 0.05 and 0.5 for $v_B(\beta)/v_B(0)$, p and λ/J respectively.

The correlation length ξ is extracted with MPO numerics as well, as the inverse spatial decay rate of connected two-point correlations $\mathrm{tr}(\rho Z_{15} Z_{15+x}) - \mathrm{tr}(\rho Z_{15})\mathrm{tr}(\rho Z_{15+x})$ in an $N = 50$ chain with operator insertions at sites 15 and $15 + x$, where $x = 0, 1, \ldots, 20$. The exponential fit is remarkably good with correlation lengths at different temperatures and longitudinal fields shown in Fig. 2.8. Given the correlation length ξ along with p and λ from Fig. 2.7, the bound is evaluated (with error estimates) in Fig. 2.9. In evaluating the inequality (2.23) we have used $v_S \leq v$ for $v = 3Ja$ and $\xi_{\mathrm{LR}} = a$ (cf. section 2.8), where a Lieb-Robinson inequality with $(v, \xi_{\mathrm{LR}}) = (3Ja, a)$ is verified in numerics and a is the lattice spacing.

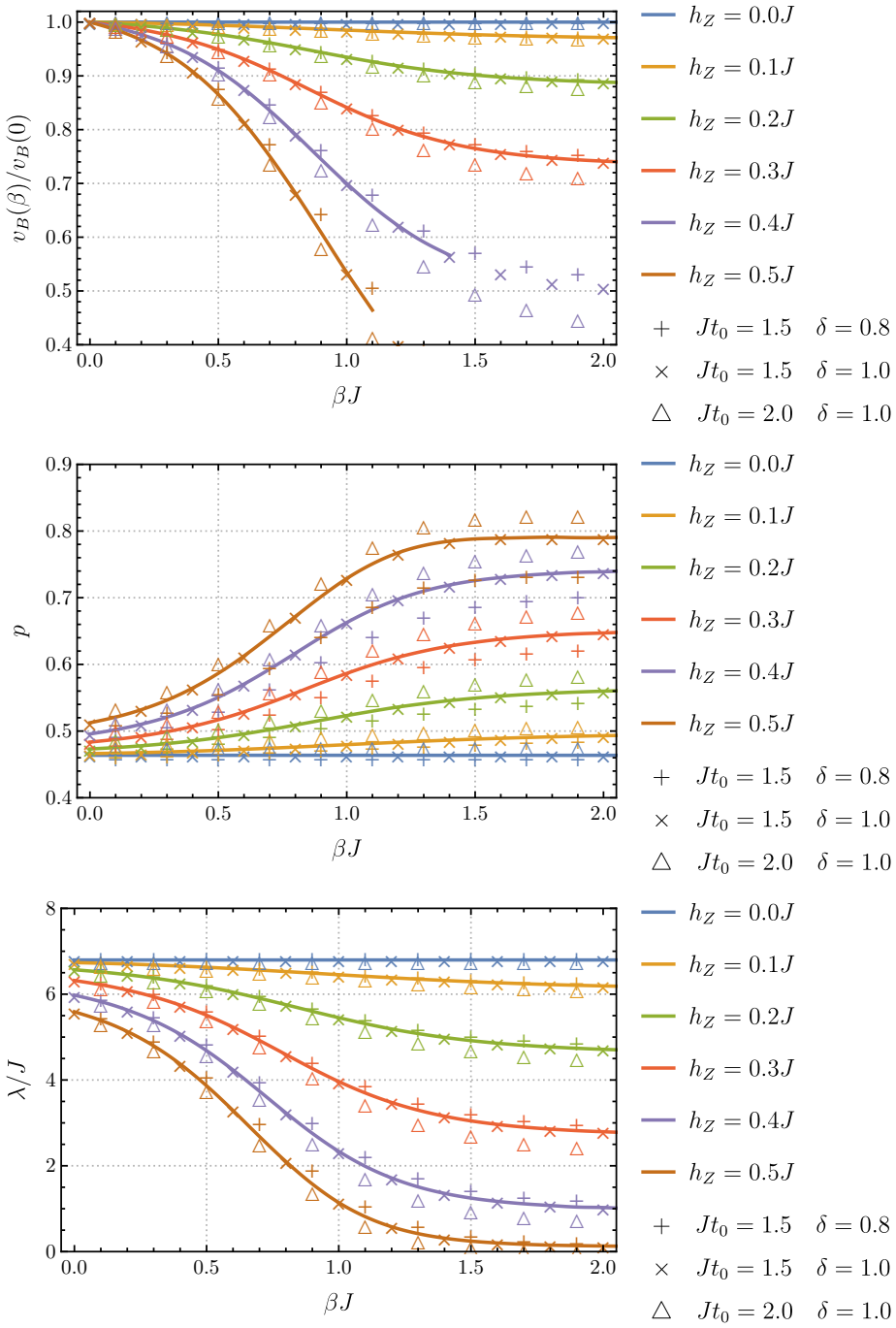


Figure 2.7: Scrambling characteristics in (2.2) fitted for numerics in mixed field Ising chain (2.21) with $h_X = 1.05J$, different longitudinal field h_Z, inverse temperature β and hyperparameters t_0 and δ (with $Jt_1 = 4.4$). Solid curves are guides to the eye of fits at $Jt_0 = 1.5$ and $\delta = 1.0$.

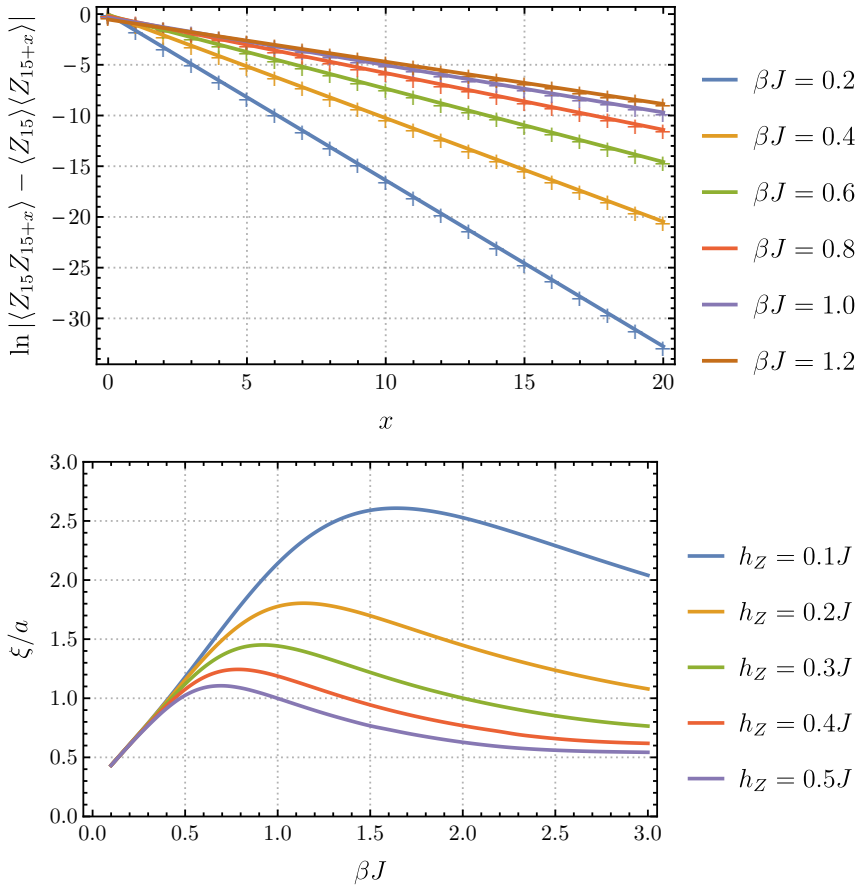


Figure 2.8: Lower plot: Correlation length ξ for different inverse temperatures β and longitudinal fields h_Z. $N = 50$, $h_X = 1.05J$ in (2.21) and a is the lattice spacing. Upper plot: As an example, details of fitting at $h_Z = 0.1J$. $+$ are numerical data and lines are linear fitting.

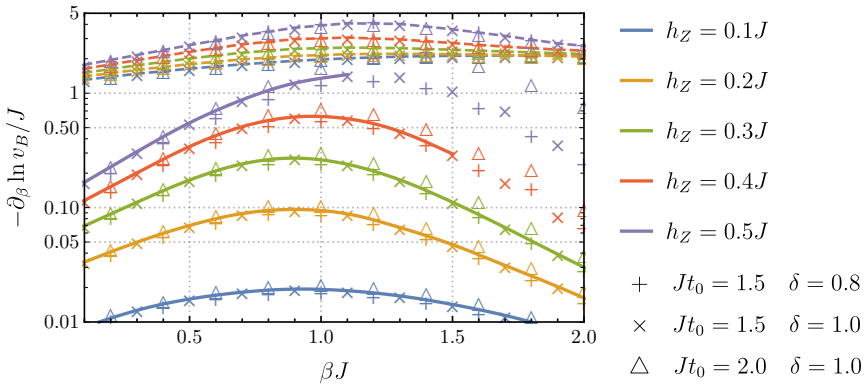


Figure 2.9: Temperature dependence of the butterfly velocity for different longitudinal fields h_Z and hyperparameters t_0 and δ with $h_X = 1.05J$. Upper bounds are evaluated according to (2.23) shown as the dashed lines in the top of the figure.

Chapter 3

Quantum Many-body Bootstrap

Abstract

A numerical bootstrap method is proposed to provide rigorous and nontrivial bounds in general quantum many-body systems with locality. In particular, lower bounds on ground state energies of local lattice systems are obtained by imposing positivity constraints on certain operator expectation values. Complemented with variational upper bounds, ground state observables are constrained to be within a narrow range. The method is demonstrated with the Hubbard model in one and two dimensions, and bounds on ground state double occupancy and magnetization are discussed.

3.1 Introduction

Understanding ground states of interacting many-body systems remains a central challenge in quantum physics. The general problem is intrinsically difficult [227] and advances are often made with the aid of symmetries, approximations and numerics. Conformal symmetry and positivity have proved to be powerful in constraining correlators of quantum fields, via the conformal bootstrap [201]. In this work the positivity constraints are applied to lattice systems without conformal invariance.

The bootstrap approach in this work is algebraic in nature, and relies only on quantum mechanical first principles. As such it is capable of addressing ground state questions in systems with unbounded local Hilbert spaces, or with fermion sign problems. For example, similar methods for solving many-body quantum mechanics with large-N matrix degrees of freedom are proposed in [149, 89]. Also no approximation or assumption about the states is necessary, and thus the results are rigorous and serve as tests for other approximate algorithms. Generality and rigor of the method are favorable in cases where approximate methods give inconsistent results [139].

The bootstrap algorithm is also a generalization of the established variational reduced density matrix theory [212, 166] to infinite lattices. In that method, the energy is minimized while the

positivity constraints are imposed for few-body reduced density matrices, yielding lower bounds for ground state energies. Previous works (e.g., [18, 245, 167, 221]) mostly deal with all two-body reduced density matrices, and hence the computational complexity is polynomial in system size. To better utilize geometric locality of the problem, I instead consider spatially local operators only. This allows me to systematically probe more-body operators and bootstrap directly in the thermodynamic limit.

In this work some ground state observables in the Hubbard model [112] are bounded, as a proof-of-principle demonstration of the method. In one dimension (see Table 3.1), exact solutions are available for comparison [147]. Significant numerical progress has been made in two-dimensional cases [139, 261, 111]. Lower bounds on ground state energies are obtained by bootstrap and are within a few percent of the state-of-the-art results (see Table 3.2). It should be interesting to compare the current algorithm with the Anderson bounds [8, 242, 241].

The lower bounds are complementary to the varitional upper bounds given by existing numerical approaches [139]. Often the ground state energy and observables are then pinned down in a narrow range. Such rigorous constraints on ground state observables are not generally accessible to variational methods. As an example, nontrivial bounds on double occupancy and antiferromagnetic ordering in the two-dimensional Hubbard model ground states are obtained in Table 3.3 and 3.4.

3.2 Method

The many-body bootstrap is based on symmetry and unitarity in quantum mechanics. Specifically, denote $\langle O\rangle = \mathrm{tr}(\rho O)$, where ρ is some density matrix, then for any operator O,

$$\langle I\rangle = 1, \quad \langle O^\dagger\rangle = \langle O\rangle^*, \quad \langle O^\dagger O\rangle \geq 0. \tag{3.1}$$

Furthermore, $\langle U^{-1}OU\rangle = \langle O\rangle$ if U is a symmetry of the state ρ, i.e., $U\rho U^{-1} = \rho$. If the symmetry is generated by a conserved charge C, also $\langle [C, O]\rangle = 0$. Thermal states and energy eigenstates are time translation invariant, so $\langle [H, O]\rangle = 0$ with H the Hamiltonian and O an arbitrary operator.

Lower bounds on ground state energies are obtained by minimizing $\langle H\rangle$ subject to the constraints (3.1). More precisely, the minimization is done over the following set $\mathcal{A}$ of linear functionals $\mathcal{F}$ of operators:

$$\begin{aligned}\mathcal{A} = \{\mathcal{F} : \mathcal{F}[I] = 1, \quad & \mathcal{F}[O^\dagger] = \mathcal{F}[O]^*, \\ & \mathcal{F}[[C_\alpha, O]] = 0, \quad \mathcal{F}[U_\alpha^{-1} O U_\alpha] = \mathcal{F}[O], \quad \forall O \in \mathcal{C}_1, \\ & \mathcal{F}[\widetilde{O}^\dagger \widetilde{O}] \geq 0, \quad \forall \widetilde{O} \in \mathcal{C}_2\}.\end{aligned} \tag{3.2}$$

Minimization over this subset of functionals is equivalent to searching for operator expectation values $\langle O\rangle = \mathcal{F}[O]$ under the constraints (3.1). Here C_α and U_α are generators of the continuous

and discrete symmetries to be imposed on the state. In practice the constraints (3.1) can only be imposed for a subset of operators $\mathcal{C}_1$ and $\mathcal{C}_2$. Choice of $\mathcal{C}_1$ and $\mathcal{C}_2$ affects computational efficiency of the algorithm, and an empirical choice in fermionic lattice models will be discussed shortly.

The true ground state energy E_0 is bounded below by the minimal value from $\mathcal{F} \in \mathcal{A}$:

$$E_0 \geq \min_{\mathcal{F}\in\mathcal{A}} \mathcal{F}[H] =: E_{\text{lb}}, \tag{3.3}$$

because the functional $\mathcal{F}[O] = \text{tr}(\rho_0 O)$ is always in $\mathcal{A}$ for a ground state ρ_0 of H that also commutes with all the charges C_α and U_α. The minimization in (3.3) can be solved efficiently and accurately by semidefinite programming (e.g., with [185, 186]).

The equality in (3.3) is reached when $\mathcal{C}_1$ and $\mathcal{C}_2$ are the full set of operators. Hence it is expected that the lower bound (3.3) becomes tight as the number of constraints is increased. Indeed, any linear functional $\mathcal{F}$ can be written as $\mathcal{F}[O] = \text{tr}(FO)$ for some operator F. And F is a density matrix (positive with unit trace) if and only if (3.1) holds for any O. Thus by the variational principle $\mathcal{F}[H] = \text{tr}(FH)$ is minimized precisely when F is a ground state, and $E_0 = \min \mathcal{F}[H]$.

The bootstrap lower bound on ground state energy is complementary to the conventional variational upper bounds. Knowing that $E_{\text{lb}} \leq E_0 \leq E_{\text{ub}}$, the ground state expectation values can be bounded as

$$\begin{aligned} \text{tr}(\rho_0 O) &\geq \min_{\mathcal{F}\in\mathcal{A}, E_{\text{lb}}\leq\mathcal{F}[H]\leq E_{\text{ub}}} \mathcal{F}[O], \\ \text{tr}(\rho_0 O) &\leq \max_{\mathcal{F}\in\mathcal{A}, E_{\text{lb}}\leq\mathcal{F}[H]\leq E_{\text{ub}}} \mathcal{F}[O]. \end{aligned} \tag{3.4}$$

The inequalities (3.4) can be restrictive when E_{lb} and E_{ub} are close (e.g., see Table 3.2 and 3.3).

The method is illustrated with the Hubbard model in one and two dimensions:

$$H = -\sum_{\langle xy\rangle\sigma} c^\dagger_{x\sigma} c_{y\sigma} + U \sum_x n_{x\uparrow} n_{x\downarrow}, \tag{3.5}$$

where $\langle xy\rangle$ runs over ordered pairs of nearest-neighbor lattice sites, and $c_{x\sigma}$ is the fermion annihilation operator on site x with spin $\sigma = \uparrow, \downarrow$. For simplicity I consider square lattices with unit spacing. The bootstrap works directly in the thermodynamic limit.

The Hamiltonian (3.5) has discrete lattice translation and rotation symmetries, along with a U(2) global symmetry generated by

$$N = \sum_x (n_{x\uparrow} + n_{x\downarrow}), \quad S_\alpha = \frac{1}{2} \sum_{x\sigma\sigma'} c^\dagger_{x\sigma} (\sigma_\alpha)_{\sigma\sigma'} c_{x\sigma'}, \tag{3.6}$$

where $\alpha = x, y, z$ and σ_α are Pauli matrices. The fermion number N, total spin-z component S_z, lattice translation and rotation will serve as C_α and U_α in (3.2) for bootstrapping.

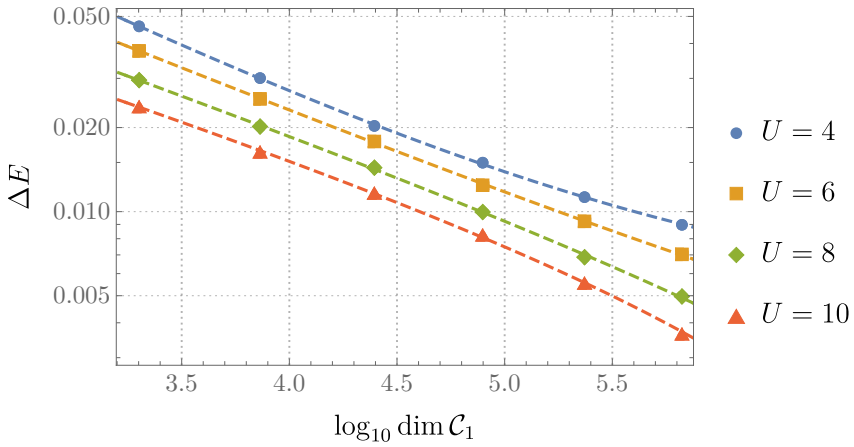


Figure 3.1: The difference $\Delta E = E_0 - E_{\text{lb}}$ as a function of $\dim \mathcal{C}_1$, the number of operators in (3.2), for the one-dimensional Hubbard model (3.5) at half filling. Dashed curves show the best fits of form $E_{\text{lb}} = A + B(\dim \mathcal{C}_1)^{-\alpha}$.

As mentioned previously, the choice of $\mathcal{C}_1$ and $\mathcal{C}_2$ in (3.2) affects performance of the algorithm. In fermionic lattice models with a local Hamiltonian, such as (3.5), it is plausible that local operators are more important. Dimensions of the subspaces $\mathcal{C}_1$ and $\mathcal{C}_2$ are controlled by a positive integer K, bounding the degree of locality of operators. The spaces are enlarged when K is increased.

To be more precise, two types of locality are present in (3.5): k-locality (H is a sum of few-body operators) and geometric locality (the interactions are short-ranged). For a string of fermion creation and annihilation operators

$$O = c^{(\dagger)}_{x_1\sigma_1} c^{(\dagger)}_{x_2\sigma_2} \cdots c^{(\dagger)}_{x_r\sigma_r}, \tag{3.7}$$

define a locality measure (with respect to a site chosen as the origin)

$$l(O) = r + \sum_{i=1}^{r} \|x_i\|. \tag{3.8}$$

The first term r is the number of fermion operators in (3.7), counting the degree of k-locality. The second term is a sum of geometric l_1-norms of the lattice vectors x_i. For any positive integer K, I choose $\mathcal{C}_2$ to be linearly spanned by fermion strings (3.7) with $l(O) \leq K$, and $\mathcal{C}_1$ spanned by the strings that appear in the products of two operators in $\mathcal{C}_2$. An ordering of fermion creation and annihilation operators is also employed and only normal ordered strings are considered to avoid unnecessary duplication.

$n = 1$	$U = 4$	$U = 6$	$U = 8$	$U = 10$
$E_{\rm lb}\vert_{K=10}$	-0.5827	-0.4271	-0.3325	-0.2708
$E_{\rm lb}\vert_{K=\infty}$	$-0.5781(7)$	$-0.4212(9)$	$-0.3260(11)$	$-0.2648(14)$
E_0	-0.5737	-0.4201	-0.3275	-0.2672
$\mathcal{F}[D]\vert_{K=10}$	0.1013	0.0592	0.0373	0.0252
$\mathcal{F}[D]\vert_{K=\infty}$	0.1015(4)	0.0588(7)	0.0371(4)	0.0248(3)
$\langle D\rangle_0$	0.1002	0.0582	0.0366	0.0248

Table 3.1: Bootstrap lower bounds $E_{\rm lb}$ of one-dimensional Hubbard model ground state energies (per site), and the double occupancy D in $\mathcal{F}$ that minimizes (3.3). Exact values E_0 and $\langle D\rangle_0$ are shown for comparison. The number of fermions per site $n = 1$. For values extrapolated to $K = \infty$, standard errors in fitting are shown in the brackets.

$n = 1$	$U = 2$	$U = 4$	$U = 6$	$U = 8$
$E_{\rm lb}\vert_{K=7}$	-1.221	-0.913	-0.705	-0.565
$E_{\rm lb}\vert_{K=\infty}$	–	–	$-0.66(2)$	$-0.54(2)$
$E_{\rm AFQMC}$	$-1.1763(2)$	$-0.8603(2)$	$-0.6568(3)$	$-0.5247(2)$
$E_{\rm DMET}$	$-1.1764(3)$	$-0.8604(3)$	$-0.6562(5)$	$-0.5234(10)$
$E_{\rm DMRG}$	$-1.176(1)$	$-0.8605(5)$	$-0.6565(1)$	$-0.5241(1)$
$n = 0.875$	$U = 2$	$U = 4$	$U = 6$	$U = 8$
$E_{\rm lb}\vert_{K=7}$	-1.316	-1.103	-0.963	-0.867
$E_{\rm lb}\vert_{K=\infty}$	–	–	$-0.86(5)$	$-0.77(3)$
$E_{\rm DMET}$	$-1.2721(6)$	$-1.031(3)$	$-0.863(13)$	$-0.749(7)$

Table 3.2: Bootstrap lower bounds $E_{\rm lb}$ of two-dimensional Hubbard model ground state energies (per site) E_0, at fillings $n = 1$ and $n = 0.875$. Solutions from AFQMC, DMET and DMRG are shown for comparison.

3.3 Result in Hubbard model

3.3.1 One dimension

Symmetries imposed in (3.2) include $C_\alpha = \{H, N, S_z\}$ from (3.5) and (3.6), and $U_\alpha = \{T, \Pi\}$. Here T is the lattice translation and Π the lattice reflection. For $5 \le K \le 10$, $E_{\rm lb}$ in (3.3) is evaluated and lower bounds the ground state energy. The best bound from $K = 10$ is shown in Table 3.1. Other expectation values are also available, for the functional $\mathcal{F}$ that minimizes (3.3). For example, $D = n_{x\uparrow}n_{x\downarrow}$ in Table 3.1 is the double occupancy. Note that $\mathcal{F}[D]$ does not necessarily bound the ground state value $\langle D\rangle_0 = \mathrm{tr}(\rho_0 D)$.

Extrapolation to $K = \infty$ is also possible. In Figure 3.1 expectation values at finite K fit well to the functional form $A + B(\dim \mathcal{C}_1)^{-\alpha}$, where $\dim \mathcal{C}_1$ is the number of operators in the constraints (3.2). The fitted $\alpha \approx 0.3$, consistent with that the algorithmic complexity is polynomial in the required accuracy. Standard errors from the fitting are included in Table 3.1. The extrapolated values agree with the exact solution.

$n=1$	$U=2$	$U=4$	$U=6$	$U=8$
$d_{\rm lb}\vert_{K=7}$	0.160	0.106	0.071	0.049
$d_{\rm lb}\vert_{K=\infty}$	0.161(6)	0.108(7)	0.072(5)	0.050(3)
$d_{\rm ub}\vert_{K=7}$	0.224	0.169	0.117	0.079
$d_{\rm ub}\vert_{K=\infty}$	0.195(14)	–	–	–
$d_{\rm DMET}$	0.1913(4)	0.1261(1)	0.08095(4)	0.05398(7)
$d_{\rm DMRG}$	0.188(1)	0.126(1)	0.0809(3)	0.0539(1)

Table 3.3: Bootstrap bounds $d_{\rm lb} \le {\rm tr}(\rho_0 D) \le d_{\rm ub}$ of ground state double occupancy (per site) $D = n_{x\uparrow} n_{x\downarrow}$, for the two-dimensional Hubbard model at half filling.

$n=1$	$U=2$	$U=4$	$U=6$	$U=8$
$m_{\rm ub}\vert_{K=7}$	0.194	0.292	0.352	0.383
$m_{\rm ub}\vert_{K=\infty}$	–	–	–	0.34(2)
$m_{\rm DMET}$	0.133(5)	0.252(9)	0.299(12)	0.318(13)

Table 3.4: Bootstrap upper bounds $m_{\rm ub}$ of ground state staggered magnetization (3.9) per site, at half filling.

3.3.2 Two dimensions

Symmetries are $C_\alpha = \{H, N, S_z\}$ along with $U_\alpha = \{T_{(1,0)}, T_{(0,1)}, \Pi, R\}$, where $T_{(1,0)}$ and $T_{(0,1)}$ are the lattice translations, Π the reflection, and R the $\pi/2$ lattice rotation. No exact solution is known for general couplings, and I will compare with the AFQMC [259], DMET [133, 260] and DMRG [251] results reviewed in [139]. Here the AFQMC solution is numerically exact at half filling without sign problems. The DMRG is a variational technique and the DMET is not variational.

The bounds from $K = 7$, along with the values extrapolated to $K = \infty$ from $4 \le K \le 7$, are obtained in Table 3.2, 3.3 and 3.4. Estimated standard errors are shown in the brackets. Some values are omitted due to deficient K and thus poor fitting quality in extrapolation. While the bounds are rigorous for any finite K, uncontrolled errors are introduced in extrapolation. The extrapolation may be further improved with more computational resources.

For smaller U in Table 3.2, the best bounds available are within a few percent of the variational energies, corroborating the effectiveness of both methods. At larger U, when extrapolation is more reliable, the extrapolated energies agree with other numerics within numerical uncertainties.

If the ground state energies in [139] are upper bounds, local observables are constrained by (3.4). For instance, in the following the DMRG energies at $n = 1$ from [139] are used as $E_{\rm ub}$. Bounds for double occupancy D are shown in Table 3.3, which are restrictive and consistent with other numerics.

As another example, consider the staggered magnetization

$$M = \frac{1}{2} \sum_x (-1)^{x_1+x_2} (n_{x\uparrow} - n_{x\downarrow}), \tag{3.9}$$

where (x_1, x_2) are coordinates of x. Discrete symmetries are reduced to $U_\alpha = \{T_{(1,1)}, T_{(1,-1)}, \Pi, R\}$, to allow for nonzero M. Upper bounds on M per site are obtained in Table 3.4. At large U the bound is also consistent with the Heisenberg limit $m \approx 0.307$ [224]. For magnetization the two inequalities in (3.4) are not independent, as $\min \mathcal{F}[M] = -\max \mathcal{F}[M]$.

3.4 Conclusion

I have shown that the idea of positivity, which is fundamental in many successful theories, can be employed to solve local lattice models. The bounds are nontrivial checks on other numerics and expand our knowledge of interacting quantum many-body systems.

It would be ideal to have a nonzero lower bound on ground state ordering as well. This is difficult in the current formalism as ground states that do not break symmetries are not ruled out by the constraints. Possibly one should consider two-point functions, by re-introducing non-local few-body operators of interest.

Other directions include generalizing the method to continuous theories, or imposing more constraints on the state (for example, that the state is thermal or a condensate). Also bootstrap bounds on spectral functions, as well as inhomogeneous phases may be useful in constraining low-energy excitations and competing orders in strongly correlated electron systems.

Acknowledgements

This work arose from discussions with Sean Hartnoll. The author acknowledges discussions with Sean Hartnoll, Jorrit Kruthoff, Edward Mazenc and Daniel Ranard on related projects.

Part II

Emergent Locality

Chapter 4

Quantum Causal Influence

This chapter is essentially the same as

- Cotler, Jordan, Xizhi Han, Xiao-Liang Qi, and Zhao Yang. "Quantum causal influence." *Journal of High Energy Physics* 2019.7 (2019): 1-67.

Abstract

We introduce a framework to study the emergence of time and causal structure in quantum many-body systems. In doing so, we consider quantum states which encode spacetime dynamics, and develop information theoretic tools to extract the causal relationships between putative spacetime subsystems. Our analysis reveals a quantum generalization of the thermodynamic arrow of time and begins to explore the roles of entanglement, scrambling and quantum error correction in the emergence of spacetime. For instance, exotic causal relationships can arise due to dynamically induced quantum error correction in spacetime: there can exist a spatial region in the past which does not causally influence any small spatial regions in the future, but yet it causally influences the union of several small spatial regions in the future. We provide examples of quantum causal influence in Hamiltonian evolution, quantum error correction codes, quantum teleportation, holographic tensor networks, the final state projection model of black holes, and many other systems. We find that the quantum causal influence provides a unifying perspective on spacetime correlations in these seemingly distinct settings. In addition, we prove a variety of general structural results and discuss the relation of quantum causal influence to spacetime quantum entropies.

4.1 Introduction

Causal structure is an essential property of spacetime geometry. In relativistic classical mechanics, the causal structure is determined by the behavior of null geodesics. The future light cone of a

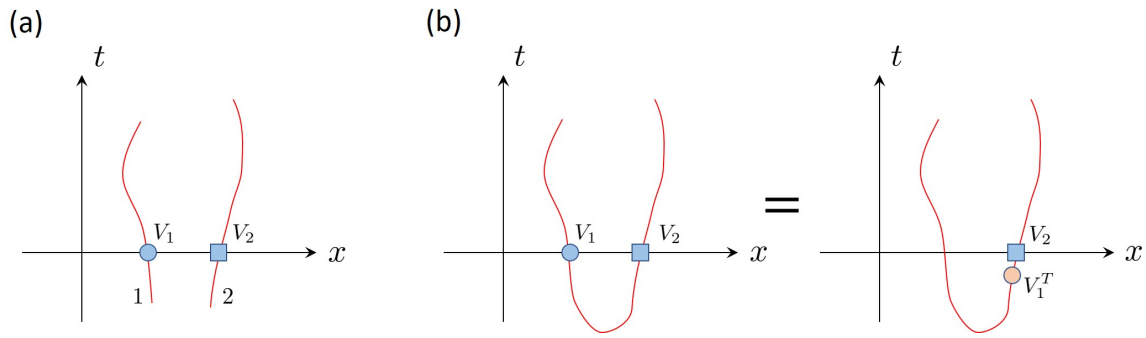


Figure 4.1: (a) The world lines of two spin-$\frac{1}{2}$ particles $1, 2$ in spacetime (red curves). Two operators V_1 and V_2 probe the spins of the two particles at time $t = 0$. (b) When the initial state of the spins of the two particles forms an EPR pair, the effect of $V_1 \otimes V_2$ on particles 1 and 2 is equivalent to applying $V_2 V_1^T$ to particle 2 alone.

point x comprises all of the points that may be influenced by an arbitrary perturbation at x. In relativistic quantum field theory, we usually treat the causal structure as classical, with well-defined light cones. In more general quantum many-body systems which may be non-relativistic or do not posses quasiparticles resembling massless excitations, there is still a generalization of the causal structure so long as there is an upper-bound on the speed of information propagation. For example, for lattice models with a local Hamiltonian, the Lieb-Robinson bound [146] gives a velocity v_{LR} which defines an analog of the speed of light. In particular, a local perturbation can only influence the region inside its future Lieb-Robinson cone.

However, beyond these familiar cases, the causal structure in quantum mechanics can be much richer. As a simple example, consider two spin-$\frac{1}{2}$ particles $1, 2$ in Figure 4.1 above. At time $t = 0$, particles 1 and 2 are at location x_1 and x_2. On a fixed time slice $t = 0$, suppose we probe the spin degrees of freedom of particles 1 and 2 with separate Hermitian operators V_1 and V_2, respectively. These two probe events are clearly spacelike separated. Now if we prepare the spin degrees of freedom of particles 1 and 2 in an EPR pair state $\frac{1}{\sqrt{2}} (|\uparrow\rangle_1 |\uparrow\rangle_2 + |\downarrow\rangle_1 |\downarrow\rangle_2)$ at an earlier time $t_i < 0$, applying V_1 and V_2 to particles 1 and 2 at time $t = 0$ is equivalent to applying $V_2 V_1^T$ only to particle 2 at $t = 0$. (V_1^T is an operator defined by the matrix transpose of V_1 in the S_z basis.) Therefore, for our particular initial state of the spin degrees of freedom, it becomes ambiguous whether the two probe events are spacelike or time-like separated.

Following the spirit of Einstein's theory of relativity, one would like an *observable* way to define the causal relation between events in a quantum many-body system, which is uniquely determined by physical correlation functions and has an unambiguous interpretation. This is the goal of the current paper. We propose a measure of quantum causal influence that determines whether a spacetime region A has nontrivial influence on another spacetime region B. The measure reproduces the ordinary causal structure for the familiar case of relativistic classical systems, but also unveils

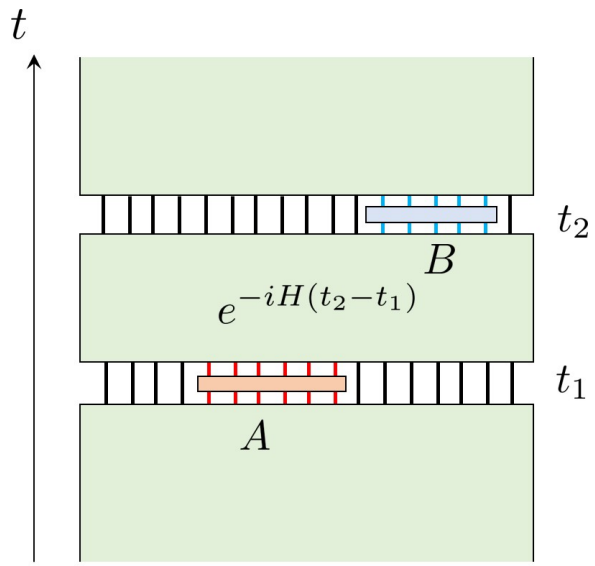


Figure 4.2: Depiction of regions A at time t_1 and B at time t_2 for a spin chain. The causal influence is measured by inserting a unitary operator U_A in region A (orange box) and studying its effect on the measurement of an arbitrary operator O_B in B (blue box).

various unconventional causal structures that are unique to quantum mechanics.

Our emphasis on correlation functions and many-body states differs from previous work on causality in quantum mechanics which emphasize few-body systems and causal inference on data from decoherent measurements [190, 42, 1, 76, 213, 199, 43, 54, 191, 214, 4, 159, 47]. We are primarily interested in the emergence of causal structure in quantum systems with many degrees of freedom, and the flow of time experienced by observers inside the systems. For related work in this direction using the quantum process tensor formalism and related formalisms, see [188, 189, 91, 92, 122, 124, 123].

To illustrate the idea of our proposal, let us first consider time evolution with a local Hamiltonian. For concreteness, we can consider a $(1+1)$-dimensional model of N spins labeled by $x = 1, 2, ..., N$ with a Hamiltonian that couples neighboring spins. We will refer to this system as the "main system." Starting from an initial state $|\psi_i\rangle$ at time $t = 0$, the main system evolves as $|\psi(t)\rangle = e^{-iHt}|\psi_i\rangle$ in the Schrödinger picture. Consider two spatial regions, A at time t_1 and B at time t_2, as shown in Figure 4.2. Now suppose there is an experimentalist who can only access the the two spacetime regions A and B, but can otherwise perform arbitrary operations. In particular, the experimentalist is a superobserver who can couple her external apparatus to region A by performing a joint unitary on A and her apparatus at time t_1, and similarly for B at time t_2. We also assume that the experimentalist has the ability to reset the whole system to the initial state $|\psi_i\rangle$ and run the experiment an unlimited

number of times. Now the question is, how can the experimentalist determine whether physical operations in A causally influence the region B? Naturally, the experimentalist can run different experiments with different perturbations on region A (by coupling her external apparatus to A in different ways) and measure some physical quantity at region B. If the result of the measurement at B depends nontrivially on the pertubation at A, we conclude that A causally influences region B.

However, it is important to distinguish causal influence and correlation. Even if A and B are spacelike separated, operators in A and B can certainly have a nontrivial connected correlation function. Measures of connected correlation (such as the quantum mutual information between the two regions, if they are spacelike separated) are symmetric between the two regions, and thus do not probe the causal structure. For instance, it may be the case that A causally influences B but B does not causally influence A, and so causal influence is necessarily an asymmetric relation. It turns out that a simple modification of the setup can distinguish causal influence from other kinds of correlation. The experimentalist can apply a unitary gate U_A to region A, which changes the state of the system but does not introduce entanglement with her apparatus. Then, the experimentalist can couple her apparatus to region B in the ordinary way, which generically entangles the main system with the apparatus. If B has no overlap with the future light cone (or for a lattice model, the Lieb-Robinson cone) of A, the unitary operator U_A does not change the reduced density matrix of B and therefore does not change any physical property there.

The procedure described above may sound a bit trivial since it is exactly how we do response theory in many-body systems. If we consider an infinitesimal unitary $U_A = \exp(-i\,\epsilon J_A)$, and measure an operator J_B at B, the linear response function is determined by the commutator $-i\,[J_A(t_1), J_B(t_2)]\,\theta(t_2 - t_1)$, which vanishes outside the light cone. However, the commutator expression depends on the Heisenberg picture, which relies on picking a choice of time slicing (i.e., Cauchy surfaces). Since we want a measure of the causal structure that is not predicated on pre-defined time slices, it is more natural to work with tensor networks, which are not endowed with a pre-defined causal structure. Indeed, our proposal allows us to study causal structure in systems with no obvious time slicings. For example, in a hyperbolic "perfect tensor network" [195], there are isometry relations between operators acting on different subsets of links, but there is no light cone or preferred time-like direction. Our proposal allows us to start from scratch and probe causal influence between different degrees of freedom in the system, without any *a priori* knowledge of a time direction. In particular, there is no need to distinguish whether some qubits (or more generally, degrees of freedom) in A and B are "the same qubits evolved in time" or "independent qubits that are entangled."

The remainder of the paper is organized as follows: We start by presenting the general setup. For concreteness, we use the language of tensor networks to describe a general quantum system, *without* needing to designate how degrees of freedom sit in a putative spacetime. This is a very

useful framework for "spacetime agnostic" descriptions of quantum systems. Even if we have a continuum of degrees of freedom, as long as we assume that accessible regions A and B comprise of discrete spacetime points, the system can be described by a tensor network. We show how a general quantum system can be considered as a tensor network with insertions of operators in links, and with a given boundary condition. For example, in the more familiar setting of a quantum system with unitary time evolution, the boundary conditions of the tensor network correspond to an initial density operator (i.e., an initial state) and optionally a final density operator (i.e., a final state). Ordinary quantum mechanics without a final state density operator is equivalent to having a *maximally mixed* final state density operator. We will discuss this in detail later.

Next, we provide the definition of quantum causal influence in the general setup. With this probe of quantum causal structure at hand, we investigate various examples and identify some key features of causal influence that are unique to quantum systems. One feature is that the causal structure generically depends on the initial state, or more generally the boundary conditions of the tensor network. In the familiar case of a quantum system with unitary time evolution, the direction of the "future" is determined by the fact that the final state is maximally mixed but the initial state is not. If the initial state contains a region with a maximally mixed reduced density operator, the future light cone of points in the domain of dependence of that region will be "erased." Another example of causal structure which is sensitive to the initial state is quantum teleportation. We show how quantum teleportation corresponds to "erasing" part of the future light cone of the teleportee due to a special initial state containing EPR pairs.

The other unique feature of quantum causal influence is that it is generically nonlocal. In classical mechanics, causal structure is determined by the causal relationships of pairs of points. Classically, a spacetime region B is influenced by a spacetime region A if and only if some points in B are in the future light cone of some points in A. This is not the case for quantum systems. To fully understand the quantum causal structure of a system, it is essential to consider the influence between regions A, B of generic size. In fact, the quantum causal influence between subsystems of A and B do not generically determine the quantum causal influence between A and B themselves. For instance, it is possible to have smaller regions B_1 and B_2 which are not *individually* influenced by A, but for which the union $B_1 \cup B_2$ is influenced by A. Such nonlocal influence is a key feature of quantum erasure codes. The encoding map of a quantum erasure code takes quantum information in a region A and maps it to $B = B_1 \cup B_2$ nonlocally. If the influence of A to each subregion B_1, B_2 is trivial, the code is immune to local errors that occur in only one of B_1 or B_2.

The nonlocality of quantum causal influence provides a new perspective on the exotic causal structures underlying holographic duality. In holographic tensor networks such as perfect tensor networks or large bond dimension random tensor networks [102], all pairs of small regions appear "spacelike separated" since no small region influences any other small region. However, a small region (or more precisely, code subspace operators in a small bulk region) can influence *large* regions

and ultimately influence the boundary in a nonlocal way, as is required by the reconstruction of bulk operators on the boundary. Using quantum causal influence, we find that holographic tensor networks can admit exotic quantum analogs of Cauchy slices comprising of concentric spheres. Another example we study is the final state projection model of black holes [108], which utilizes post-selected quantum mechanics. We discuss how causal influence between small regions does not know about a post-selected random final state, while regions that are large enough have abnormal causal relations and do detect the violation of unitarity by the final state.

After discussing various features and examples of quantum causal influence, we turn to some more quantitative properties. We define a "superdensity operator" [56] of regions A, B which determines all correlation functions involving these two regions. With this tool, we investigate the averaged quantum causal influence by averaging over unitaries in A and generic operators in B. The averaged causal influence is a quantum information theoretic property of the superdensity operator. As two examples, we numerically computed the averaged causal influence in quantum Ising spin chains and stabilizer code models.

We find that quantum causal influence provides a new probe of many-body chaos since the influence between two small regions decays in a chaotic system even if the regions are time-like separated. This is a consequence of operator scrambling and thermalization – a local perturbation becomes non-local and at a later time has little effect on local regions except by contributing to conserved quantities such as energy. We also discuss an upper bound of the causal influence by spacetime quantum mutual information (which is again defined for the superdensity operator) [56]. Finally, we discuss some open questions and future directions.

Below is a brief summary, section by section:

- In Section 4.2, we provide definitions of general tensor networks, graphical tensor networks, and quantum causal influence.
- In Section 4.3, we explore how quantum causal influence depends on boundary conditions. We provide many examples, and prove general, structural results.
- In Section 4.4, we discuss the nonlocality of quantum causal influence in the context of quantum error correction codes, scrambling, and quantum teleportation.
- In Section 4.5, we give examples in the context of quantum gravity, specifically for holographic tensor networks and models of a black hole final state.
- In Section 4.6, we establish the relationship between the averaged quantum causal influence and spacetime quantum entropies and mutual information. We use our results to analyze quantum causal influence in quantum spin chains and stabilizer tensor networks.
- In Section 4.7, we make concluding remarks and discuss future directions.

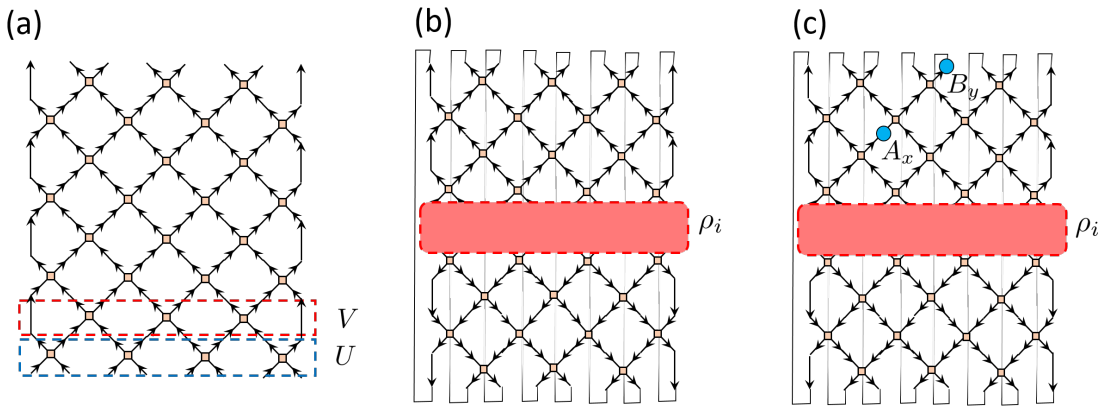


Figure 4.3: (a) An example of a tensor network describing a unitary operator $W = (VU)^M$. Each vertex is a two-qubit unitary gate with the inputs and outputs indicated by arrows pointing toward or away from the vertex, respectively. (b) This is the tensor network obtained by contracting W and $W^\dagger$ with an initial state ρ_i (the red box), and then taking a trace. In other words, the tensor network computes $\mathrm{tr}(W\rho_i W^\dagger) = \mathrm{tr}(\rho_i) = 1$. (c) The tensor network representation of a two-point function defined in Eqn. (4.1).

- In the Appendices, we provide classical and quantum generalizations of causal influence, review the superdensity operator formalism, and also review stabilizer tensor networks.

4.2 General setup

4.2.1 General tensor networks

In order to define characteristics of quantum causal structure, we need to start from a description of a quantum many-body system that does not pick out a time direction. A suitable framework is general tensor networks [247, 246, 144, 56, 209]. Even though popular examples of tensor networks often have a constrained form, the framework of general tensor networks is far broader and encompasses the entire scope of familiar (and unfamiliar) quantum many-body systems.

We start from a simple example of a tensor network, before providing the most general definition. Consider N qubits, where N is even, arranged in a line. First, we apply in parallel two-qubit gates to adjacent qubits via the unitary $U = U_{12} \otimes U_{34} \otimes \cdots \otimes U_{N-1,N}$. Next, we apply another unitary on a different pairing of adjacent qubits, namely $V = V_{23} \otimes V_{45} \otimes \cdots \otimes V_{N-2,N-1}$. Afterwards, we again apply U followed by V, and so on a total of M times, as illustrated in Figure 4.3(a). This procedure yields the unitary operator $W = (VU)^M$. The discrete time evolution implemented by sequential applications of U and V can be considered as a discretization of a continuous time evolution operator e^{-iHt} where H is a local Hamiltonian. Indeed, we can find U and V via a Suzuki-Trotter decomposition of e^{-iHt}.

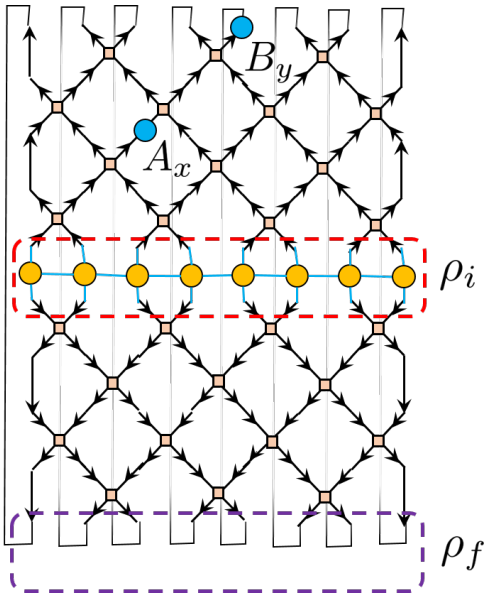


Figure 4.4: This tensor network is a special case of the one in Figure 4.3(c). The network here specifies a particular choice of ρ_i, namely a matrix product operator (MPO), which is depicted within the dashed red lines. We have also put in purple dashed lines to illustrate the fact that taking a trace is equivalent to taking an inner product with a maximally mixed density matrix $\rho_f = \mathbf{1}/d$, up to a normalization d (i.e., the Hilbert space dimension of a spatial slice).

Mathematically, the matrix W is obtained by contracting indices of small matrices $U_{2k-1,2k}$ and $V_{2k,2k+1}$ along all internal links of the network in Figure 4.3(a). We can then contract W and a $W^\dagger$ with some initial state ρ_i, and then take a trace. This yields the tensor network in Figure 4.3(b), which computes $\mathrm{tr}(W\rho_i W^\dagger) = \mathrm{tr}(\rho_i) = 1$. The tensor network is a discrete analog of a partition function, which can be used to compute physical correlation functions. For example, the time-ordered two-point function

$$\langle \mathcal{T}\, B_y(t_2) A_x(t_1)\rangle = \mathrm{tr}\left[B_y (VU)^{t_2-t_1} A_x (VU)^{t_1} \rho_i (U^\dagger V^\dagger)^{t_2}\right] , \tag{4.1}$$

where for concreteness we suppose $t_2 > t_1$, can be computed from the tensor network in Figure 4.3(b) by inserting the operators A_x, B_y into links corresponding to x and y which yields the tensor network Figure 4.3(c). Indeed, the tensor network in Figure 4.3(c) evaluates to the two-point function in Eqn. (4.1) above.

For concreteness, in Figure 4.4 we have chosen an initial density matrix ρ_i which is a matrix product operator (MPO). We will not use MPO's later in the paper, but it suffices to say that the state ρ_i is represented by the partially contracted tensors in the red dashed box in Figure 4.4. The tensor network representation of Figure 4.3(b) also highlights the fact that taking the trace in Eqn. (4.1) is, up to a normalization, equivalent to taking an inner product with another density

matrix $\rho_f = \mathbf{1}/d$, which is the maximal entropy state on a spatial slice (here, we suppose that the Hilbert space dimension of a spatial slice is d). This is just to say that correlation functions, such as the two-point function in Eqn. (4.1), can be written as

$$\mathrm{tr}\left[\rho_f B_y (VU)^{t_2-t_1} A_x (VU)^{t_1} \rho_i (U^\dagger V^\dagger)^{t_2}\right] \propto \mathrm{tr}\left[B_y (VU)^{t_2-t_1} A_x (VU)^{t_1} \rho_i (U^\dagger V^\dagger)^{t_2}\right] \tag{4.2}$$

since ρ_f is proportional to the identity. Although this may seem like a trivial rewriting, we will see later that it is significant.

By making both ρ_i and ρ_f explicit, we see that ρ_i and ρ_f play symmetric roles. More general tensor networks with insertions on links provide a powerful framework for describing physical processes of quantum many-body systems. Much like a partition function, a tensor network is an object into which operators can be inserted to compute correlation functions. However, partition functions require a Hamiltonian or action that implicitly or explicitly specifies spatial and temporal degrees of freedom. For instance, Hamiltonians and actions specify dynamical degrees of freedom such as spins, particles or fields, and designate both spatial and temporal coordinates. By contrast, a tensor network is a completely general contraction of quantum operators which is a priori agnostic to distinctions of space and time.

Going back to our example, we have so far viewed the network in Figure 4.3(b) as an initial state with unitary time evolution vertically and two operator insertions at A_x, B_y. However, the tensor network is agnostic to the words we use to describe it: we could instead equivalently say that the tensor network implements non-unitary evolution *horizontally*, and that what we formerly called *spatial* open boundary conditions correspond here to *temporal* boundary conditions (such as initial and final states). From this perspective, ρ_i and ρ_f now play the role of *spatial* boundary conditions. Also from this point of view, the operator insertions A_x, B_y compute a two-point correlation function in a different physical system.

This example may seem somewhat contrived, since we intuitively know that viewing the tensor network as implementing evolution vertically yields the familiar form of unitary time evolution, whereas viewing the tensor network as implementing evolution horizontally leads to peculiar non-unitary evolution. Thinking carefully about this distinction, we might ask: what precisely makes the "vertical" point of view more natural than the "horizontal" point of view, for this example? More generally, we may have a tensor network that does not have an obvious causal structure. So then we may ask, how do we diagnose the causal structure of a general tensor network? Which tensor networks yield familiar causal structures, either exactly or approximately? Are there new kinds of causal structures which are natural but specific to quantum systems? These are the questions which we begin to study in this paper.

Now, let us give the most general definition of a tensor network:

Definition (general tensor network): *A tensor network is specified by a triple* $\{\{\mathcal{H}_i\}, |L\rangle, \rho_P\}$

comprised of:

1. *A set of Hilbert spaces* $\{\mathcal{H}_i\}$ *which each correspond to a spacetime subsystem* i ,
2. *A link state* $|L\rangle \in \mathcal{H} = \bigotimes_i \mathcal{H}_i$,
3. *A density operator* ρ_P *acting on the same Hilbert space* $\mathcal{H}$.

The most general correlation function of the tensor network is computed by $\langle L|Q_1\,\rho_P\,Q_2|L\rangle$ *where* Q_1, Q_2 *are operators acting on* $\mathcal{H}$.

In other words, a general tensor network is like a quantum many-body state given by $|L\rangle$, except that the inner product is defined by positive semi-definite quadratic form ρ_P instead of the ordinary inner product in the Hilbert space. Furthermore, a general tensor network can encode correlations in time, since we regard each tensor factor $\mathcal{H}_i$ as a subsystem in spacetime. For instance, if our tensor network described standard unitary time evolution, the contracted tensor network would have unitary time evolution operators connecting subsystems corresponding to adjacent times.

4.2.2 Tensor networks based on graphs

Here we explain a useful type of tensor network, called a *graphical tensor network* (GTN). We will utilize GTN's throughout the paper. A GTN is defined for an undirected graph $G = (V, E)$ where V is the set of vertices and E is the set of edges. For a given vertex v, let $\deg(v)$ (i.e., the degree of v) denote the number of edges which attach to it. The GTN corresponding to G has a Hilbert space

$$\mathcal{H} = \bigotimes_{v \in V} \mathcal{H}_v \tag{4.3}$$

where $\mathcal{H}_v \simeq (\mathbb{C}^d)^{\otimes \deg(v)}$. In words, each Hilbert space $\mathcal{H}_v$ corresponding to a vertex v comprises of $\deg(v)$ tensored copies of $\mathbb{C}^d$, also known as $\deg(v)$ qudits.[1] It will be convenient to write the full Hilbert space as

$$\mathcal{H} = \bigotimes_{v \in V} \bigotimes_{j=1}^{\deg(v)} \mathcal{H}_{v_j} \tag{4.4}$$

where $\mathcal{H}_{v_j} \simeq \mathbb{C}^d$, and v_j denotes the jth qudit of $\mathcal{H}_v$.

Then $|L\rangle$ is a "link state" comprised of a tensor product of EPR pair states as follows. (The explanation of the construction of $|L\rangle$ is slightly involved, but has a simple pictographic interpretation given in Figure 4.5 above). Let us denote by (v, w) an edge e of the graph which connects the vertices v and w. Since our graph G is undirected, (v, w) is an unordered pair. Now we define a function f which assigns a pair of qudits to each edge e. The function f has two properties:

[1] A qudit is a d-level system (hence qu$d$it), whereas a qubit is a 2-level system.

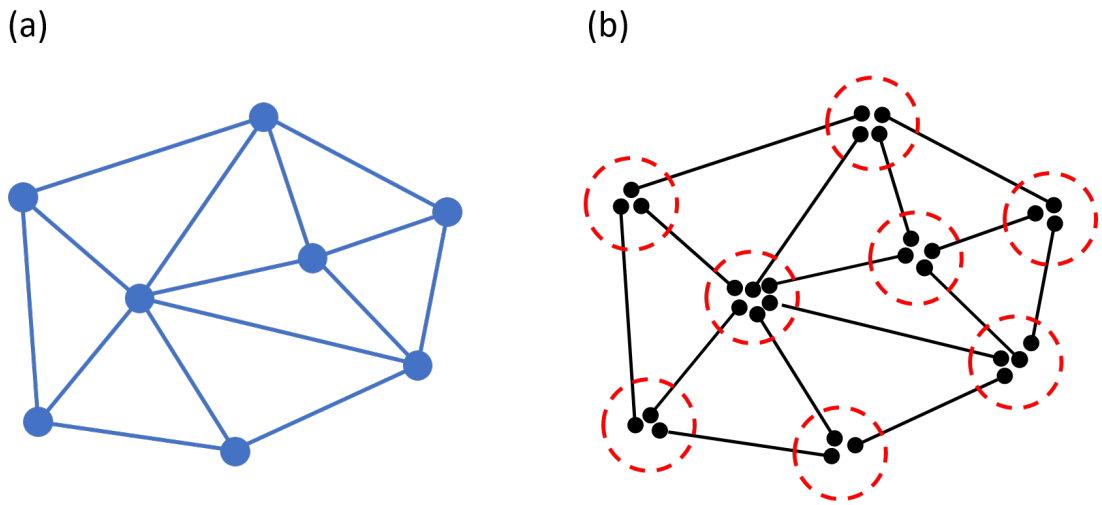


Figure 4.5: (a) A graph $G = (V, E)$ is shown in blue. (b) A representation of the link state $|L\rangle$. Each line with a dot at each end represents an EPR pair, with the dots corresponding to qudits. The dotted red circles designate the collections of qudits corresponding to vertices v of the graph G. The number of qudits at a vertex of the graph is the same as the degree of that vertex.

1. $f((v, w)) = \{v_m, w_n\}$ for some m, n with $1 \leq m \leq \deg(v)$ and $1 \leq n \leq \deg(w)$. In words, in this case f assigns (v, w) to the mth qudit of $\mathcal{H}_v$ and the nth qudit of $\mathcal{H}_w$.

2. For every pair of distinct edges e, e', we have $f(e) \cap f(e') = \emptyset$. In words, f assigns to each edge e a unique pair of qudits which does not intersect with the qudits assigned to any other edge.

Let $|\mathrm{EPR}_{v_m w_n}\rangle$ denote some EPR state, say $\frac{1}{\sqrt{d}} \sum_{i=1}^{d} |i\rangle|i\rangle$, between the mth qudit of $\mathcal{H}_v$ and the nth qudit of $\mathcal{H}_w$. Then $|L\rangle$ is given by

$$|L\rangle = \bigotimes_{e \in E} |\mathrm{EPR}_{f(e)}\rangle \,. \tag{4.5}$$

For clarity, consider the graph in Figure 4.5(a) above. Then we can visualize $|L\rangle$ by EPR pairs organized as in Figure 4.5(b) above. Indeed, we can imagine that the edges of the graph have been "replaced" by EPR pairs. Finally, the state ρ_P has the structure

$$\rho_P = \bigotimes_{v \in V} P_v \tag{4.6}$$

where P_v is a projector on $\mathcal{H}_v$. Hence, ρ_P is furnished with a subscript P (for "projector"). In some graph-based tensor networks, ρ_P is not restricted to comprise of a tensor product of projectors, and can instead be any density matrix on $\bigotimes_{v \in V} \mathcal{H}_v$.

As an example of a GTN, we consider correlation functions in a matrix product state (MPS)

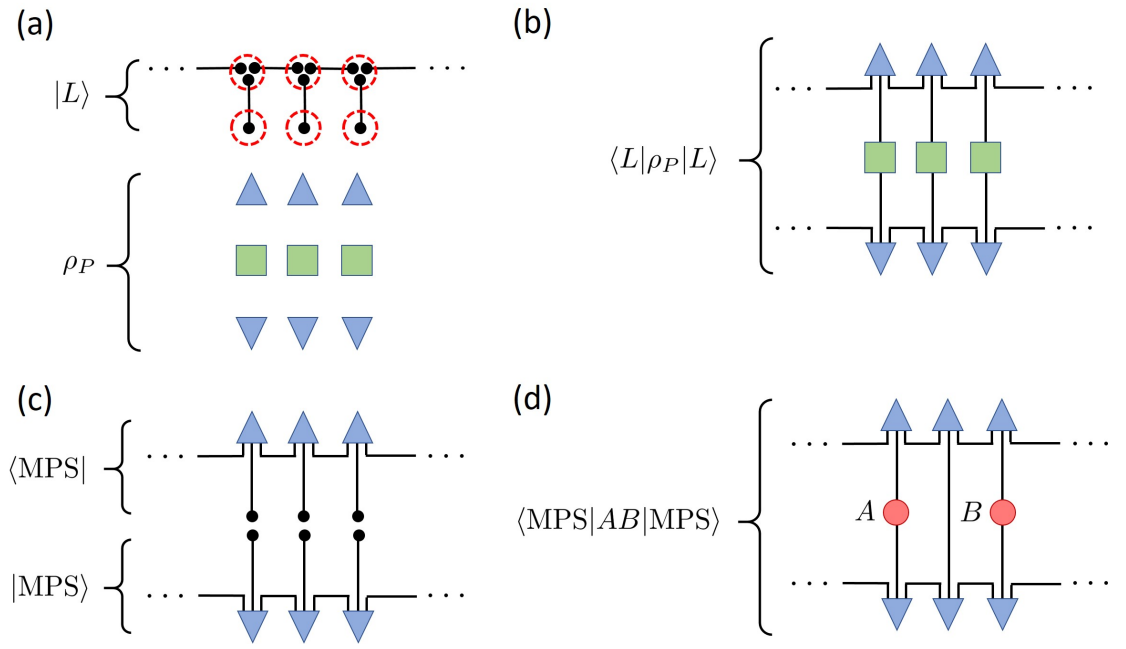


Figure 4.6: (a) A diagrammatic representation of $|L\rangle$ and ρ_P for a nascent MPS tensor network. The blue triangles represent the 3-qudit pure states $\langle\varphi|$ (each upper triangle) and $|\varphi\rangle$ (each lower triangle), and the green boxes are 1-qudit identity operators. Therefore, $\rho_P = (|\varphi\rangle\langle\varphi| \otimes \mathbf{1})^{\otimes N}$ for some N. (b) A diagrammatic representation of $\langle L|\rho_P|L\rangle$. The green boxes can be omitted since they are identity operators. (c) If we split $\langle L|\rho_P|L\rangle$ by cutting through the vertical links, we obtain two MPS states. (d) A diagram of the two-point function $\langle\text{MPS}|\, AB\, |\text{MPS}\rangle$.

tensor network. To construct the MPS tensor network, we start with a link state $|L\rangle$ and density operator $\rho_P = (|\varphi\rangle\langle\varphi|\otimes\mathbf{1})^{\otimes N}$ for some N, as depicted in Figure 4.6(a). Here, $\langle\varphi|$ and $|\varphi\rangle$ are 3-qudit states, and are each represented, respectively, by an upper and lower blue triangle in Figure 4.6(a). The identity operator $\mathbf{1}$ acts on one qudit, and is depicted as a blue box in Figure 4.6(a). Contracting $\langle L|$ and ρ_P and $|L\rangle$ as $\langle L|\rho_P|L\rangle$, we obtain the tensor network in Figure 4.6(b). Here, the green boxes can be omitted since they are just identity operators. We can sever the vertical links to obtain two MPS states $|\text{MPS}\rangle$ and $\langle\text{MPS}|$, as in Figure 4.6(c). Indeed, we have $\langle\text{MPS}|\text{MPS}\rangle = \langle L|\rho_P|L\rangle$. Finally, to compute correlation functions of the MPS state $|\text{MPS}\rangle$, we contract $\langle\text{MPS}|$ and A and B and $|\text{MPS}\rangle$ to obtain $\langle\text{MPS}|AB|\text{MPS}\rangle$, which is depicted by the tensor network in Figure 4.6(d).

The Trotter networks in Fig.'s 4.3(b), 4.3(c), and 4.4 are also examples of GTN's. For these GTN's, the state ρ_P is

$$\rho_P = \rho_i \otimes \bigotimes_{t=1}^{M} \left(\bigotimes_i |U_{i,i+1}\rangle\langle U_{i,i+1}| \otimes \bigotimes_j |V_{j,j+1}\rangle\langle V_{j,j+1}| \right) \tag{4.7}$$

where $|U_{i,i+1}\rangle$ and $|V_{j,j_1}\rangle$ are Choi-Jamiolkowski representations of the local unitary operators $U_{i,i+1}$ and $V_{j,j+1}$. For instance, for a unitary two-qubit gate $U_{i,i+1}$ with matrix elements $[U_{i,i+1}]^{\gamma\delta}_{\alpha\beta}$ in some basis, one can define its Choi-Jamiolkowski representation which is the four-qubit state

$$|U_{i,i+1}\rangle = \frac{1}{2}\sum_{\alpha\beta\gamma\delta}[U_{i,i+1}]^{\gamma\delta}_{\alpha\beta}\,|\alpha\rangle\,|\beta\rangle\,|\gamma\rangle\,|\delta\rangle\ .$$

The states $|V_{j,j+1}\rangle$ are represented similarly.

Then $|L\rangle$ comprises of qubit EPR pairs which link together the Choi-Jamiolkowski representations of the local unitary operators $\{U_{i,i+1}\}$ and $\{V_{j,j+1}\}$, as well as the initial state ρ_i, to form the tensor networks in Fig.'s 4.3(b), 4.3(c), and 4.4. Here, the role of $|L\rangle$ is to "unwrap" the Choi-Jamiolkowski isomorphism and glue the the appropriate unitaries together in space (for instance, $U_{i,i+1}$ should linked on the right with $U_{i+1,i+2}$) and in time (for instance, U's are followed in the next time step by V's).

Although much of the tensor network literature is centered around GTN's, our discussion of quantum causal influence below applies to general tensor networks.

4.2.3 Defining quantum causal influence

In the framework of general tensor networks, we now define our measures of quantum causal influence. Roughly speaking, the key idea is to distinguish causal influence from other forms of correlation by using unitary operators. The causal influence of a region R_1 on a region R_2 is characterized by how correlations within R_2 can be changed by arbitrarily varying a unitary operator acting on R_1. As a prerequisite for this discussion, a unitary acting on R_1 has to preserve the norm of the tensor network, namely

$$\langle L|U_{R_1}\,\rho_P\,U^\dagger_{R_1}|L\rangle = \langle L|\rho_P|L\rangle\,, \tag{4.8}$$

which is generically not true due to the "metric" ρ_P. Therefore we define the concept of *unitary regions*.

Consider a tensor network with a Hilbert space decomposition into subsystems as $\mathcal{H} = \bigotimes_{i\in\Omega}\mathcal{H}_i$, where Ω indexes the subsystems. Let us call the subsystems indexed by Ω the fundamental subsystems, since they are prescribed by the definition of the tensor network. A *unitary region* is a subsystem R, with $R \subseteq \Omega$, and an associated Hilbert space $\mathcal{H}_R = \bigotimes_{i\in R}\mathcal{H}_i$ such that

$$\langle L|\,U_R\,\rho_P\,U^\dagger_R\,|L\rangle = \langle L|\,\rho_P\,|L\rangle \tag{4.9}$$

for arbitrary unitaries U_R supported on R. In other words, a unitary region is a subsystem for which acting with local unitaries preserves the norm of the tensor network. We also say that two regions

R_1, R_2 are *mutually unitary regions* if

$$\langle L| U_{R_1} U_{R_2} \rho_P U_{R_2}^\dagger U_{R_1}^\dagger |L\rangle = \langle L| \rho_P |L\rangle \tag{4.10}$$

for arbitrary unitaries U_{R_1} supported on R_1 and arbitrary unitaries U_{R_2} supported on R_2. Notice that if R_1, R_2 are mutually unitary regions, then they are each unitary regions individually. The converse is not generally true.

For concreteness, in the Trotter networks in Fig.'s 4.3(a), 4.3(b), 4.3(c) and 4.4, we can define 45^o lines as "light cones." Using these light cones, it is easy to see that all regions that only contain only "spacelike" separated points are unitary regions. All pairs of such regions are in fact mutually unitary regions. In contrast, a region with two time-like separated points x, y is not a unitary region. As another example, for a general MPS tensor network as depicted in Figure 4.6(d), only the sites obtained by breaking apart vertical links are unitary regions.

Given a unitary region R_1, its causal influence on another region R_2 is reflected in the following quantity:

$$M(U_{R_1} : O_{R_2}) := \langle L| (U_{R_1} \otimes O_{R_2}) \rho_P (U_{R_1}^\dagger \otimes O_{R_2}^\dagger) |L\rangle \tag{4.11}$$

If $M(U_{R_1} : O_{R_2})$ has nontrivial dependence on U_{R_1}, this means that physical operations on region R_1 have a nontrivial causal influence on physical observables in region R_2.

Using $M(U_{R_1} : O_{R_2})$, one can define different measures of quantum causal influence that are independent from the choice of operators U_{R_1}, O_{R_2}. For example, one can define the *maximal quantum causal influence* (henceforth, mQCI)

$$\mathrm{CI}(R_1 : R_2) = \sup_{U_{R_1}, O_{R_2}} \frac{1}{||O_{R_2}||_2^2} \left| M(U_{R_1} : O_{R_2}) - \int dU_{R_1}\, M(U_{R_1} : O_{R_2}) \right|, \tag{4.12}$$

and the *averaged quantum causal influence* (henceforth, aQCI)

$$\overline{\mathrm{CI}}(R_1 : R_2) = \int dU_{R_1} \int_{||O_{R_2}||_2^2 = 1} dO_{R_2} \left| M(U_{R_1} : O_{R_2}) - \int dU_{R_1}\, M(U_{R_1} : O_{R_2}) \right|^2 \tag{4.13}$$

where in Eqn.'s (4.12) and (4.13), U_{R_1} is integrated via the Haar measure, and in Eqn. (4.13) O_{R_2} is averaged with the uniform measure on the unit sphere defined by $||O_{R_2}||_2^2 = 1$ in the linear space of operators O_{R_2}. In the rest of the paper, when we discuss whether the quantum causal influence is zero or non-zero, we do not need to distinguish between the mQCI and aQCI, and so will refer to the QCI more broadly. In Section 4.6, we will discuss more quantitative properties of the aQCI. Variations of quantum causal influence for non-unitary regions can be found in Appendix A. A discussion of causal influence for classical systems is in Appendix C.

With our definitions at hand, we would like to gain more intuition about quantum causal influence

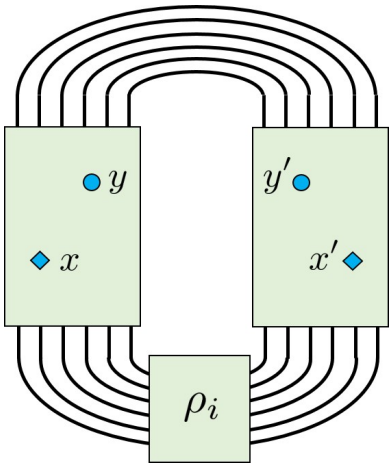


Figure 4.7: A spacetime state with initial state ρ_i. Two spacetime points x and y are designated, along with their mirror copies x' and y'.

by studying some of its key features through various examples.

4.3 Boundary condition dependence of quantum causal influence

Before discussing more abstract properties of quantum causal influence for general tensor networks, we first present examples which exhibit interesting causal features. Our examples in Fig.'s 4.3(b), 4.3(c) and 4.4 in the previous section have a natural form which can be abstracted as follows. They comprise of some initial state ρ_i conjugated by some (not necessarily unitary) operator W which implements evolution, followed by a trace.

A more abstract representation is drawn in Figure 4.7. We call such a representation a "spacetime state" to distinguish it from other kinds of tensor networks. The green boxes on either side of ρ_i represent W (on the left) and $W^\dagger$ (on the right). The tensor contractions at the top of the diagram represent a trace. Analogously to Fig.'s 4.3(b), 4.3(c) and 4.4 which comprise of a mesh of links (i.e., EPR pairs), we treat the W and $W^\dagger$ boxes in Figure 4.7 as comprised of a mesh of links which we can break open to insert operators. For instance, in Figure 4.7 we label the positions of two (hidden) links x and y, which can be broken to insert operators. We imagine that x and y are spacetime points. Likewise, x' and y' are mirroring spacetime points. By inserting A into x, B into y, $A^\dagger$ into x' and $B^\dagger$ into y', the tensor network computes

$$\langle \mathcal{P}\, B_y\, A_x\, \rho_i\, A_x^\dagger\, B_y^\dagger \rangle \tag{4.14}$$

where the path ordering $\mathcal{P}$ is defined by the contracted tensor network. Indeed, if W corresponds

to Hamiltonian time evolution or some discrete-time analog thereof, then Eqn. (4.14) is merely a standard correlation function with an initial state ρ_i. In this case, we imagine that slicing the W or $W^\dagger$ boxes along a horizontal line and contracting operators with the exposed links corresponds to operator insertions at a fixed intermediate time. This is directly analogous to Fig.'s 4.3(b), 4.3(c) and 4.4.

The causal structure of a spacetime state can depend on its boundary conditions – namely the initial state ρ_i, and the trace taken over $W\rho_i W^\dagger$. In this section, we illustrate the boundary condition dependence of causal influence in spacetime states in several examples. Our results suggest an explanation of "time's arrow" in a quantum many-body system.

4.3.1 Initial state dependence

Suppose we have a spacetime state comprised of an initial state $\rho_i = |\psi\rangle\langle\psi|$ which is then unitarily evolved in time. In other words, W implements unitary time evolution. As mentioned above, slicing the W or $W^\dagger$ boxes along a horizontal line and contracting operators with the exposed links corresponds to operator insertions at a fixed intermediate time. In Figure 4.7, we allow insertions of operators into the spacetime points x and y, and then contract the spacetime state (i.e., take its trace) at some later time. Unpacking Eqn. (4.9) for our case, we find that x is a unitary region if

$$\langle \mathcal{P}\, U_x\, \rho_i\, U_x^\dagger \rangle = \langle \mathcal{P} \rho_i \rangle\,, \tag{4.15}$$

and similarly for y,

$$\langle \mathcal{P}\, U_y\, \rho_i\, U_y^\dagger \rangle = \langle \mathcal{P} \rho_i \rangle\,. \tag{4.16}$$

Each of the above equations is satisfied, and so any such points x and y are unitary regions. In fact, we have also

$$\langle \mathcal{P}\, U_x\, U_y \rho_i\, U_y^\dagger\, U_x^\dagger \rangle = \langle \mathcal{P} \rho_i \rangle\,, \tag{4.17}$$

for all such pairs x, y, and so all pairs of points x, y form mutually unitary regions.

Say that we insert a unitary U_y at y and $U_y^\dagger$ at y'. This U_y and $U_y^\dagger$ will cancel one another along the upper contraction of the spacetime state in Figure 4.7. The reason is that the unitary evolution that occurs after y and y' cancels across the trace – see, for instance, the red boxes in Figure 4.8. These red boxes clearly cancel across the trace (i.e., the upper contracted legs), and so allow U_y at y and $U_y^\dagger$ at y' to similarly cancel. If we insert some Hermitian operator O_x at x and $O_x^\dagger$ at x', then these operators will be unaffected by the cancellation of U_y and $U_y^\dagger$. Therefore,

$$M(U_y : O_x) = \langle \mathcal{P}\, U_y\, O_x\, \rho_i\, O_x^\dagger\, U_y^\dagger \rangle \tag{4.18}$$

is independent of U_y, and thus

$$\mathrm{CI}(y\,:\,x) = 0$$

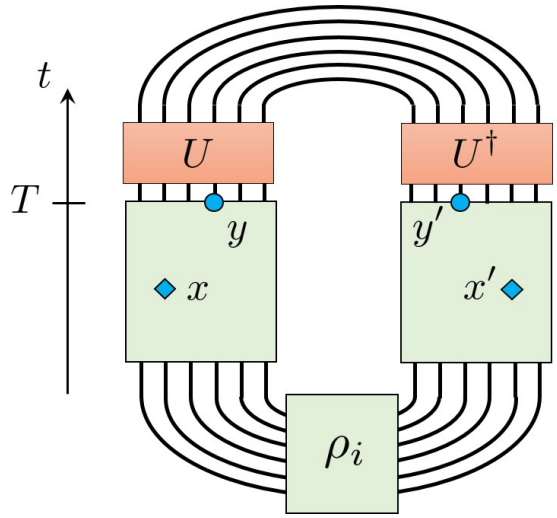


Figure 4.8: A spacetime state, such that operators are not inserted later than a time T. Then the unitary evolution U after time T cancels out with the corresponding unitary evolution $U^\dagger$.

meaning that y does not influence x. Similarly, $\overline{\text{CI}}(y : x) = 0$, although we will focus on the mQCI in this section.

But now suppose that we insert U_x at x and O_y at y, and $U_x^\dagger$ at x' and $O_y^\dagger$ at y'. We cannot cancel out U_x with $U_x^\dagger$ along the lower contraction of the spacetime state, since we are obstructed by the boundary condition ρ_i (i.e., the initial state). We might be able to cancel U_x with $U_x^\dagger$ along the upper contraction of the spacetime state, but the operator insertions O_y and $O_y^\dagger$ may obstruct us. If O_y and $O_y^\dagger$ *obstruct* the cancellation of U_x and $U_x^\dagger$ along the upper contraction, then $M(U_x : O_y)$ *would* depend on U_x, and thus $\text{CI}(x : y) \neq 0$. In summary, we would have

$$\text{CI}(y : x) = 0 \quad \text{and} \quad \text{CI}(x : y) \neq 0 \quad \text{implies } y \text{ is in the } \textit{future} \text{ of } x\,.$$

If instead O_y and $O_y^\dagger$ do *not* obstruct cancellation of U_x and $U_x^\dagger$ along the upper contraction, then $M(U_x : O_y)$ would *not* depend on U_x, and so $\text{CI}(x : y) = 0$. Then in this case, we would have

$$\text{CI}(y : x) = 0 \quad \text{and} \quad \text{CI}(x : y) = 0 \quad \text{implies } x \text{ and } y \text{ are } \textit{spacelike} \text{ separated.}$$

The interesting feature here is that the state ρ_i induces a causal structure in which time flows *away from* ρ_i via the unitary evolution comprising the spacetime state. In other words, the initial state has picked out a preferred arrow of time. Crucially, there is not a "final state" at the top contraction of the spacetime state. This is perfectly physical, since we often start in an initial state and evolve it up to some time, perhaps making operator insertions intermediately. If we only

consider operator insertions up to a finite time T, then we only have to consider the spacetime state evolved up until that T. If we evolve the state further thereafter, when computing expectation values this additional time evolution would cancel out, as depicted in Figure 4.8. In the Figure, the time evolution U in the left red box cancels out the time evolution $U^\dagger$ in the right red box.

There is another complementary perspective which is useful. Instead of thinking of the upper end of the spacetime state (where the trace is) as a "cutoff time" after which we do not care about making operator insertions, we can instead imagine that we are inserting a *maximally mixed state* $\mathbf{1}/d$ as a final-time state. Here, d is the Hilbert space dimension of a spatial slice. As far as any of our analysis is concerned, these two perspectives are mathematically equivalent, up to an overall multiplicative rescaling of the spacetime state by d. The benefit of this change of perspective is that we can think about ρ_i and $\mathbf{1}/d$ on more equal footing. In particular, we can say:

- The initial state ρ_i can obstruct unitary cancellation across the initial-time boundary.
- The final state $\mathbf{1}/d$ can allow unitary cancellation across the final-time boundary.

In this manner, the initial state ρ_i acts as a barrier and a *source* of causal flow, and the final state $\mathbf{1}/d$ acts as a passageway or *sink* of causal flow. It is no coincidence that the flow of time coincides with the disparity between the entropy of the initial and final states: namely, we have the von Neumann entropies $S[\rho_i] = 0$ and $S[\mathbf{1}/d] = \log(d)$ and so time is flowing from a lower entropy state to a higher entropy state. One might naïvely guess that more generally, given an initial state ρ_i and final state ρ_f, there would be a forward arrow of time if $S[\rho_i] < S[\rho_f]$, but this is not generally true. There needs to be additional relations between ρ_i and ρ_f to get a forward arrow of time, but we will leave this for future work.

Now suppose that we choose both the initial state ρ_i and the final state ρ_f to be the maximally mixed state, namely $\rho_i = \mathbf{1}/d$, and that we multiplicatively rescale the resulting spacetime state by d. Then we have

$$\mathrm{CI}(x : y) = 0$$

meaning that x does not causally influence y. Similarly, we also have

$$\mathrm{CI}(y : x) = 0$$

meaning that y does not causally influence x. Then x and y are spacelike separated. Indeed, when the past and future are maximally mixed states, the unitary evolution in between does not impose a particular directionality of time.

4.3.2 Conceptual remarks

In standard discussions of the arrow of time, a key ingredient is that the initial conditions of the universe provide a low-entropy initial state.[2] Tied to the arrow of time is the production of coarse-grained entropy, and ultimately the universe becomes a high-entropy equilibrium state. Once the universe has reached equilibrium, there ceases to be an arrow of time in any conventional sense, since there is no longer entropy growth. In blunt terms, there are no local clocks in thermal equilibrium.

In the context of this paper, we find a new twist on these ideas. Above, we found that when both boundaries of a spacetime state are maximally mixed, which we can think of as infinite temperature (or maximum entropy) states, all pairs of spacetime points in between are spacelike separated. If we attach the word "past" to one of the boundaries and attach the word "future" to the other boundary, we can say: *If the putative past and future have maximal entropy, then all spacetime points in between are spacelike separated and there is no flow of time.*

We also saw that by fixing one of the boundaries to be a low-entropy state, such as a pure state, we can induce an arrow of time. We will later show that by imposing more interesting boundary conditions on *both* boundaries, we can have even richer causal structures and local arrows of time. Intuitively, we will see that for fine-tuned boundary conditions, regions of boundary states which have higher and lower entropies act as sinks and sources for causal flow, respectively, which is consistent with more conventional intuitions from thermodynamics. Presumably some version of our analysis applies to more general initial and final states, but such a generalization is beyond the scope of this work.

4.3.3 Trotterized tensor network

A nice example of a spacetime state which implements the above constructions is a Trotterized tensor network, such as in Fig.'s 4.3(b), 4.3(c) and 4.4 above. For example, consider Figure 4.9 below which is a spacetime state with Trotterized time evolution and initial state ρ_i. We see that in the contracted network, $\mathrm{CI}(x : y) = 0$ unless y is in a future cone of x, which is in fact the future light cone of x. Notice that Figure 4.9 is folded relative to the spacetime states in Fig.'s 4.7 and 4.8 – in particular, ρ_i is in the middle, W is on top, $W^\dagger$ is on the bottom, and the trace is looped behind.

As we discussed earlier, the quantum causal structure generically depends on the initial state. For example, consider the spacetime state in Figure 4.10, which has an initial state $\mathbf{1}_R/d_R \otimes \rho_{\overline{R}}$. The figure only displays part of the tensor network, namely $W\left(\mathbf{1}_R/d_R \otimes \rho_{\overline{R}}\right)$, and we have not depicted $W^\dagger$ or the trace.[3] Since the initial state is maximally mixed on a subregion R, the spacetime has an interesting causal structure. For instance, applying a unitary U to x_1 can cancel with a $U^\dagger$ applied to x_1' across the R region at the initial time, rather than canceling across the trace at the final

[2] In our universe, it seems that cosmic inflation provides us with such a low-entropy initial state.

[3] The full diagram would give us $\mathrm{tr}(W\left(\mathbf{1}_R/d_R \otimes \rho_{\overline{R}}\right) W^\dagger)$.

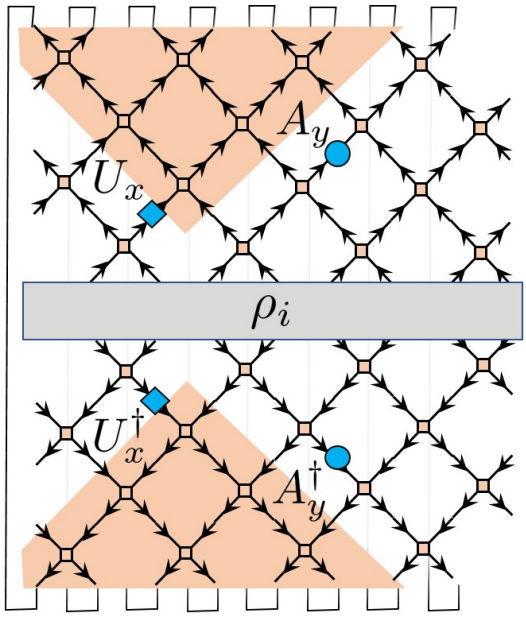


Figure 4.9: In the Trotter network, $\mathrm{CI}(x : y) = 0$ unless y is in the future light cone of x.

time. Consequently, the quantum causal influence of x_1 on any point in its usual future light cone[4] vanishes. Similarly, x_1 does not causally influence any point in its usual past light cone because unitaries acting at x_1 can still be canceled at the future boundary. Therefore, x_1 does not causally influence any *single site* regions. However, x_1 can have a quantum causal influence on larger regions. When we consider a spacetime region that overlaps with both the usual future light cone and usual past light cone of x_1, such as $y_1 \cup y_2$, the quantum causal influence $\mathrm{CI}(x_1, y_1 \cup y_2)$ is generically non-zero since it is not possible to push a unitary operator at x_1 to either the future boundary or the past boundary (since it is obstructed by the operators inserted at both y_1 and y_2) to cancel with a corresponding Hermitian conjugate unitary.

More generally, any region A in the domain of dependence of R (the red shaded region in Figure 4.10) does not causally influence its usual causal future $I^+(A)$. The only regions that are causally influenced by A are those that overlap with both the usual causal future $I^+(A)$ and the usual causal past $I^-(A)$. Thus, we see that specifying a special initial state may erase some regions from the causal future of a given region. Although some of the causal future of a given region may be erased (such as y_2), nonlocal regions can still remain in the causal future (such as $y_1 \cup y_2$). These observations are quite general, and we will see them instantiated in many contexts throughout the paper.

[4]The usual future light cone of a point is defined by extending 45^o lines from that point, as per Figure 4.9. This "usual" future light cone is in fact the region which a point can causally influence if the initial state is pure.

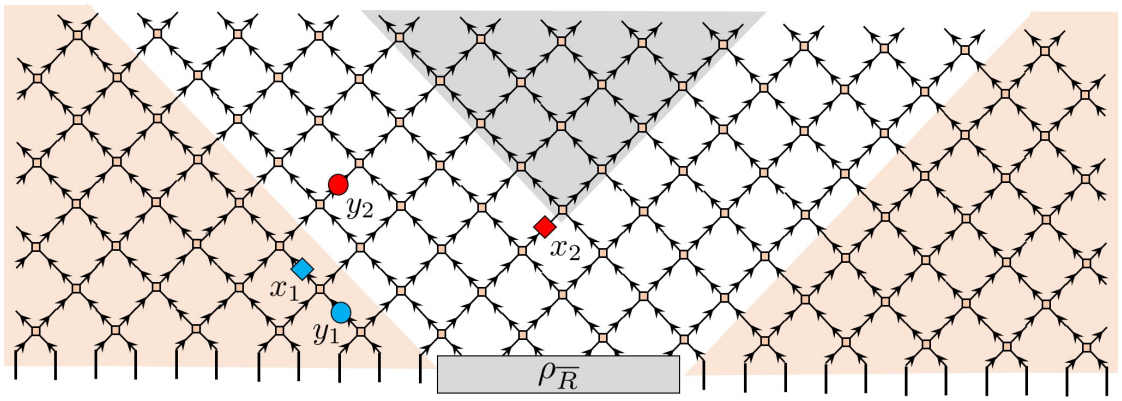


Figure 4.10: In a Trotter network with a special initial state, some causal influence can be lost.

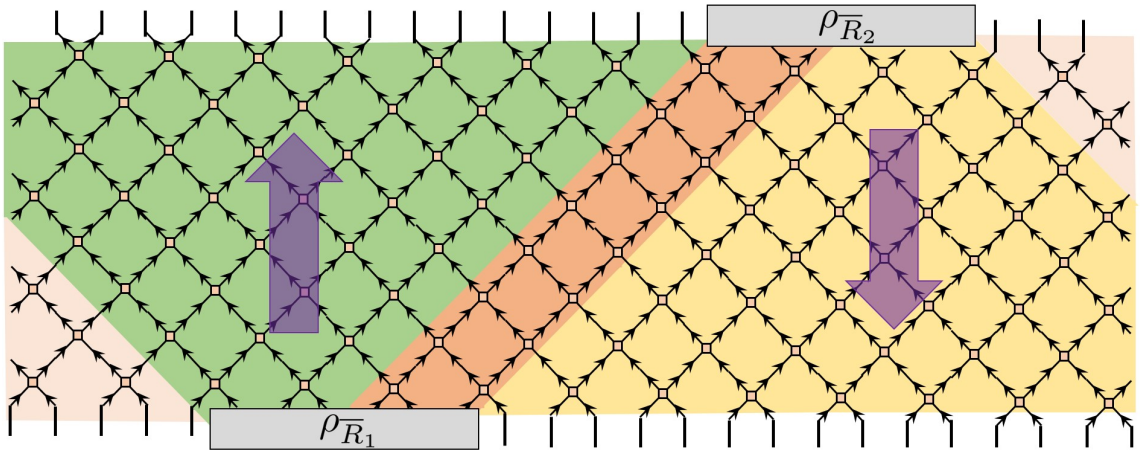


Figure 4.11: In a Trotter network with a special initial *and* final state, there can be multiple arrows of time, as depicted by the arrows.

4.3.4 Final state dependence (post-selection)

There are many possibilities for including both initial and final states (i.e., pre-selection and post-selection), but we will only examine one case here to give a general flavor for the sorts of causal structures that can occur. Consider the spacetime state comprised of Trotterized time evolution in Figure 4.11, with initial state $\mathbf{1}_{R_1}/d_{R_1} \otimes \rho_{\overline{R}_1}$ and final state $\mathbf{1}_{R_2}/d_{R_2} \otimes \rho_{\overline{R}_2}$. Similar to the previous figure, this figure only displays part of the tensor network, namely

$$(\mathbf{1}_{R_2}/d_{R_2} \otimes \rho_{\overline{R}_2})\, W\, (\mathbf{1}_{R_1}/d_{R_1} \otimes \rho_{\overline{R}_1})\,.$$

Accordingly, we have not depicted $W^\dagger$ or the trace.[5] Suppose that $\overline{R}_1$ and $\overline{R}_2$ are regions of the same size, and that $\rho_{\overline{R}_1} = \rho_{\overline{R}_2}$ are pure states. Then we see than there is a flow of time from bottom to top in the region shaded in green, but there is a flow of time from top to bottom in the region shaded in yellow. Then every pair of points in the pink regions are spacelike separated, and the region in orange is not even a unitary region (and so, in a sense, does not have any preferred direction of time at all). (See Appendix A for diagnostics quantum causal influence within nonunitary regions.) This example emphasizes that pure states act as sources of causal flow, and maximally mixed states act as sinks of causal flow. The pink regions are created by two sinks of causal flow (i.e., the maximally mixed states on each boundary), whereas the orange region is due to the interplay of two sources of causal flow (i.e., the pure states $\rho_{\overline{R}_1}$ and $\rho_{\overline{R}_2}$).

4.3.5 General results

In this subsection we summarize some generic features that can be observed from examples above, and describe them more quantitatively.

Sinks of causal flow

Having worked through explicit examples of the interplay between the initial and final states of a spacetime state and its causal structure, we now move towards more general and abstract results. First, we present a result about GTN's that has played a role in all of the above examples. The result generalizes the observed fact that in spacetime states, maximally mixed subsystems of initial and final states act as sinks of causal flow.

Suppose we have a GTN on a graph $G = (V, E)$, with the structure specified in Section 4.2.2. As per Eqn. (4.4), the corresponding Hilbert space is

$$\mathcal{H} = \bigotimes_{v \in V} \bigotimes_{j=1}^{\deg(v)} \mathcal{H}_{v_j} \, .$$

Let $\Sigma \subseteq V$ be a subset of the vertices (which may correspond to a subregion in a putative spacetime), and partition V as $V = \Sigma \cup \overline{\Sigma}$. We can write the link state $|L\rangle$ as

$$|L\rangle = |L_{\Sigma \leftrightarrow \Sigma}\rangle \otimes |L_{\Sigma \leftrightarrow \overline{\Sigma}}\rangle \otimes |L_{\overline{\Sigma} \leftrightarrow \overline{\Sigma}}\rangle \, . \tag{4.19}$$

In the above equation,

- $|L_{\Sigma \leftrightarrow \Sigma}\rangle$ are the EPR pairs associated with edges $e = (v, w)$ with $v, w \in \Sigma$;
- $|L_{\Sigma \leftrightarrow \overline{\Sigma}}\rangle$ are the EPR pairs associated with edges $e = (v, w)$ with $v \in \Sigma$ and $w \in \overline{\Sigma}$;

[5] Here, the full diagram would give us $\mathrm{tr}\left[(\mathbf{1}_{R_2}/d_{R_2} \otimes \rho_{\overline{R}_2}) \, W \, (\mathbf{1}_{R_1}/d_{R_1} \otimes \rho_{\overline{R}_1}) \, W^\dagger\right]$.

- $|L_{\overline{\Sigma}\leftrightarrow\overline{\Sigma}}\rangle$ are the EPR pairs associated with edges $e = (v, w)$ with $v, w \in \overline{\Sigma}$.

See Fig.'s 4.12(a) and 4.12(b) for a diagrammatic depiction. So, for instance, each EPR pair in $|L_{\Sigma\leftrightarrow\overline{\Sigma}}\rangle$ comprises of one qudit in Σ and one qudit in $\overline{\Sigma}$. Let the Hilbert space of the qudits in $|L_{\Sigma\leftrightarrow\overline{\Sigma}}\rangle$ which lie in $\overline{\Sigma}$ be denoted by $\mathcal{H}_{\partial\Sigma}$. Then the total Hilbert space $\mathcal{H}$ decomposes as

$$\mathcal{H} = \mathcal{H}_{\Sigma} \otimes \mathcal{H}_{\partial\Sigma} \otimes \mathcal{H}_{\overline{\Sigma\cup\partial\Sigma}}\,. \tag{4.20}$$

Now, let $\rho_P^{\Sigma} := \mathrm{tr}_{\overline{\Sigma}}(\rho_P)$, and consider the state

$$\sigma^{\partial\Sigma} := \mathrm{tr}_{\Sigma}\left[(\rho_P^{\Sigma} \otimes \mathbf{1}_{\partial\Sigma})\ \left(|L_{\Sigma\leftrightarrow\Sigma}\rangle\langle L_{\Sigma\leftrightarrow\Sigma}| \otimes |L_{\Sigma\leftrightarrow\overline{\Sigma}}\rangle\langle L_{\Sigma\leftrightarrow\overline{\Sigma}}|\right)\right]\,. \tag{4.21}$$

This state $\sigma^{\partial\Sigma}$ is a density matrix on $\mathcal{H}_{\partial\Sigma}$. Now we make the following proposition:

Proposition: *Suppose we decompose $\mathcal{H}_{\partial\Sigma}$ into subsystems as*

$$\mathcal{H}_{\partial\Sigma} = \mathcal{H}_R \otimes \mathcal{H}_{\overline{R}}\,. \tag{4.22}$$

If we have

$$\sigma^{\partial\Sigma} = \frac{\mathbf{1}_R}{d_R} \otimes \rho_{\overline{R}} \tag{4.23}$$

for some $\rho_{\overline{R}}$, then

$$\mathrm{CI}(R : S) = 0 \tag{4.24}$$

for any region S such that $S \cap (\Sigma \cup \partial\Sigma) = \emptyset$, i.e., S does not intersect Σ or $\partial\Sigma$.

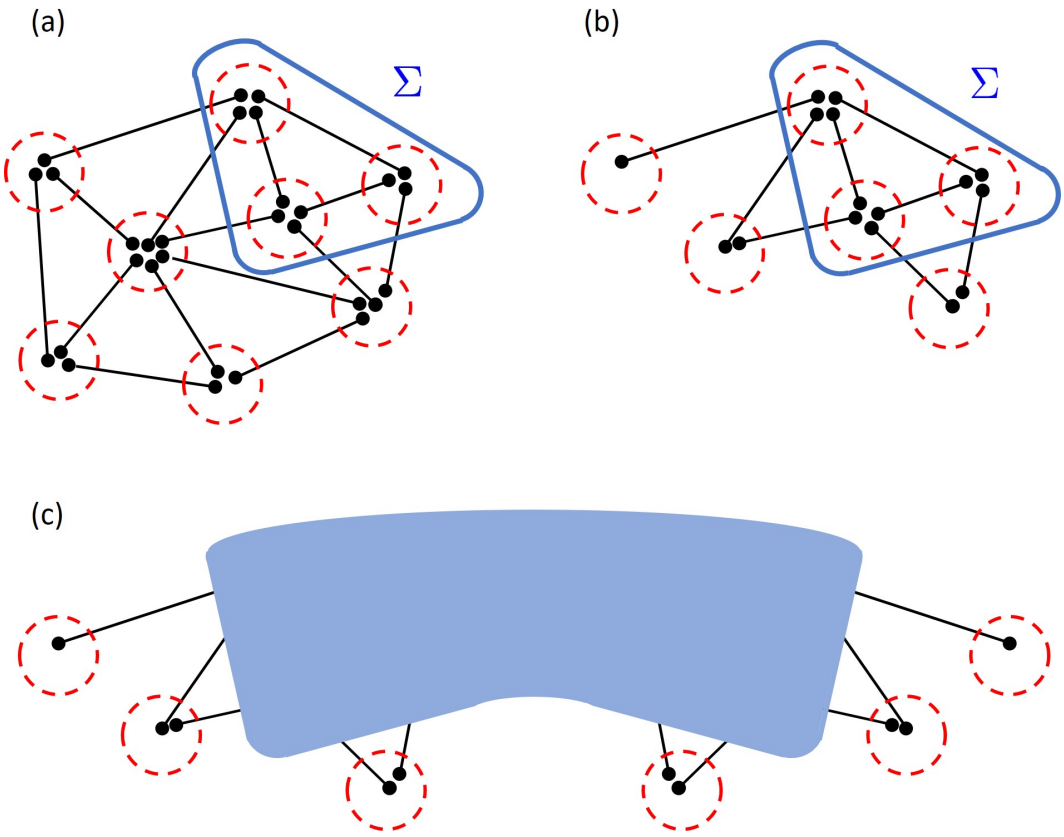


Figure 4.12: (a) The link state $|L\rangle$ is depicted. The region Σ is outlined in blue, and in this case contains 3 vertices and 11 qudits. (b) The state $|L_{\Sigma\leftrightarrow\Sigma}\rangle \otimes |L_{\Sigma\leftrightarrow\overline{\Sigma}}\rangle$ is shown. The EPR pairs lying within Σ form $|L_{\Sigma\leftrightarrow\Sigma}\rangle$, and the EPR pairs crossing the boundary of Σ form $|L_{\Sigma\leftrightarrow\overline{\Sigma}}\rangle$. The qudits lying outside of Σ form the Hilbert space $\mathcal{H}_{\partial\Sigma}$. (c) By taking two copies of $|L_{\Sigma\leftrightarrow\Sigma}\rangle \otimes |L_{\Sigma\leftrightarrow\overline{\Sigma}}\rangle$ and partially contracting their Σ regions with the state ρ_P^{Σ}, we obtain the density matrix $\sigma^{\partial\Sigma}$, which is depicted in the Figure. The light blue region represents the contraction of the Σ regions of $|L_{\Sigma\leftrightarrow\Sigma}\rangle \otimes |L_{\Sigma\leftrightarrow\overline{\Sigma}}\rangle$ and $\langle L_{\Sigma\leftrightarrow\Sigma}| \otimes \langle L_{\Sigma\leftrightarrow\overline{\Sigma}}|$ with ρ_P^{Σ}. We see that the density matrix $\sigma^{\partial\Sigma}$ maps $\mathcal{H}^*_{\partial\Sigma} \otimes \mathcal{H}_{\partial\Sigma} \to \mathbb{C}$, since a state on $\mathcal{H}_{\partial\Sigma}$ can be contracted with the exposed legs on the right-hand side, and a dual state on $\mathcal{H}_{\partial\Sigma}$ can be contracted with the exposed legs on the left-hand side.

Proof. Let us compute $M(U_R : O_S)$, where U_R is a unitary on R and O_S is some Hermitian operator

on S. Let $\rho_P^{\overline{\Sigma\cup\partial\Sigma}} = \mathrm{tr}_{\Sigma\cup\partial\Sigma}(\rho_P)$. Then

$$\begin{aligned}\langle L|U_R\, O_S\, \rho_P\, O_S^\dagger\, U_R^\dagger|L\rangle &= \mathrm{tr}\left[U_R\, O_S\, \rho_P\, O_S^\dagger\, U_R^\dagger\, |L\rangle\langle L|\right] \\ &= \mathrm{tr}_{\overline{\Sigma\cup\partial\Sigma}}\left[U_R\, O_S\, \rho_P^{\overline{\Sigma\cup\partial\Sigma}}\, O_S^\dagger\, U_R^\dagger\ \left(\sigma^{\partial\Sigma}\otimes|L_{\overline{\Sigma}\leftrightarrow\overline{\Sigma}}\rangle\langle L_{\overline{\Sigma}\leftrightarrow\overline{\Sigma}}|\right)\right] \\ &= \mathrm{tr}_{\overline{\Sigma\cup\partial\Sigma}}\left[O_S\, \rho_P^{\overline{\Sigma\cup\partial\Sigma}}\, O_S^\dagger\ \left(U_R\sigma^{\partial\Sigma}U_R^\dagger\otimes|L_{\overline{\Sigma}\leftrightarrow\overline{\Sigma}}\rangle\langle L_{\overline{\Sigma}\leftrightarrow\overline{\Sigma}}|\right)\right].\end{aligned}$$

But since $\sigma^{\partial\Sigma} = \frac{\mathbf{1}_R}{d_R}\otimes\rho_{\overline{R}}$ we have $U_R\sigma^{\partial\Sigma}U_R^\dagger = \sigma^{\partial\Sigma}$ and so the U_R dependence drops out of the above equation. Then

$$\langle L|U_R\, O_S\, \rho_P\, O_S^\dagger\, U_R^\dagger|L\rangle = \langle L|O_S\, \rho_P\, O_S^\dagger|L\rangle$$

and so $M(U_R : O_S)$ does not depend on U_R. Therefore, $\mathrm{CI}(R:S) = 0$, as claimed. □

The proposition is a technical way of saying that we can cancel out a U_R with a $U_R^\dagger$ in a GTN if there is a bridge (built out of tensor contractions) between them which is a maximally mixed state. Thus, the proposition specifies how maximally mixed states are sinks of causal flow in GTN's. In the special case of spacetime states, we see that initial and final states with maximally mixed subsystems act as sinks of causal flow since they provide a pathway for unitary cancellation.

Structure theorem

It is interesting to consider how causal relationships between regions of spacetime points affect the structure of correlation functions comprised of operator insertions at those points. A particular question along these lines is:

Suppose we have two spacetime points x and y, where x is a unitary region. If x does not causally influence y so that $\mathrm{CI}(x:y) = 0$, then what restrictions does this impose on the structure of spacetime correlation functions of the form $\langle L|A_x\, B_y\, \rho_P\, B_y^\dagger\, A_x^\dagger|L\rangle$ for a general tensor network, or as a special case $\langle \mathcal{P}\, A_x\, B_y\, \rho_i\, B_y^\dagger\, A_x^\dagger\rangle$ for a spacetime state?

To answer such a question, we need to utilize a formalism which organizes the data of spacetime correlation functions for spacetime states. This is called the "superdensity operator formalism" [56], which is reviewed in Appendix B. In short, a superdensity operator ϱ is a multilinear map taking operators to correlation functions (which evaluate to complex numbers). In our question of interest,

we will use a superdensity operator

$$\varrho : \mathcal{B}^*(\mathcal{H}_x) \otimes \mathcal{B}^*(\mathcal{H}_y) \otimes \mathcal{B}(\mathcal{H}_x) \otimes \mathcal{B}(\mathcal{H}_y) \longrightarrow \mathbb{C} \tag{4.25}$$

defined by

$$\varrho[A_x^\dagger\,, B_y^\dagger\,;\, A_x\,, B_y] := \langle L|A_x\, B_y\, \rho_P\, B_y^\dagger\, A_x^\dagger|L\rangle\,. \tag{4.26}$$

In the special case of spacetime states, the right-hand side of the above equation becomes $\langle \mathcal{P}\, A_x\, B_y\, \rho_i\, B_y^\dagger\, A_x^\dagger\rangle$.

As an example, in Figure 4.13(a), we depict ϱ diagramatically for a spacetime state with Trotterized time evolution. This tensor network can be more abstractly represented by the diagram in Figure 4.13(b). The diagram in Figure 4.13(b) is completely general for spacetime states, and simply expresses that the superdensity operator is a multilinear object which takes as input operators on $\mathcal{B}(\mathcal{H}_x) \otimes \mathcal{B}(\mathcal{H}_y)$ as well as dual operators on the dual space $\mathcal{B}^*(\mathcal{H}_x) \otimes \mathcal{B}^*(\mathcal{H}_y)$, and outputs a complex number.

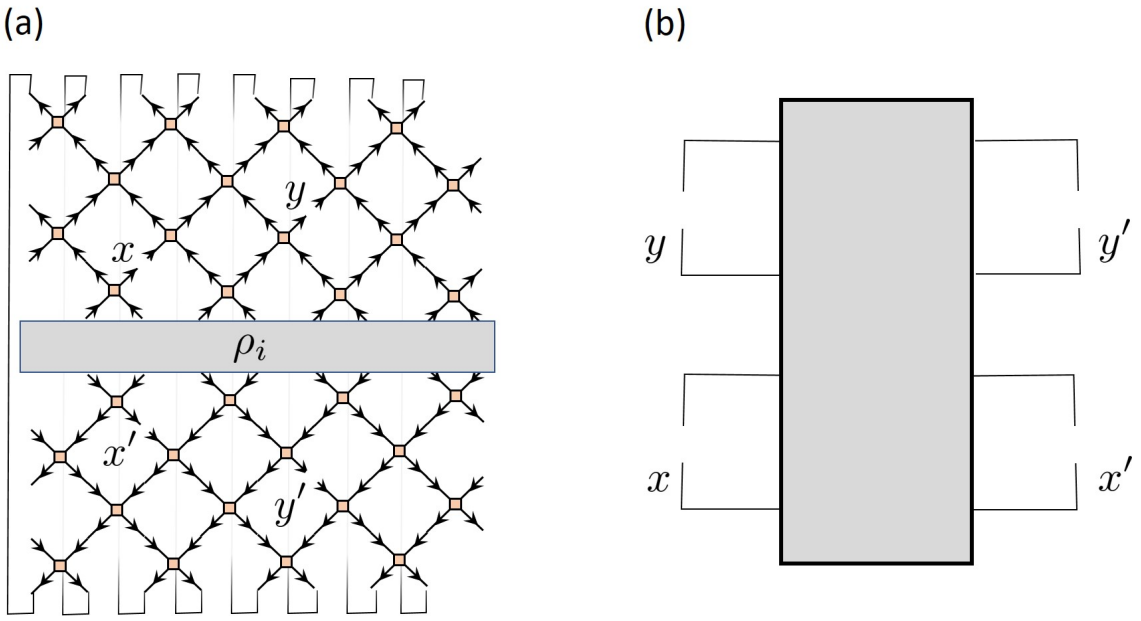


Figure 4.13: (a) A Trotterized network comprised of a spacetime state contracted with its Hermitian conjugate with initial state ρ_i, and broken legs to allow the insertion of operators into x and y as well as x' and y'. (b) A more abstract superdensity operator, allowing for operator insertions at x and y as well as x' and y'.

Using the superdensity setup, we prove the following structure theorem about general tensor networks:

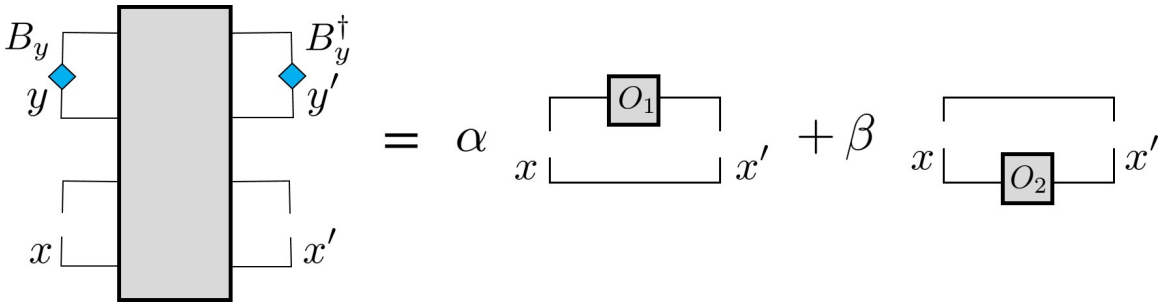


Figure 4.14: If $\mathrm{CI}(x : y) = 0$, then the superdensity operator with fixed y insertions can be written as a linear combination of a tensor network with a maximally mixed past, and a tensor network with a maximally mixed future.

Structure theorem: *If and only if* $\mathrm{CI}(x : y) = 0$*, then for fixed* B_y*, the spacetime correlation function* $\langle L|A_x\, B_y\, \rho_P\, B_y^\dagger\, A_x^\dagger|L\rangle$ *can be written as*

$$\langle L|A_x\, B_y\, \rho_P\, B_y^\dagger\, A_x^\dagger|L\rangle = \alpha\, \mathrm{tr}(O_1\, A_x\, A_x^\dagger) + \beta\, \mathrm{tr}(A_x^\dagger\, A_x\, O_2) \tag{4.27}$$

for all A_x*, where* α *and* β *are complex numbers and* O_1 *and* O_2 *are operators which are independent of* A_x*.*

Let us give a more intuitive interpretation of this theorem. First, we note that we can rewrite Eqn. (4.27) in terms of the superdensity operator ϱ given in Eqn. (4.26) as

$$\varrho[A_x^\dagger\,,\, B_y^\dagger\,;\, A_x\,,\, B_y] = \alpha\, \mathrm{tr}(O_1\, A_x\, A_x^\dagger) + \beta\, \mathrm{tr}(A_x^\dagger\, A_x\, O_2)\,. \tag{4.28}$$

This equivalence is depicted diagrammatically in Figure 4.14. We see from the figure a nice interpretation of the result: *the causal influence is trivial if and only if the two-site superdensity operator is a linear superposition of a tensor network with the final state being maximally mixed and another tensor network with the initial state being maximally mixed.* With this in mind, we prove the theorem.

Proof. For fixed B_y, we can generically write

$$\langle L|A_x\, B_y\, \rho_P\, B_y^\dagger\, A_x^\dagger|L\rangle = \sum_{i,j=0}^{d^2-1} K_{ij}\, \mathrm{tr}(M^i\, A_x\, M^{j\,\dagger}\, A_x^\dagger) \tag{4.29}$$

where K_{ij} are complex numbers, $\{M^i\}$ is a complete set of orthonormal operators satisfying $\mathrm{tr}(M^i M^{j\,\dagger}) = \delta_{ij}$, and $\mathcal{H}_x$ is a d-dimensional Hilbert space. Note that the K_{ij}'s depend on B_y, but not on A_x.

If $\mathrm{CI}(x : y) = 0$, then

$$\langle L|U_x\, B_y\, \rho_P\, B_y^\dagger\, U_x^\dagger|L\rangle = \langle L|\widetilde{U}_x\, B_y\, \rho_P\, B_y^\dagger\, \widetilde{U}_x^\dagger|L\rangle$$

for all unitaries U_x and $\widetilde{U}_x$. Therefore,

$$\sum_{i,j=0}^{d^2-1} K_{ij}\,\mathrm{tr}(M^i\, U_x\, M^{j\,\dagger}\, U_x^\dagger) = \sum_{i,j=0}^{d^2-1} K_{ij}\,\mathrm{tr}(M^i\, \widetilde{U}_x\, M^{j\,\dagger}\, \widetilde{U}_x^\dagger) \tag{4.30}$$

for all $U_x, \widetilde{U}_x$. In Eqn. (4.30) above, the terms for which either i or j is zero have vanishing trace. Also, the $i = j = 0$ term evaluates to one. Then Eqn. (4.30) simplifies to

$$\sum_{i,j=1}^{d^2-1} K_{ij}\,\mathrm{tr}(M^i\, U_x\, M^{j\,\dagger}\, U_x^\dagger) = \sum_{i,j=1}^{d^2-1} K_{ij}\,\mathrm{tr}(M^i\, \widetilde{U}_x\, M^{j\,\dagger}\, \widetilde{U}_x^\dagger) \tag{4.31}$$

where the sums now run from $i, j = 1, ..., d^2 - 1$. Letting $\widetilde{U}_x = \mathbf{1}$, we find that

$$\sum_{i,j=1}^{d^2-1} K_{ij}\,\mathrm{tr}(M^i\, U_x\, M^{j\,\dagger}\, U_x^\dagger) = \sum_{i=1}^{d^2-1} K_{ii} = C \tag{4.32}$$

for all U_x and some constant C. Using the Haar unitary integral

$$\int dU\, U^*_{nm} U_{k\ell} = \frac{1}{d}\,\delta_{nk}\delta_{m\ell} \tag{4.33}$$

we find

$$\int dU_x \sum_{i,j=1}^{d^2-1} K_{ij}\,\mathrm{tr}(M^i\, U_x\, M^{j\,\dagger}\, U_x^\dagger) = 0 \tag{4.34}$$

Therefore $C = 0$, implying that

$$\sum_{i,j=1}^{d^2-1} K_{ij}\,\mathrm{tr}(M^i\, U_x\, M^{j\,\dagger}\, U_x^\dagger) = 0 \tag{4.35}$$

for all U_x. Then we have

$$\left|\sum_{i,j=1}^{d^2-1} K_{ij}\,\mathrm{tr}(M^i\, U_x\, M^{j\,\dagger}\, U_x^\dagger)\right|^2 = 0 \tag{4.36}$$

for all U_x. Using the Haar unitary integral

$$\int dU\, U^*_{n_1 m_1} U^*_{n_2 m_2} U_{k_1 \ell_1} U_{k_2 \ell_2} = \frac{1}{d^2-1}\big[\delta_{n_1 k_1}\delta_{m_1 l_1}\delta_{n_2 k_2}\delta_{m_2 \ell_2} + \delta_{n_1 k_2}\delta_{m_1 \ell_2}\delta_{n_2 k_1}\delta_{m_2 \ell_1} - \frac{1}{d}\,\delta_{n_1 k_1}\delta_{n_2 k_2}\delta_{m_1 \ell_2}\delta_{m_2 \ell_1} - \frac{1}{d}\,\delta_{n_1 k_2}\delta_{n_2 k_1}\delta_{m_1 \ell_1}\delta_{m_2 \ell_2}\big] \tag{4.37}$$

we obtain

$$\int dU_x \left| \sum_{i,j=1}^{d^2-1} K_{ij}\,\mathrm{tr}(M^i\, U_x\, M^{j\,\dagger}\, U_x^\dagger) \right|^2 = \frac{1}{d^2-1} \sum_{i,j=1}^{d^2-1} |K_{ij}|^2 = 0 \tag{4.38}$$

so that $K_{ij} = 0$ for $i, j = 1, ..., d^2 - 1$.

It follows that

$$\langle L| A_x\, B_y\, \rho_P\, B_y^\dagger\, A_x^\dagger |L\rangle = \frac{1}{d}\, K_{00}\,\mathrm{tr}(\mathbf{1}\, A_x\, \mathbf{1}\, A_x^\dagger) + \sum_{i=0}^{d^2-1} \frac{K_{i0}}{\sqrt{d}}\,\mathrm{tr}(M^i\, A_x\, \mathbf{1}\, A_x^\dagger) + \sum_{j=0}^{d^2-1} \frac{K_{0j}}{\sqrt{d}}\,\mathrm{tr}(\mathbf{1}\, A_x\, M^{j\,\dagger}\, A_x^\dagger) \tag{4.39}$$

which we can repackage into the desired equation

$$\langle L| A_x\, B_y\, \rho_P\, B_y^\dagger\, A_x^\dagger |L\rangle = \alpha\,\mathrm{tr}(O_1\, A_x\, A_x^\dagger) + \beta\,\mathrm{tr}(A_x^\dagger\, A_x\, O_2)\,.$$

Conversely, if $\langle L| A_x\, B_y\, \rho_P\, B_y^\dagger\, A_x^\dagger |L\rangle = \alpha\,\mathrm{tr}(O_1\, A_x\, A_x^\dagger) + \beta\,\mathrm{tr}(A_x^\dagger\, A_x\, O_2)$, then $\langle L| U_x\, B_y\, \rho_i\, B_y^\dagger\, U_x^\dagger |L\rangle$ is independent of unitaries U_x which implies $\mathrm{CI}(x : y) = 0$. □

4.4 Nonlocality of the quantum causal influence

Quantum causal influence captures the ability of one subsystem of a tensor network to affect another subsystem. As remarked above, the quantum causal influence can behave in a peculiar way under the union of subsystems: in particular, we can have $\mathrm{CI}(R : S_1) = \mathrm{CI}(R : S_2) = 0$, whereas $\mathrm{CI}(R : S_1 \cup S_2) > 0$. In words, R does not influence either S_1 or S_2 individually, but R does influence their union $S_1 \cup S_2$. More modest cases are also possible – we may simply have that $\mathrm{CI}(R : S_1)$, $\mathrm{CI}(R : S_2)$ are close to zero whereas $\mathrm{CI}(R : S_1 \cup S_2) > 0$ is significantly larger than zero.

How do we interpret the above cases, especially in the context of spacetime? We will find that a core mechanism is the non-local encoding of information in spacetime. For instance, in the spacetime setting, perturbations at R can be non-locally encoded in the spacetime region $S_1 \cup S_2$, but not in the spacetime regions S_1 or S_2 alone. We can find natural examples in which S_1 and S_2 can be vastly separated in both space and time. Our analysis indicates that the non-local encoding of information in spacetime is a ubiquitous phenomenon.

A key tool for analyzing non-local quantum causal influence is the theory of quantum error correction codes. We begin by discussing quantum error correction, and show how quantum error correction codes allow us to construct examples of non-local causal influence. We then give a natural example of scrambling in a chaotic quantum many-body system. Finally, we explore the causal structure of quantum teleportation.

4.4.1 Quantum error correction codes

Nonlocal features of quantum causal influence are intimately related to quantum error correction codes. First, we briefly review quantum error correction codes, and quantum erasure codes in particular. A nice overview written for high energy physicists is given in [5].

There are many equivalent definitions of quantum error correction codes, so we choose one which is most convenient for our analysis here. Consider two Hilbert spaces $\mathcal{H}_A$, $\mathcal{H}_B$ with $\dim \mathcal{H}_A < \dim \mathcal{H}_B$. We may think of A as subsystem of B, so that $\mathcal{H}_B = \mathcal{H}_A \otimes \mathcal{H}_{\overline{A}}$. Intuitively, imagine we have a noisy quantum system B, and that we want to construct a protocol which protects the state of some subsystem A against our particular form of noise. The idea is to redundantly *encode* the state of the subsystem A into a state of the larger system B, in such a way that the larger encoded state is robust to our form of noise. Then we can subsequently *decode* the larger encoded state to obtain the original state on B.

Now we formalize this intuition. The space of density matrices on each Hilbert space $\mathcal{H}_A$, $\mathcal{H}_B$ are $\mathcal{S}(\mathcal{H}_A)$ and $\mathcal{S}(\mathcal{H}_B)$, respectively. Suppose we have three quantum channels (i.e., completely positive trace-preserving (CPTP) maps):

$$\mathcal{E} : \mathcal{S}(\mathcal{H}_A) \longrightarrow \mathcal{S}(\mathcal{H}_B) \tag{4.40}$$

$$\mathcal{N} : \mathcal{S}(\mathcal{H}_B) \longrightarrow \mathcal{S}(\mathcal{H}_B) \tag{4.41}$$

$$\mathcal{R} : \mathcal{S}(\mathcal{H}_B) \longrightarrow \mathcal{S}(\mathcal{H}_A)\,. \tag{4.42}$$

The channel $\mathcal{E}$ is the "encoding" channel, which maps density matrices on the subsystem A to density matrices on the larger system B. The channel $\mathcal{N}$ is the "noise" channel, which induces errors on density matrices on B. Finally, the channel $\mathcal{R}$ is the "recovery" channel, which decodes density matrices on B to density matrices on A. Then we have a quantum error correction code if

$$(\mathcal{R} \circ \mathcal{N} \circ \mathcal{E})(\rho) = \rho\,, \quad \text{for all } \rho \in \mathcal{S}(\mathcal{H}_A)\,. \tag{4.43}$$

In words, the above equation means that for all states on the subsystem A, applying the encoding channel $\mathcal{E}$, the noise channel $\mathcal{N}$, and finally the recovery channel $\mathcal{R}$ gives back the state that we

started with.

Notice that the description of a quantum error correction code depends on a specified form of noise, as provided by the given noise channel $\mathcal{N}$. There are many kinds of quantum error correction codes which protect against varied forms of noise. For our purposes, we will be most interested in noise which erases information. The corresponding form of quantum error correction code which is robust to erasure errors is called a quantum erasure code. These kinds of code are robust to an entire collection of noise channels $\{\mathcal{N}_S\}$, which we will define shortly.

To formally define a noise channel which causes erasure errors, consider again the Hilbert space $\mathcal{H}_B$, and let S be a subsystem of B with Hilbert space $\mathcal{H}_S$. Then let $\mathcal{N}_S$ be a channel taking $\mathcal{S}(\mathcal{H}_B) \to \mathcal{S}(\mathcal{H}_B)$ which erases all information on the subsystem S. The channel $\mathcal{N}_S$ is given by

$$\mathcal{N}_S(\rho) = \mathrm{tr}_S(\rho) \otimes \frac{\mathbf{1}_S}{\dim(\mathcal{H}_S)} \tag{4.44}$$

where $\mathbf{1}_S/\dim(\mathcal{H}_S)$ is the maximally mixed state on the subsystem S.

Now supposing that our system is a collection of qudits, let $|S|$ denote the number of qudits comprising the subsystem S. Equivalently, $|S| = \log_d(\dim(\mathcal{H}_S))$. Then a k–qudit quantum error correction code is given by quantum channels $\mathcal{E} : \mathcal{S}(\mathcal{H}_A) \to \mathcal{S}(\mathcal{H}_B)$, $\mathcal{R}_S : \mathcal{S}(\mathcal{H}_B) \to \mathcal{S}(\mathcal{H}_A)$ such that

$$(\mathcal{R}_S \circ \mathcal{N}_S \circ \mathcal{E})(\rho) = \rho\,, \quad \text{for all } S \text{ such that } |S| \leq k, \text{ and all } \rho \in \mathcal{S}(\mathcal{H}_A)\,. \tag{4.45}$$

In words, the k–qudit quantum error correction code can correct for the erasure of at most k qudits of B. Hence, the k–qudit quantum error correction code corrects for the entire collection of noise channels $\{\mathcal{N}_S\}_{|S|\leq k}$. Notice that the recovery channel $\mathcal{R}_S$ depends on the choice of subsystem S that is erased.

Now we provide an example of a 1–qutrit[6] quantum erasure code, called the "three qutrit code" [53, 20, 21, 5]. This code protects against the erasure of a single qutrit, among three qutrits. Let $\mathcal{H}_A = \mathrm{span}\{|0\rangle, |1\rangle, |2\rangle\}$ be the space of a single qutrit (so that $\dim \mathcal{H}_A = 3$) and let $\mathcal{H}_B$ be the space of three qutrits (so that $\dim \mathcal{H}_B = 27$). The encoding channel $\mathcal{E}$ is a unitary channel

$$\mathcal{E}(\rho) = U_{\text{encode}}\, \rho\, U_{\text{encode}}^{\dagger} \tag{4.46}$$

where U_{encode} acts by

$$U_{\text{encode}} \sum_{i=0}^{3} c_i\, |i\rangle = \sum_{i=0}^{3} c_i\, |\widetilde{i}\rangle \tag{4.47}$$

[6] A qutrit is a three-level system, i.e. a qudit with $d = 3$.

and

$$U_{\text{encode}}\left(|0\rangle \otimes |00\rangle\right) = |\widetilde{0}\rangle = \frac{1}{\sqrt{3}}\left(|000\rangle + |111\rangle + |222\rangle\right) \tag{4.48}$$

$$U_{\text{encode}}\left(|1\rangle \otimes |00\rangle\right) = |\widetilde{1}\rangle = \frac{1}{\sqrt{3}}\left(|012\rangle + |120\rangle + |201\rangle\right) \tag{4.49}$$

$$U_{\text{encode}}\left(|2\rangle \otimes |00\rangle\right) = |\widetilde{2}\rangle = \frac{1}{\sqrt{3}}\left(|021\rangle + |102\rangle + |210\rangle\right) . \tag{4.50}$$

Then the noise channels $\mathcal{N}_S$ have the form of Eqn. (4.44), where S is either $\{1\}$, $\{2\}$ or $\{3\}$, corresponding to erasing either the first, second or third qutrits. Then the recovery maps $\mathcal{R}_S$ are

$$\mathcal{R}_S(\rho) = \text{tr}_{g(\overline{S})\cup S}\Big(\left(U_{\overline{S}} \otimes \mathbf{1}_S\right)\, \rho\, \Big(U_{\overline{S}}^{\dagger} \otimes \mathbf{1}_S\Big)\Big) \tag{4.51}$$

where $\overline{S}$ can be $\{1,2\}$, $\{2,3\}$ or $\{1,3\}$, and $g(\{1,2\}) = \{1\}$, $g(\{2,3\}) = \{2\}$, and $g(\{1,3\}) = \{3\}$. Here $U_{\overline{S}}$ is a unitary that takes

$$U_{\overline{S}}|00\rangle = |00\rangle\,, \quad U_{\overline{S}}|11\rangle = |01\rangle\,, \quad U_{\overline{S}}|22\rangle = |02\rangle\,, \tag{4.52}$$

$$U_{\overline{S}}|01\rangle = |12\rangle\,, \quad U_{\overline{S}}|12\rangle = |10\rangle\,, \quad U_{\overline{S}}|20\rangle = |11\rangle\,, \tag{4.53}$$

$$U_{\overline{S}}|02\rangle = |21\rangle\,, \quad U_{\overline{S}}|10\rangle = |22\rangle\,, \quad U_{\overline{S}}|21\rangle = |20\rangle\,. \tag{4.54}$$

This code has the property that for any operator O on a qutrit state $|\psi\rangle$ in $\mathcal{H}_A$, we have the equivalences

$$U_{\text{encode}}\, O|\psi\rangle = (\widetilde{O}_{12} \otimes \mathbf{1}_3)|\widetilde{\psi}\rangle = (\widetilde{O}_{23} \otimes \mathbf{1}_1)|\widetilde{\psi}\rangle = (\widetilde{O}_{13} \otimes \mathbf{1}_2)|\widetilde{\psi}\rangle \tag{4.55}$$

for some operators $\widetilde{O}_{12}$, $\widetilde{O}_{23}$ and $\widetilde{O}_{13}$. This result expresses that the effect of any operation on the original state can be expressed by an equivalent operator on any two of the three qutrits of the encoded state.

Now let us consider the three qutrit code in spacetime. A diagram of a spacetime state which implements the three qutrit code is shown in Figure 4.15. The initial state of the qutrit we wish to encode is ρ_i, and the other two qutrits are initialized to $|0\rangle$. From Eqn. (4.55), it immediately follows that

$$\text{CI}(x : y_1) = \text{CI}(x : y_2) = \text{CI}(x : y_3) = 0\,. \tag{4.56}$$

However, we have

$$\text{CI}(x : y_1\, y_2) = \text{CI}(x : y_2\, y_3) = \text{CI}(x : y_1\, y_3) > 0 \tag{4.57}$$

and

$$\text{CI}(x : y_1\, y_2\, y_3) > 0 \quad \text{is maximal.} \tag{4.58}$$

By "maximal," we mean that $\text{CI}(x : y_1\, y_2\, y_3)$ is as large as possible. Taken together, Eqn's (4.56), (4.57)

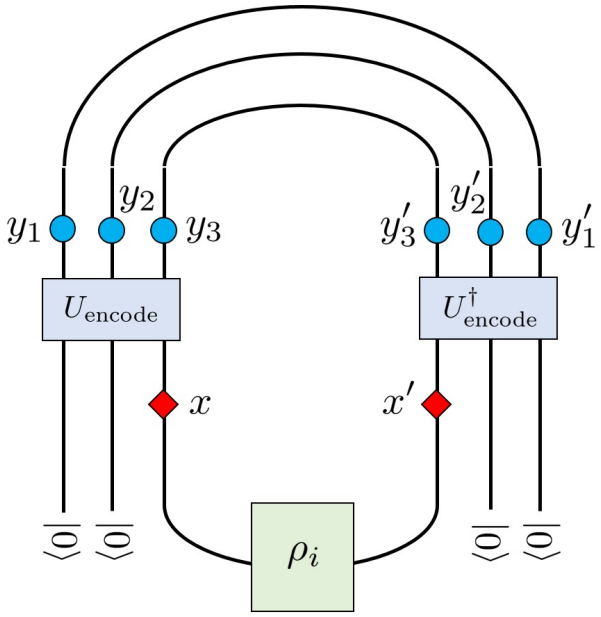


Figure 4.15: The spacetime state for the three qutrit code.

and (4.58) demonstrate how a peturbation at x can be non-locally encoded in space so that in the future the perturbation can be detected by any two (or more) qutrits, but not any single qutrit. More generally, all quantum erasure codes have non-local quantum causal influence between appropriate combinations of subsystems before and after the encoding.

4.4.2 Scrambling

While engineered quantum erasure codes provide examples of systems with nonlocal quantum causal influence, they are somewhat fine-tuned examples. However, *approximate* quantum error correction codes occur in various contexts in more natural systems. The simplest example is that of a chaotic quantum many-body system which scrambles information. The scrambling of information is ubiquitous in nature, since most all physical systems exhibit many-body chaos. However, the most extreme examples of scrambling systems are black holes, which are the fastest scramblers in nature [228, 138, 229, 162]. It was in the context of black holes that scrambling was first explored. We will not focus on any particular scrambling system, but instead use generic features of scrambling for our analysis.

There are many definitions of information scrambling in the literature. (See, for instance, [103, 95, 40, 162]. For a short review of diagnostics of scrambling at infinite temperature, see Appendix A of [55]). Suppose we have a system with a large number N of sites, and that the initial state of the system is ρ_i. If the time evolution $U(t)$ of the system is chaotic, then the scrambling time t_{scr} is

the smallest time such that for any subsystem a of $\mathcal{O}(1)$ size and any subsystem B of size $N/2+1$, there exists a quantum channel $\mathcal{R}_{B\to a}$ such that

$$\mathcal{R}_{B\to a}\left[\mathrm{tr}_{\overline{B}}\left(U(t_{\mathrm{scr}})\,\rho_i\,U^\dagger(t_{\mathrm{scr}})\right)\right]\approx \mathrm{tr}_{\overline{a}}(\rho_i)\,. \tag{4.59}$$

In other words, any $\mathcal{O}(1)$–sized subsystem can be approximately recovered from just over half of the state after a scrambling time. In this sense, unitary evolution for a scrambling time in a chaotic quantum system creates an (approximate) erasure code for initial subsystems of $\mathcal{O}(1)$ size. The length of the scrambling time t_{scr} depends on the types of interactions in the system, and typically scales with the number of degrees of freedom N either polynomially in N (if the interactions are geometrically local) or logarithmically in N (for instance, if the interactions are k-local for $k\sim\mathcal{O}(1)$).

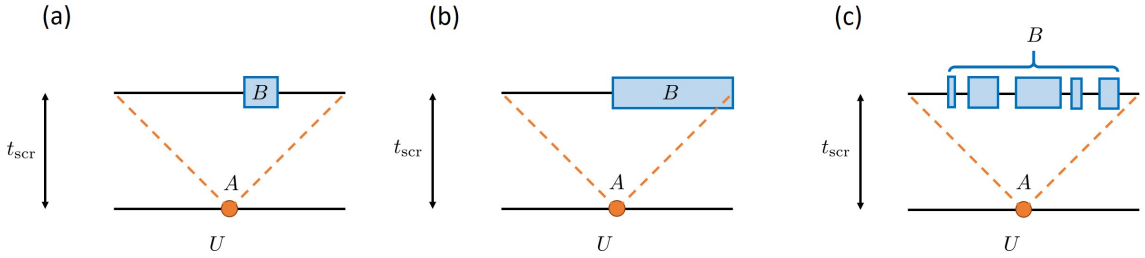


Figure 4.16: A perturbation is made at some initial time, which then spreads out over a scrambling time t_{scr}, inside a cone (shown as dotted orange lines) bounded by the butterfly velocity v_B [229, 216, 172]. Here, time runs from bottom to top. (a) A perturbation at a barely causally influences the subregion B, since B is less than half of the system size. (b) and (c) A perturbation at a strong causally influences B, if B is greater than half of the system size. The two figures illustrate the cases when region B is a contiguous spatial region or the union of many contiguous regions. The conclusion applies to both cases.

Now consider Figure 4.16 below, which shows a system scrambling (time goes from bottom to top). In Figure 4.16(a), we see that the causal influence $\mathrm{CI}(a:B)\approx 0$ since B is less than half of the system size. However, in Figure 4.16(b), the causal influence $\mathrm{CI}(a:B)$ is sizeable, since B is greater than half of the system size. Finally, in Figure 4.16(c), we have that $\mathrm{CI}(a:B)$ is sizeable since B is greater than half the system size, even though B is not a spatially contiguous subregion.

We emphasize that any $\mathcal{O}(1)$–sized region at the initial time will have a negligible causal influence with any $\mathcal{O}(1)$–sized region in the future after the scrambling time, and conversely as well. Relatedly, from the point of view of quantum causal influence, local subsystems in the present will appear approximately spacelike separated with local subsystems in the future after the entire system has thermalized. Indeed, local notions of time disappear after a system thermalizes – local properties of the past only weakly influence local properties of the far future.

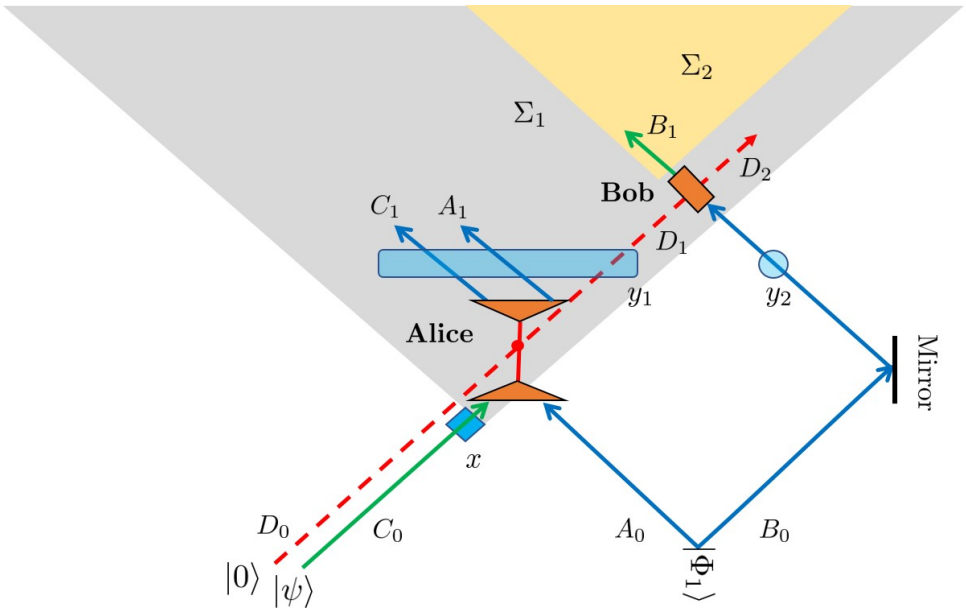


Figure 4.17: A spacetime diagram of the quantum teleportation protocol. Space runs horizontally, and time runs vertically from bottom to top. The consequence of teleportation is that the future of x (the teleportee C_0 right before teleportation happens) shrinks from the ordinary future light cone Σ_1 to a subset Σ_2 (the future of the point where teleportation is finished). Note that the EPR pair of A_0, B_0 is still outside the future of x, as is expected by microscopic causality.

4.4.3 Quantum teleportation

Now we explore how quantum teleportation [19] nonlocally encodes information in spacetime. Quantum teleportation can be described by a tensor network, as shown in Figure 4.17 below. In the Figure, space runs horizontally, and time runs vertically from bottom to top. Let us walk through the protocol step by step.

Consider the setup in Figure 4.17. We suppose that all of the states involved are encoded into photons (say, in their polarization degrees of freedom), which have lightlike trajectories. We start with a Bell state (i.e., an EPR pair of two qubits)

$$|\Phi_1\rangle_{AB} = \frac{1}{\sqrt{2}}\Big(|0\rangle_{A_0}|0\rangle_{B_0} + |1\rangle_{A_0}|1\rangle_{B_0}\Big),$$

a state $|\psi\rangle_C$ that we wish to teleport (i.e., the teleportee), and an ancillary qubit $|0\rangle_{D_0}$. One qubit of the Bell pair, as well as the joint state $|\psi\rangle_{C_0} \otimes |0\rangle_{D_0}$, are fed into a "teleporter" owned by Alice,

denoted in the Figure by an orange triangle. Letting

$$|\Phi_1\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle|0\rangle + |1\rangle|1\rangle\right)$$
$$|\Phi_2\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle|0\rangle - |1\rangle|1\rangle\right)$$
$$|\Phi_3\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle|1\rangle + |1\rangle|0\rangle\right)$$
$$|\Phi_4\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle|1\rangle - |1\rangle|0\rangle\right)$$

denote the Bell states (which are the basis vectors of the Bell basis), the teleporter implements the unitary

$$U_{1,\,A_0C_0D_0} = \sum_{j=1}^{4} |\Phi_j\rangle\langle\Phi_j|_{A_0C_0} \otimes |j\rangle_{D_0}\langle 0| + \cdots \tag{4.60}$$

which couples the A_0C_0 state in the Bell basis to the ancillary qubit D_0. The teleporter then outputs the A_0, C_0 and D_0 subsystems, now denoted A_1, C_1 and D_1. The A_1 and C_1 subsystems are discarded, while the D_1 subsystem goes on to Bob. In the meantime, the B subsystem of the Bell state is directed towards Bob with a mirror. When Bob receives B_0 and D_1, he applies the unitary

$$U_{2,\,B_0D_1} = \sum_{j=1}^{4} U_{B_0,\,j} \otimes |j\rangle_{D_1}\langle j| \tag{4.61}$$

which is denoted by an orange box. The unitary $U_{2,\,B_0D_1}$ applies the unitary $U_{B_0,\,j}$ to the B_0 subsystem, controlled by the state of D_1. The output of the B_1 subsystem will be the original state of A_0, namely $|\psi\rangle$, which has successfully been teleported to Bob.

Now we analyze the causal future of the initial state $|\psi\rangle_{C_0}$, denoted by the initial subsystem C_0. Apparently in the protocol, the future of C_0 is B_1. In fact it can be checked that

$$\mathrm{CI}(C_0 : B_1) > 0 \quad \text{is maximal.} \tag{4.62}$$

(As before, "maximal" means that the quantum causal influence is as large as possible.) However, denoting $y_1 = A_1 \cup C_1 \cup D_1$, we also have that

$$\mathrm{CI}(C_0 : y_1) = 0 \tag{4.63}$$

and thus C_0 is spacelike separated from $A_1 \cup C_1 \cup D_1$ and any subset thereof. We also have that

$$\mathrm{CI}(C_0 : y_2) = 0 \tag{4.64}$$

which means that C_0 is spacelike separated from B_0. This is consistent with the causal structure

which Figure 4.17 inherits from Minkowski space.

In summary, even though it appears that C_0 should be able to influence its whole future light cone Σ_1, it can only causally influence the subset Σ_2. In words:

$$C_0 \textit{ cannot influence any local region while it is being teleported.} \tag{4.65}$$

Even though $\mathrm{CI}(C_0 : y_1) = 0$ and $\mathrm{CI}(C_0 : y_2) = 0$, we still have that

$$\mathrm{CI}(C_0 : y_1 \cup y_2) > 0 \tag{4.66}$$

which is in fact maximal. Thus, while the state of C_0 is not encoded in either $A_1 \cup C_1 \cup D_1$ alone or B_0 alone, C_0 *is* encoded in $(A_1 \cup C_1 \cup D_1) \cup B_0$.

From another point of view, the example of quantum teleportation shows again that the causal structure depends on properties of the initial state, in this case the presence of the Bell state $|\Phi_1\rangle$. Fine-tuning of the initial state can only reduce the size of the putative future of spatial subregions. Said simply, special initial states can *remove* regions from the future.

4.5 Quantum gravity examples

In this section, we discuss several examples in holography as well as models of black holes for which quantum causal influence is a useful measure. In Section 4.5.1 we discuss holographic tensor networks and show how the causal influence correctly reproduces the bulk causal structure. In Section 4.5.2 we discuss the causal structure in the Horowitz-Maldacena final state projection model of black hole.

4.5.1 Holographic tensor networks

Holographic states

An interesting instantiation of quantum error correction codes in high energy physics is in holographic systems, and specifically AdS-CFT [161, 252]. In AdS-CFT, there is a duality between a $(d+1)$–dimensional quantum gravity theory in AdS space (i.e., the bulk theory), and a d–dimensional conformal field theory which lives on a space isomorphic to the conformal boundary of AdS (i.e., the boundary theory). There is necessarily an intricate relationship between degrees of freedom in the bulk and the boundary, and in fact, low-energy degrees of freedom in the bulk are non-locally encoded in the boundary theory in the form of a quantum erasure code [5]. In particular, a local low energy operator acting in the bulk can be reconstructed from many distinct spatial regions in the boundary theory.

The quantum error correction property of AdS-CFT duality can be captured in toy models known as holographic tensor networks [195, 102]. We will consider quantum causal influence in holographic

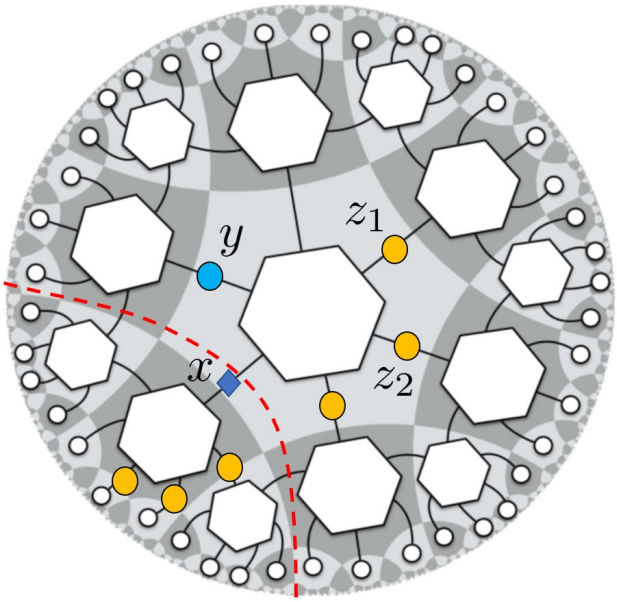


Figure 4.18: A hyperbolic perfect tensor network in which y is not in the future of x, as U_x can be pushed to the boundary only using the circuit on the left-side of the geodesic (red dashed line). A multi-site region such as z (6 yellow links) can be causally influenced by x since there is no way to push U_x to boundary without passing through z. This type of non-local causal influence is characteristic of holographic systems, and does not occur, for example, on a fixed Cauchy slice for a quantum field theory. This Figure is adapted from [195].

tensor networks, and study its relation to the bulk causal structure.

As an example, we consider the hyperbolic perfect tensor network state defined by the work of Pastawski et al. [195], shown in Figure 4.18. (All the discussion in the following also applies to the random tensor networks in large bond dimension limit proposed in Ref. [102].) A perfect tensor network state represents a many-body quantum state of the boundary legs, with its wavefunction defined by contracting perfect tensors. Each perfect tensor is a rank $2n$ tensor $T_{a_1 \cdots a_{2n}}$ such that the bipartition of its indices into sets A and A^c with $|A| \leq |A^c|$ defines an isometry from A to A^c up to a normalization constant. In Figure 4.18, we have considered the case $n = 3$, and the only uncontracted legs of the tensor network state live near the boundary of a hyperbolic disk.[7] Thus, the tensor network state in Figure 4.18 forms a so-called "holographic state", which we denote by $|\Psi\rangle$. The essential feature of this state is that if we break open any bulk leg (i.e., a non-boundary leg) of the tensor network state and stick in an operator, we can (non-uniquely) push it through the isometries out to the boundary, and so rewrite the operator as a "boundary" operator. This mimics the AdS-CFT correspondence: operators inserted into the bulk can be rewritten non-uniquely as

[7]The hyperbolic disk has infinite area. We have imposed a radial cutoff so that it has finite area. The uncontracted tensors live on the radial cutoff.

operators applied to some boundary state.

Suppose we break open two links x and y of $|\Psi\rangle$ to insert operators. If we insert operators A_x and B_y into x and y, respectively, we denote the resulting state by $|\Psi[A_x, B_y]\rangle$. While we can express $\langle\Psi|\Psi\rangle$ as

$$\langle\Psi|\Psi\rangle = \langle L|\rho_P|L\rangle \tag{4.67}$$

and similarly express $\langle\Psi[A_x, B_y]|\Psi[A_x, B_y]\rangle$ as

$$\langle\Psi[A_x, B_y]|\Psi[A_x, B_y]\rangle = \langle L|B_y^\dagger\, A_x^\dagger\, \rho_P\, A_x\, B_y|L\rangle \tag{4.68}$$

As per our definition of GTN's, ρ_P is the tensor product of vertex tensors (where we choose the boundary vertex tensors to be identity operators) and $|L\rangle$ is the link state comprised of EPR pairs.

We usually speak of the causal structure of a *fully contracted* tensor network (such as the one which computes $\langle\Psi|\Psi\rangle$), but here is it convenient to speak of the causal structure of the *state* $|\Psi\rangle$ (which has uncontracted legs). This is purely for terminological convenience – we always have in mind computing expectation values like $\langle\Psi[A_x, B_y]|\Psi[A_x, B_y]\rangle$. So when we say "the causal structure of $|\Psi\rangle$," we mean "the causal structure of $|\Psi\rangle$ contracted with itself."

With our terminology defined, we now discuss quantum causal influence for the holographic state $|\Psi\rangle$ in Figure 4.18. For any two links x and y, as long as they can be separated by a geodesic line on the hyperbolic disk, a unitary U_x inserted at x can be pushed to the boundary without using the y link, so that y is not in the causal future of x. Examination of the holographic state reveals that any two links can be separated by a geodesic line on the holographic disk, and therefore

$$\langle\Psi[U_x, O_y]|\Psi[U_x, O_y]\rangle \text{ is independent of } U_x\,,$$

$$\langle\Psi[O_x, U_y]|\Psi[O_x, U_y]\rangle \text{ is independent of } U_y\,.$$

It follows that $\mathrm{CI}(x : y) = \mathrm{CI}(y : x) = 0$, so that any two links x and y in network are "spacelike separated."

Our operational definition of causal structure explains why perfect tensor network states should be understood as *spatial* tensor network states even if their isometry conditions allow one to push operators around. Indeed, the perfect tensor network state is an example where all small enough regions are spacelike separated, but larger size regions may be causally dependent (i.e., if such regions cannot be separated by a geodesic line on the hyperbolic disk). For example, in Figure 4.18, x does not influence y, or any of the yellow points $z_1,\ z_2, ...$ individually. Furthermore, x does not influence the pair $z_1 \cup z_2$, since x can be separated from $z_1 \cup z_2$ by a geodesic on the hyperbolic disk. However, x does causally influence the subregion that is the union of *all* the yellow dots, since there is no way to push operators at x to the boundary without overlapping with this subregion.

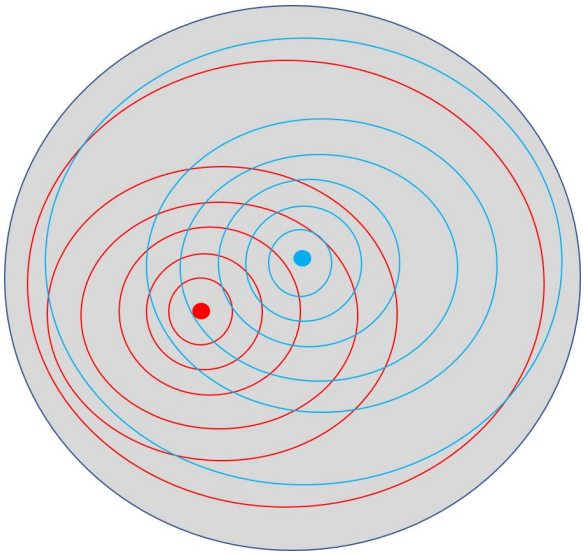


Figure 4.19: Exotic quantum Cauchy slices of the HaPPY code holographic state.

Exotic quantum Cauchy slicings of holographic states

In Figure 4.19, we provide some further illustration of the unconventional causal structure in the holographic tensor network state $|\Psi\rangle$. In the Figure, the tensor network has been abstracted to a gray disk. Consider a set of concentric rings on the hyperbolic disk (the red circles in Figure 4.19). Each red ring defines a subsystem into which we can insert operators (i.e., corresponding to inserting operators into all links that the red ring cuts through). Then we find that the subsystem corresponding to a red ring R_1 causally influences a subsystem corresponding to any bigger red ring R_2 that encloses R_1. The influence is in fact maximal since there is an isometry from R_1 to R_2. Indeed, a pair of subsystems corresponding to a pair of concentric red rings has timelike separation with respect to the QCI. Therefore, the concentric red rings are quantum analogs of Cauchy slicing of the holographic state. We will not attempt to define quantum Cauchy slices in full generality, but will comment further in Section 4.7. The concentric ring subsystems provide an exotic causal structure where the radial direction acts as time – this is dramatically different from more familiar examples. For instance, this exotic causal structure does not admit light cones.

There are many possible, incompatible Cauchy slicings of the holographic state, corresponds to different sets of concentric rings. For instance, in Figure 4.19, the set of blue rings is another Cauchy slicing with the same property as the red rings. However, the red and blue Cauchy slicings are not compatible with each other, since the subsystem corresponding to some red ring may not be time-like separated with the subsystem corresponding to some blue ring. This situation never occurs with standard Cauchy slicings of a classical spacetime with Lorentzian signature. The exotic Cauchy slicing found here is essential for bulk reconstruction to be consistent with the homogeneity

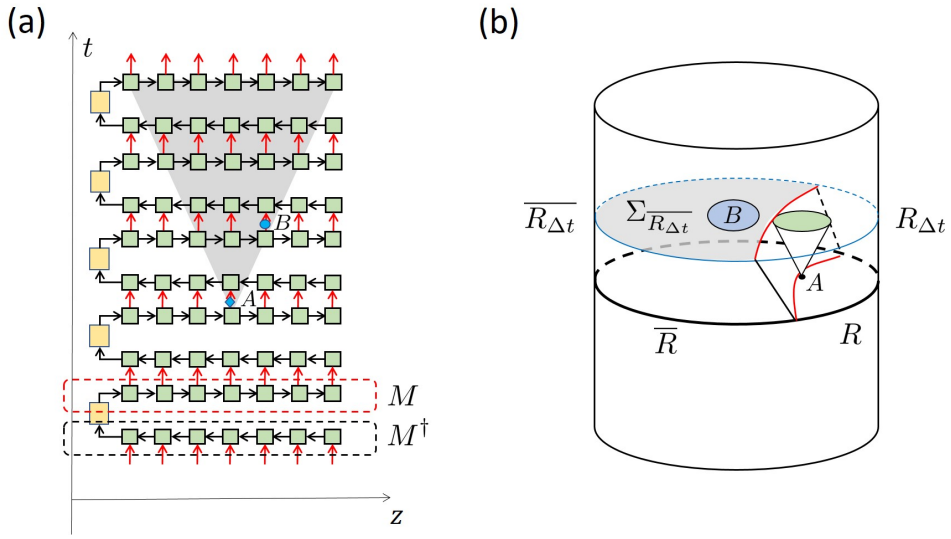


Figure 4.20: Causal influence for two bulk regions at different times. (a) Bulk time evolution is defined by pulling back the boundary time evolution using a holographic tensor network (see text). (b) An illustration in a (2+1)d bulk. An operator on a small region A can be reconstructed in a boundary region R, which evolves into a slightly bigger region $R_{\Delta t}$ after a short time Δt. Therefore, all operators in the complement entanglement wedge $\Sigma_{\overline{R_{\Delta t}}}$ still commute with the operator at A, which proves that A has no causal influence on any region $B \subset \Sigma_{\overline{R_{\Delta t}}}$.

of the bulk (i.e., there is no preferred point or preferred direction on the hyperbolic disk), which is the key difference between perfect tensor network states (as well as random tensor network states) and earlier proposals of MERA [248, 239].

In summary, the nonlocality of quantum causal influence characterizes how bulk locality is consistent with bulk reconstruction, as a consequence of the bulk's quantum error correction properties. The bulk contains a redundant encoding of boundary quantum information as is evident in the Cauchy surface structure, but this redundancy is invisible for local observers.

Explicit time direction

The discussion above can be further generalized by considering an *explicit* time direction via unitary evolution of the holographic state $|\Psi\rangle$. This section will be more technical, and we refer readers to [102] and [208] for details. To describe the bulk dynamics of low-energy degrees of freedom, consider the holographic mapping (or holographic code) defined by a random tensor network with bulk and boundary indices. Such a network defines a linear map

$$M : \mathcal{H}_{\text{bulk}} \longrightarrow \mathcal{H}_{\text{bdy}} \tag{4.69}$$

from low-energy bulk degrees of freedom to the boundary. The map is an isometry when the included bulk degrees of freedom have low enough dimension [102]. We call the image of $\mathcal{H}_{\text{bulk}}$ under M the "code subspace" of $\mathcal{H}_{\text{bdy}}$, which we denote by $\mathcal{H}_{\text{code}} := M(\mathcal{H}_{\text{bulk}})$. Indeed, we have $\mathcal{H}_{\text{code}} \subset \mathcal{H}_{\text{bdy}}$.

In Figure 4.20, we illustrate such a mapping M in the red dashed box. (The drawing is for a (1+1)d bulk for convenience, but the setup applies to arbitrary dimensions.) With this mapping M, boundary time evolution can be "pulled back" to the bulk and to define the bulk time evolution. With the boundary time evolution operator $e^{-iH\Delta t}$ for small Δt, the bulk time evolution is given by $U_{\text{bulk}} = Me^{-iH\Delta t}M^\dagger$ (which is unitary in the code subspace if the boundary time evolution preserves the code subspace). Naïvely, this time evolution is very nonlocal in the bulk, since we have to map all operators to (non-local) operators on the boundary and then map them back after the time evolution. However, the quantum error correction properties and locality of boundary dynamics actually guarantees that the bulk evolution also has a local causal structure [208].

The basic idea is illustrated in Figure 4.20(b). An operator ϕ_A in a small bulk region A can be reconstructed in a boundary region R. Then due to boundary locality, the operator ϕ_A at a slightly later time Δt will live in a slightly larger region $R_{\Delta t}$. Consequently, all bulk operators in the entanglement wedge $\Sigma_{\overline{R_{\Delta t}}}$ of the complement $\overline{R_{\Delta t}}$ still commute with the (slightly) Heisenberg-evolved operator ϕ_A. This implies that for any bulk region $B \in \Sigma_{\overline{R_{\Delta t}}}$, we have $\text{CI}(A : B) = 0$. Since the reconstruction can be done on different boundary regions R, the argument applies to each possible R. As long as B is included in the complement of the entanglement wedge of some $R_{\Delta t}$, there will be no causal influence from B to R or $R_{\Delta t}$.

If we consider regions B that are infinitesimal disks on the Δt time slice, any B that is outside the domain of support of A at time Δt is not influenced by A. In Figure 4.20(b), we see that any small blue disk B which does not intersect the green disc (which is the domain of support of A at time Δt) is spacelike separated from the green disc. Therefore, we recover the ordinary causal structure expected for the bulk theory. The boundary of the domain of support of A at time Δt (i.e., the green region in the Figure) defines an upper bound of the bulk speed of light [208].

Now, if we consider more generic regions B that are not small discs, the influence of B with the domain of support of A at time Δt can be nontrivial even there is no intersection between these regions. For example, if B is a ring enclosing the domain of support of A at time Δt, the causal influence will be nontrivial, since the reconstruction of operators in boundary region $R_{\Delta t}$ must use a bulk region that overlaps with B. This is similar to the exotic quantum Cauchy surfaces discussed above for the equal-time case.

4.5.2 Black hole final state

In Section 4.3.4 we discussed how for spacetime states, the causal influence depends in a similar manner on both the initial and final states. The initial and final states act as boundary conditions for the spacetime state. An interesting example of a nontrivial final state is the final state projection

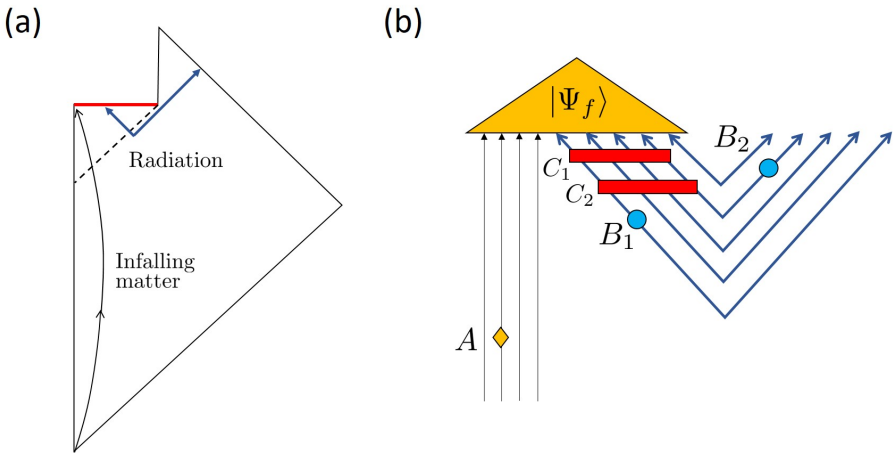


Figure 4.21: (a) The Penrose diagram of a Schwarzchild black hole with infalling matter (black curve) and infalling and outgoing Hawking radiation (blue lines). The red line represents singularity. (b) The Horowitz-Maldacena final state projection model, with the infalling matter and infalling Hawking qubits projected to a pure state at the singularity. A, B_1, B_2 are small regions of infalling matter, infalling radiation and outgoing radiation, respectively. C_1, C_2 are bigger regions of the infalling radiation, for which artifacts of the final state projection are detectable.

model of the black hole singularity, proposed by Horowitz and Maldacena [108]. This model is illustrated in Figure 4.21. There is infalling matter (the black curve), as well as infalling and outgoing radiation. The outgoing radiation is Hawking radiation, and the infalling radiation can be thought of as the "Hawking partner" of the Hawking radiation [93]. The outgoing and infalling radiation form a maximally entangled state.[8] The hypothesis is that there is a (post-selected) final state at the singularity, and that all matter and radiation falling into the singularity are projected onto that fixed final state. Such a projection will generically violate unitarity, but when the final state is chosen properly, the information content of infalling matter is mapped unitarily to outgoing radiation. This is much like quantum teleportation: a desired state (infalling matter) and half of a maximally entangled state (infalling radiation) are jointly measured (projection onto black hole final state), and the desired state is teleported to the other half of the maximally entangled state (outgoing Hawking radiation).

For example, suppose the black hole final state $|\Psi_f\rangle$ is a Haar random state. The state $|\Psi_f\rangle$ lives on the Hilbert space $\mathcal{H}_M \otimes \mathcal{H}_R$ where $\mathcal{H}_M$ is the Hilbert space of the infalling matter and $\mathcal{H}_R$ is the Hilbert space of the infalling radiation. Then $|\Psi_f\rangle$ has the form

$$|\Psi_f\rangle = \sum_{i,j} c_{ij} |i\rangle_M \otimes |j\rangle_R \,.$$

[8]The situation is more complicated when the entanglement is not maximal, but we will not discuss this here.

By dualizing $\mathcal{H}_M$ to $\mathcal{H}_M^*$ (and thus $|i\rangle_M \to \langle i|_M$), we can re-express $|\Psi_f\rangle$ as a mapping $V_{\Psi_f} : \mathcal{H}_M \to \mathcal{H}_R$ from the infalling matter to the infalling radiation as

$$V_{\Psi_f} = \sum_{i,j} c_{ij} |j\rangle_R \langle i|_M \,. \tag{4.70}$$

Indeed, if $|\Psi_f\rangle$ is Haar random and $\dim \mathcal{H}_M < \dim \mathcal{H}_I$ (i.e., the Hilbert space dimension of infalling matter is smaller than that of the infalling radiation), the mapping V_{Ψ_f} is an isometry (up to exponentially small corrections in the number of degrees of freedom). Since the infalling and outgoing radiation are maximally entangled, the net effect is that the information in the infalling matter is preserved in the outgoing Hawking radiation, and the unitarity of the quantum mechanics of the exterior region is restored (up to exponentially small corrections in the number of degrees of freedom) [155, 244].

Since the final state plays the role of an (approximately) isometric mapping from the infalling matter to infalling radiation, unitary operations at A have nontrivial causal influence on both the infalling and outgoing radiation. However, when $|\Psi_f\rangle$ is a Haar random state, its corresponding (approximately) isometric mapping V_{Ψ_f} is a *random* (approximate) isometry, and so the quantum causal influence of A is highly nonlocal. Accordingly, the quantum causal influence of A on any small subsystem such as B_1, B_2 nearly vanishes. The influence due to A is only nontrivial on large enough regions such as C_1, C_2. This is the same phenomenon as the nonlocal causal influence we observed in quantum error correction codes (see [244] for a related discussion).

The near vanishing of both $\mathrm{CI}(A : B_1)$ and $\mathrm{CI}(A : B_2)$ is consistent with the causal structure in the Penrose diagram in Figure 4.21(a), since the Penrose diagram suggests that A is spacelike separated from both B_1 and B_2. When we consider the quantum causal influence from A to larger regions such as C_1 and C_2, we can observe abnormal causal structure that is at odds with the Penrose diagram. For example, we have $\mathrm{CI}(A : C_1) \neq 0$ and $\mathrm{CI}(A : C_2) \neq 0$. Furthermore, the quantum causal influence between pairs of large regions also unveils abnormal quantum causal influence, for instance $\mathrm{CI}(C_1, C_2) \neq 0$ and $\mathrm{CI}(C_2, C_1) = 0$, which means that the time ordering of big regions C_1, C_2 for infalling radiation has been reversed due to the final state projection. The reverse time ordering is consistent with the observation that measurements involving large regions can detect violations of standard (non-post-selected) quantum mechanics [84, 35].

4.6 Averaged quantum causal influence and spacetime quantum entropies

In this section, we perform a more quantitative analysis of the averaged quantum causal influence (aQCI) and discuss its relation to spacetime quantum entropies in the superdensity operator formalism. We also use our results to analyze the quantum causal structure of evolving quantum spin chains as

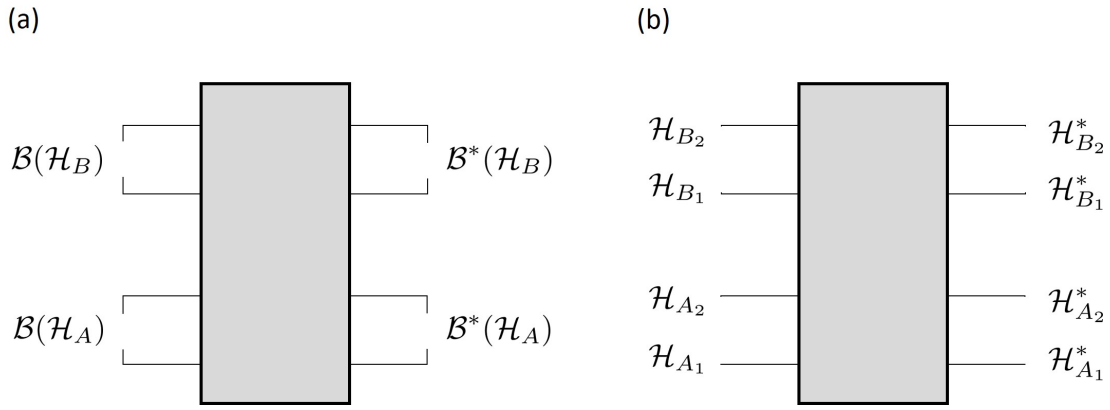


Figure 4.22: (a) A diagrammatic representation of the R tensor, as per Eqn. (4.73). (b) An equivalent diagrammatic representation of the R tensor, where tensor legs have been relabelled by the isomorphism of Hilbert spaces as per Eqn. (4.74).

well as stabilizer tensor networks.

4.6.1 Relation to spacetime quantum Rényi entropies

In Section 4.2 we presented two measures of quantum causal influence. The aQCI defined in Eqn. (4.13) is easier to compute than the mQCI defined in Eqn. (4.12). For the aQCI, we can in fact explicitly carry out the average over U_A and O_B. The aQCI can be written as

$$\overline{\mathrm{CI}}(A:B) = \int_{\|O_B\|_2^2=1} dO_B \int dU_A \; |M(U_A:O_B)|^2 - \int_{\|O_B\|_2^2=1} dO_B \left| \int dU_A \, M(U_A:O_B) \right|^2 \quad (4.71)$$

To obtain a more explicit expression of $\overline{\mathrm{CI}}(A:B)$, we define an orthonormal basis $\{|n_A\rangle\}$ of $\mathcal{H}_A$, and similarly $\{|n_B\rangle\}$ of $\mathcal{H}_B$. Since $M(U_A:O_B)$ is quadratic in U_A and in O_B, we can define a tensor $R^{k\ell\gamma\delta}_{nm\alpha\beta}$, such that

$$M(U_A:O_B) = U_A^{nm*} O_B^{\alpha\beta *} R^{k\ell\gamma\delta}_{nm\alpha\beta} U_A^{k\ell} O_B^{\gamma\delta} \quad (4.72)$$

Here, R can be thought of as a positive semidefinite operator mapping

$$R : \mathcal{B}(\mathcal{H}_A) \otimes \mathcal{B}^*(\mathcal{H}_A) \otimes \mathcal{B}(\mathcal{H}_B) \otimes \mathcal{B}^*(\mathcal{H}_B) \longrightarrow \mathbb{C} \quad (4.73)$$

which is depicted in Figure 4.22(a). Since $\mathcal{B}(\mathcal{H}_A) \simeq \mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}$ where $\mathcal{H}_{A_1} \simeq \mathcal{H}_A$ and $\mathcal{H}_{A_2} \simeq \mathcal{H}_A$ (and similarly for $\mathcal{B}^*(\mathcal{H}_A)$, $\mathcal{B}(\mathcal{H}_B)$, $\mathcal{B}^*(\mathcal{H}_B)$), we can treat R as a mapping

$$R : (\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}) \otimes (\mathcal{H}^*_{A_1} \otimes \mathcal{H}^*_{A_2}) \otimes (\mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2}) \otimes (\mathcal{H}^*_{B_1} \otimes \mathcal{H}^*_{B_2}) \longrightarrow \mathbb{C} \tag{4.74}$$

which is depicted in Figure 4.22(b). If A and B are each unitary regions, with proper normalization, R for a spacetime tensor network is an example of a superdensity operator [56]. The Haar average of U_A and O_B can be carried out with the following identities:

$$\begin{aligned}
\int dU\, U^*_{nm} U_{k\ell} &= \frac{1}{d_A}\, \delta_{nk}\delta_{m\ell} \\
\int_{\|O\|_2^2=1} dO\, O^*_{\alpha\beta} O_{\gamma\delta} &= \frac{1}{d_B^2}\, \delta_{\alpha\gamma}\delta_{\beta\delta} \\
\int dU\, U^*_{n_1 m_1} U^*_{n_2 m_2} U_{k_1 \ell_1} U_{k_2 \ell_2} &= \frac{1}{d_A^2 - 1} \left[\delta_{n_1 k_1}\delta_{m_1 l_1}\delta_{n_2 k_2}\delta_{m_2 \ell_2} + \delta_{n_1 k_2}\delta_{m_1 \ell_2}\delta_{n_2 k_1}\delta_{m_2 \ell_1}\right. \\
&\qquad \left. - \frac{1}{d_A}\, \delta_{n_1 k_1}\delta_{n_2 k_2}\delta_{m_1 \ell_2}\delta_{m_2 \ell_1} - \frac{1}{d_A}\, \delta_{n_1 k_2}\delta_{n_2 k_1}\delta_{m_1 \ell_1}\delta_{m_2 \ell_2}\right] \\
\int_{\|O\|_2^2=1} dO\, O^*_{\alpha_1\beta_1} O^*_{\alpha_2\beta_2} O_{\gamma_1\delta_1} O_{\gamma_2\delta_2} &= \frac{1}{d_B^4 + d_B^2} \left[\delta_{\alpha_1\gamma_1}\delta_{\alpha_2\gamma_2}\delta_{\beta_1\delta_1}\delta_{\beta_2\delta_2} + \delta_{\alpha_1\gamma_2}\delta_{\alpha_2\gamma_1}\delta_{\beta_1\delta_2}\delta_{\beta_2\delta_1}\right]
\end{aligned}$$

Using these identities, $\overline{\mathrm{CI}}(A:B)$ can be written as

$$\begin{aligned}
\overline{\mathrm{CI}}(A:B) = \frac{1}{(d_A^2-1)\,(d_B^4+d_B^2)}\, \mathrm{tr}\Bigg[&\left(X_{A_1} - \frac{\mathbf{1}_{A_1}}{d_A} \otimes \frac{\mathbf{1}_{A_1}}{d_A}\right)\left(X_{A_2} - \frac{\mathbf{1}_{A_2}}{d_A} \otimes \frac{\mathbf{1}_{A_2}}{d_A}\right) \\
&\times \left((\mathbf{1}_{B_1} \otimes \mathbf{1}_{B_1}) \otimes (\mathbf{1}_{B_2} \otimes \mathbf{1}_{B_2}) + X_{B_1} \otimes X_{B_2}\right) R^{\otimes 2}\Bigg]
\end{aligned} \tag{4.75}$$

where X_{A_1} is the swap operator $[X_{A_1}]_{n_1 n_2, k_1 k_2} = \delta_{n_1 k_2}\delta_{n_2 k_1}$ on $\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_1}$ (and so swaps the A_1 subsystem of of the first copy of R with the A_1 subsystem of the second copy of R), and X_{A_2}, X_{B_1}, X_{B_2} are defined similarly.

If A and B are mutually unitary regions, we can relate R to the superdensity operator ϱ for operator insertions on the regions A and B. (For a review of superdensity operators, see Appendix B.) In this case, if we multiplicatively normalize the tensor network so that

$$M(\mathbf{1}_A : \mathbf{1}_B) = 1\,, \tag{4.76}$$

then by the definition of mutually unitary regions in Section 4.2 and Eqn. (4.10), we have

$$1 = \int dU_A \, dU_B \, M\left(U_A : U_B\right) = \frac{1}{d_A d_B} \operatorname{tr}(R) \,, \tag{4.77}$$

and thus

$$\varrho = \frac{1}{d_A d_B} R \tag{4.78}$$

is a superdensity operator. As per Eqn. (4.74), we can treat ϱ as a density operator on $\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2} \otimes \mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2}$. Interestingly, we can write $\overline{\mathrm{CI}}(A:B)$ in terms of Rényi-2 entropies $S^{(2)}$ of ϱ as

$$\begin{aligned}\overline{\mathrm{CI}}(A:B) = \frac{d_A^2}{\left(d_B^2+1\right)\left(d_A^2-1\right)} \Bigg[& e^{-S^{(2)}_{A_1}} + e^{-S^{(2)}_{A_1A_2B_1B_2}} - \frac{1}{d_A}\left(e^{-S^{(2)}_{A_2}} + e^{-S^{(2)}_{A_2B_1B_2}}\right) \\ & - \frac{1}{d_A}\left(e^{-S^{(2)}_{A_1}} + e^{-S^{(2)}_{A_1B_1B_2}}\right) + \frac{1}{d_A^2}\left(1 + e^{-S^{(2)}_{B_1B_2}}\right)\Bigg] \,.\end{aligned} \tag{4.79}$$

In the above equation, we have, for instance

$$S^{(2)}_{A_1} := -\log \operatorname{tr}(\varrho^2_{A_1})$$

where $\varrho_{A_1} = \operatorname{tr}_{A_2B_1B_2}(\varrho)$. The Rényi-2 entropies of other combinations of subsystems are defined similarly. Note that Eqn. (4.79) is particularly interesting since it relates causality to *spacetime* entropies.

4.6.2 Spin chain examples

The aQCI, $\overline{\mathrm{CI}}(A:B)$, serves as an unbiased measure of causal influence, which only depends on the A, B regions and the tensor network. To obtain more intuition about its behavior, we study $\overline{\mathrm{CI}}(A:B)$ in an example system. Consider a spin chain with continuous time evolution. Here, A and B are single-site subsystems at two different times t_1, t_2, as is illustrated earlier in Figure 4.2. It should be noted that the tensor network description and the definition of causal influence apply to continuous time evolution, since we can treat a time evolution operator such as $U(t_2, t_1) = e^{-iH(t_2-t_1)}$ as a big tensor with $2L$ legs (i.e., L input legs and L output legs), when the spin chain has L sites. Our numerical results for $\overline{\mathrm{CI}}(A:B)$ are shown in Figure 4.23. We studied the dependence of $\overline{\mathrm{CI}}(A:B)$ on initial states and the Hamiltonian. The model we consider is an Ising model with a generic magnetic field:

$$H = J \sum_{n=1}^{L} \sigma_n^z \sigma_{n+1}^z + \sum_{\alpha=x,y,z} h_\alpha \sum_{n=1}^{L} \sigma_n^\alpha \tag{4.80}$$

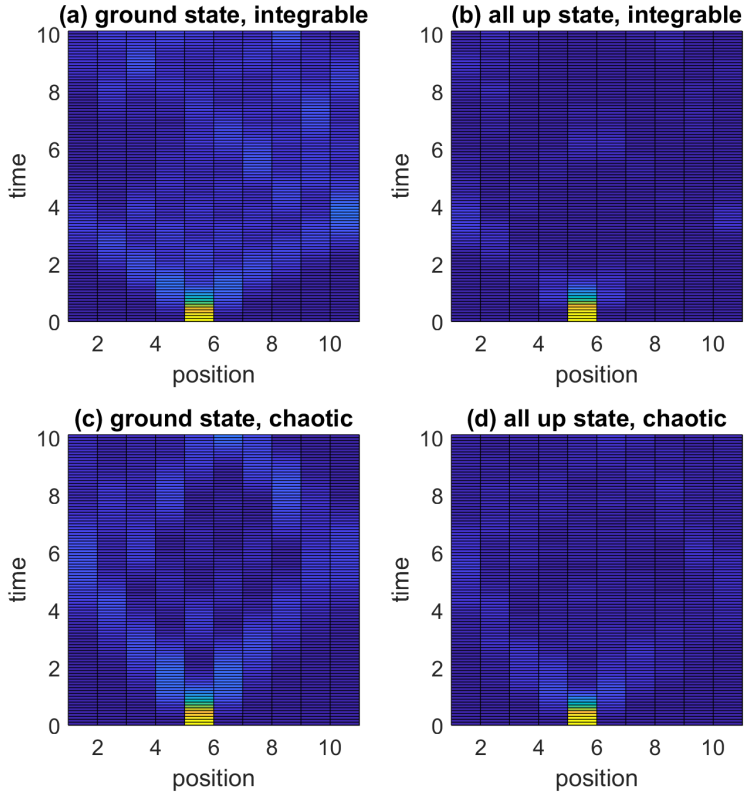


Figure 4.23: The averaged quantum causal influence $\overline{\mathrm{CI}}(A:B)$ for a quantum spin chain for length-1 regions A and B. Region A is at site 5 in the middle of the chain at time $t = 0$. The heat maps depict $\overline{\mathrm{CI}}(A:B)$ as a function of the position and time of B. Results are obtained for two different initial states: the ground state and the "all-up" state. The calculation is done for a quantum Ising model with 10 sites. The Hamiltonian has nearest neighbor ZZ interactions with coupling $J = 1$, a transverse field, and open boundary conditions. The coupling for the transverse field is $\vec{h} = (1, 0, 0)$ for the integrable model (see (a) and (b)) and $\vec{h} = (1.48, 0, -0.7)$ for the chaotic model (see (c) and (d)).

The model is integrable if the magnetic field $\vec{h}$ is in the xy–plane, and the model is chaotic otherwise.

As seen in Figure 4.23, the aQCI is strong and long-lasting if the system is integrable and the initial state is the ground state. If the system is chaotic and the initial state is the ground state, the aQCI is a bit weaker, but still lasts for long times. In contrast, if the system is integrable and the initial state is a finite energy density state (here we use the "all-up" state as an example), the causal influence has some revivals but otherwise decays. Finally, if the system is chaotic and the initial state is a finite energy density state, the causal influence decays uniformly with time.

To further investigate the initial state dependence of quantum causal influence, we start from the ground state $|G\rangle$ of the spin chain and apply a Haar random unitary U_R to the right half of

the system (see Figure 4.24). The resulting state $U_R|G\rangle$ has a high energy density in its right half (sites 6 through 10 in the Figure) and the ground state energy density in its left half (sites $1-5$). Evolving the system in time, energy propagates into the left half and ultimately heats up the whole system. Consequently, the quantum causal influence of a region in the left half, such as site 1 at $t = 0$, behaves like quantum causal influence in the ground state until the "heat wave" arrives. This is consistent with the numerical results in Figure 4.24 (b).

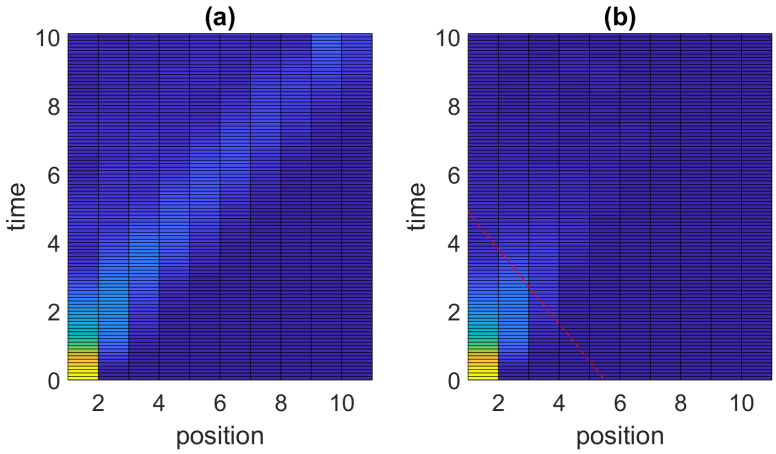


Figure 4.24: Initial state dependence of the aQCI. (a) The aQCI, $\overline{CI}(A:B)$, of the quantum Ising model with region A being site 1 at $t = 0$, as a function of the position and time of the single site region B. The initial state is the ground state $|G\rangle$. (b) The same quantity with the initial state $U_R|G\rangle$, where U_R is a Haar random unitary operator acting on the right half of the system. The red dashed line is a visual guide of the "heat wavefront." The calculation is performed for the quantum Ising model with $J = 1$, $\vec{h} = [1.48, 0, 0.70]$, with open boundary conditions.

4.6.3 Stabilizer tensor network examples

Here we apply our formula for the aQCI to stabilizer tensor networks [83], which provide a numerically tractable toy model for Trotterized Hamiltonian evolution. Stabilizer tensor networks are reviewed in Appendix D. In such networks, the entanglement entropy of any subsystem, as well as reduced density matrices of small subsystems, can be evaluated exactly in polynomial time in the network size [72]. Our chosen geometry is shown in Figure 4.9, where every vertex tensor is a stabilizer code. The horizontal direction is viewed as space (with periodic boundary conditions) and the vertical direction is viewed as time. As the network structure is periodic with respect to pairs of layers of tensors, the time is set to increase by one for every two layers. Furthermore, links in each layer are positioned at 1, 2, ... so that the speed of light in Figure 4.9 is $c = 2$.

In the following, we will consider two examples of qutrit stabilizer tensor networks (i.e., there is a three-dimensional Hilbert space assigned to each link in the network) with stabilizer initial states

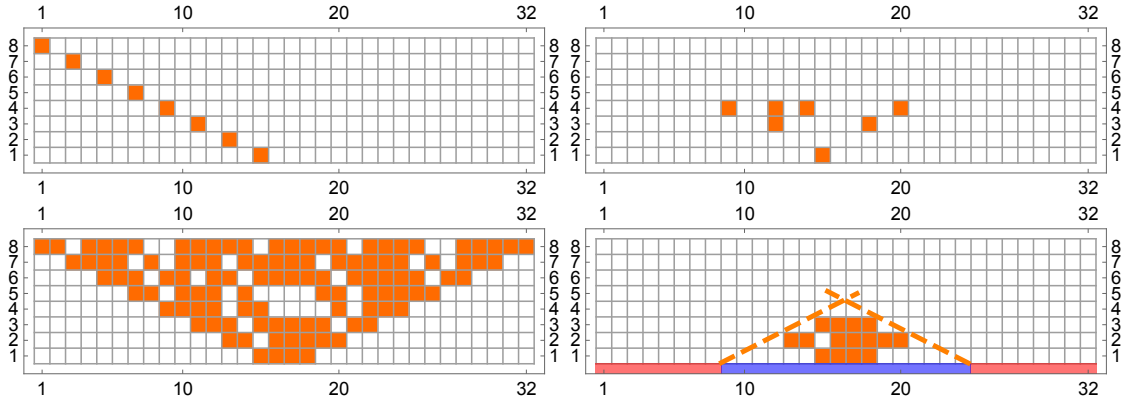


Figure 4.25: The causal future of a link at $t = 0$ (pointing to the upper left in the center of the lowest layer) is colored orange. In particular, the orange points are individually causally influenced by the the link at the initial time. The vertical axis is time and the horizontal axis is space (links) with periodic boundaries. Top-left: the integrable swap code with a random stabilizer initial state; top-right: the perfect [[4,0,3]] qutrit code with the same initial state; bottom-left: the same perfect code with initial state $\bigotimes |0\rangle\langle 0|$; bottom-right: the same perfect code with an infinite temperature initial state in the region marked red and $\bigotimes |0\rangle\langle 0|$ marked blue. Dashed lines are visual guides for the light cones of the red regions.

ρ_i. For clarity, details of the stabilizers and algorithms are recapitulated in Appendix D and only physically relevant features of these codes will be discussed here. In the first example, all tensors are chosen to be the swap code; as a unitary two-to-two gate each tensor is written as $|i\rangle|j\rangle \mapsto |j\rangle|i\rangle$ where $i, j \in \mathbb{F}_3$. This may serve as a toy model for integrable systems where particles propagate ballistically without scattering.

In the second example, all tensors are chosen to be the perfect [[4, 0, 3]] code, $|i\rangle|j\rangle \mapsto |\frac{i+j}{2}\rangle|\frac{i-j}{2}\rangle$ where division by two is evaluated in $\mathbb{F}_3$. It is straightforward to verify that the tensor, viewed as a gate from any two of the four links to the other two, is unitary (such tensors are called perfect, as mentioned in Section 4.5.1). Interestingly, the Heisenberg evolution of operators in such networks exhibits the growth of operator length (linearly in time), which captures some salient physics of scrambling in systems with spatial locality.

For a fixed U_x insertion at time $t = 0$, all positions y for which $\mathrm{CI}(x : y) > 0$ are colored orange in Figure 4.25. In the case of swap codes, the information from the U_x insertion propagates ballistically and the causal future coincides with the future light cone of x. The specific direction of information propagation in the Figure depends on which link (left- or right-moving) U_x acts on.

Results for the perfect code are remarkably different. For a generic initial state, as shown in the top-right panel of Figure 4.25, the causal influence of a point x at $t = 0$ on local regions in the future is small and vanishes for late times, which shows that information at x spreads into nonlocal degrees of freedom. However, for the special initial state $\rho_i = \bigotimes |0\rangle\langle 0|$, the causal future of x (with respect to local subregions) is the filled future light cone. Although there is not a sharp notion of

thermal initial states in stabilizer tensor networks, such a causal influence structure suggests that ρ_i is similar to a "cold" low-energy state of a local Hamiltonian (although energy is not well-defined in this Trotterized tensor network) because the causal influence does not decay substantially in the future (and hence does not quickly "thermalize"). Previously, we saw that low energy states of a quantum Ising model exhibit similar behavior, justifying our use of "cold" and "low-energy" in describing $\bigotimes |0\rangle\langle 0|$ for our stabilizer tensor network.

In Figure 4.25 we have implemented an initial state $\rho_i = \bigotimes_{\text{hot}} \frac{1}{3} I \otimes \bigotimes_{\text{cold}} |0\rangle\langle 0|$ where in "hot" regions the initial state is at infinite temperature and in "cold" regions it is the product state. The causal future of x terminates when it is engulfed by heatwaves from the infinite temperature subsystem. The initial state dependence of quantum causal influence is manifest in these examples.

4.6.4 An upper bound by spacetime quantum mutual information

Recall that the quantum mutual information provides a bound on spacelike connected correlation functions [253]. An analogous bound on *spacetime* correlation functions was given in terms of superdensity operators in [56] (a short discussion of this can be found in Appendix A). It is natural that the causal influence between two regions is bounded by the spacetime mutual information of a corresponding superdensity operator. Here, we will prove such an inequality:

Bound on causal influence by spacetime quantum mutual information: Consider two spacetime subregions A and B corresponding to Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, and a corresponding superdensity operator ϱ_{AB}. If A, B are mutually unitary regions, we have

$$\mathrm{CI}(A:B)^2 \leq 2\, d_A^2\, I_{\varrho_{AB}}(A:B) \tag{4.81}$$

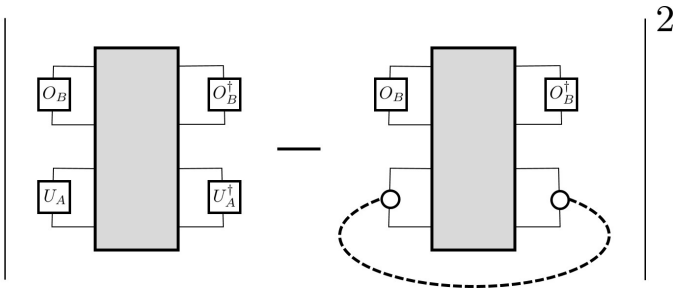
where $I_{\varrho_{AB}}(A:B)$ is the superdensity quantum mutual information between A and B.

Proof. The proof of the inequality is easiest to understand diagrammatically. First we write $\mathrm{CI}(A:B)^2$ as

$$\mathrm{CI}(A:B)^2 = \sup_{U_A, O_B} \frac{1}{\|O_B\|_2^4} \left| M(U_A : O_B) - \int dU_A\, M(U_A : O_B) \right|^2 \tag{4.82}$$

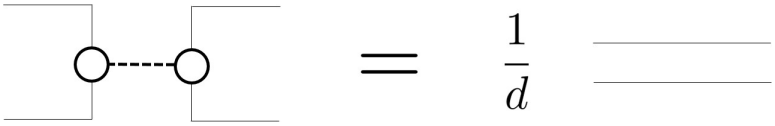
which can be expressed diagrammatically in superdensity operator notation as

$$\mathrm{CI}(A:B)^2 = \sup_{U_A, O_B} \frac{1}{\|O_B\|_2^4}$$



(4.83)

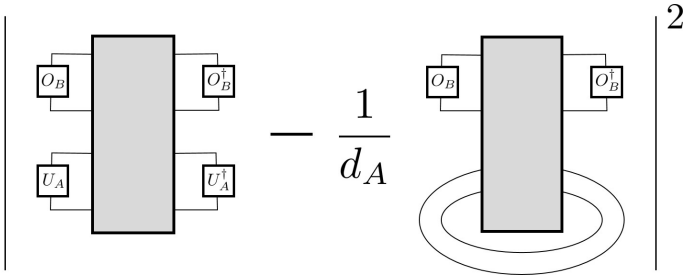
The dotted lines denote the $\int dU_A$ integration. The identity $\int dU\, U_{ij} U^\dagger_{k\ell} = \frac{1}{d}\,\delta_{i\ell}\delta_{jk}$ is depicted by



(4.84)

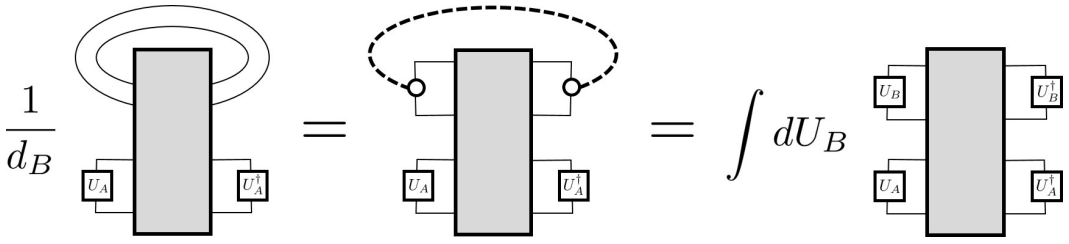
and so our diagram for $\mathrm{CI}(A:B)$ becomes

$$\mathrm{CI}(A:B)^2 = \sup_{U_A, O_B} \frac{1}{\|O_B\|_2^4}$$



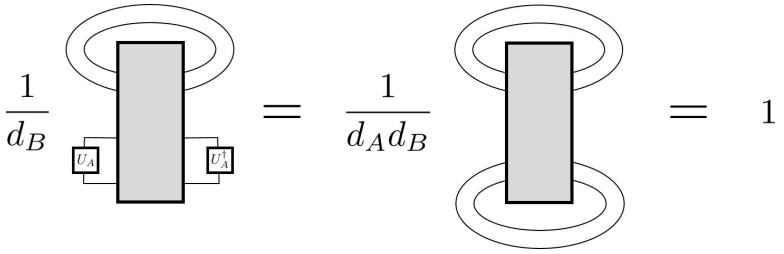
(4.85)

Now consider the identity



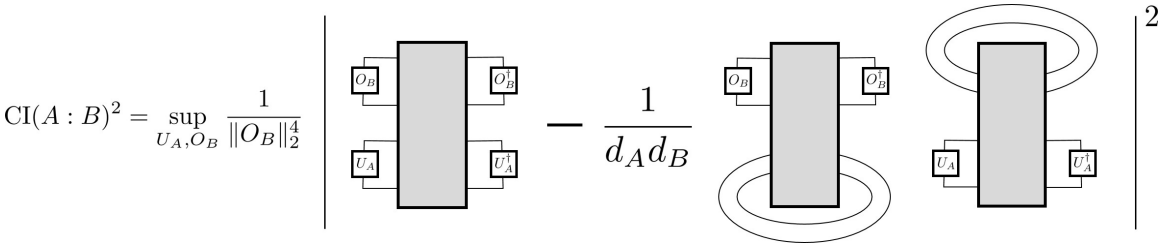
(4.86)

The term inside the integral on the right-hand side is actually independent from U_A and U_B if A, B are mutually unitary regions. Then we can replace the U_A contractions by an average over U_A,



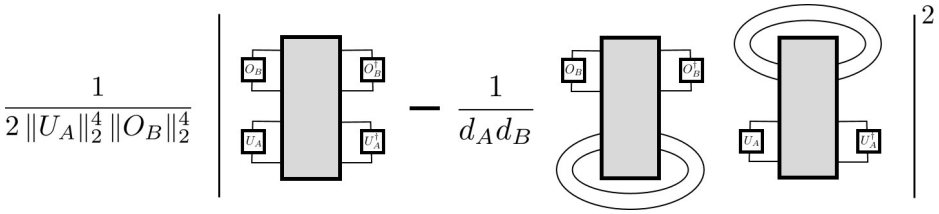
(4.87)

where the last equivalence is just the statement $\mathrm{tr}(\varrho_{AB}) = 1$. Then we can insert this factor of unity into our expression for $\mathrm{CI}(A:B)^2$ to obtain



(4.88)

The term inside the absolute value bars is a connected correlation function with respect to the superdensity operator ϱ_{AB}. Thus, we can use the superdensity operator quantum mutual information bound on connected correlation functions (see [56] and Appendix B for a review), which gives us



(4.89)

Comparing to our expression for $\mathrm{CI}(A:B)^2$, we obtain the desired inequality $\mathrm{CI}(A:B)^2 \leq 2\, d_A^2 I_{\varrho_{AB}}(A:B)$. □

4.7 Conclusion and further discussion

In this paper, we have proposed a new measure of causal structure, the quantum causal influence, in quantum many-body systems. We used the framework of general tensor networks to describe quantum many-body systems without a pre-fixed causal structure. In this framework, we showed how the causal influence between two spacetime regions A, B can be probed by the effect of unitary operations in region A on observables in region B. Unitarity plays an essential role in the asymmetry of the causal influence between two regions. Accordingly, the entanglement inherent in a general tensor network can be seen as building up space, time, and the causal relationships between local and collective spacetime degrees of freedom. Our definition of quantum causal influence provides a new unified perspective on many seemingly disconnected phenomena.

Through examples and more abstract results, we have shown that the quantum causal influence, and therefore the direction of "time's arrow," depends on the initial state and final state of the time evolution. In particular, a maximally mixed subregion of either the initial or final state cannot causally influence other regions. It would be interesting to understand in detail what happens when the initial or final states have subsystems that merely have high entropy (instead of having maximal entropy by virtue of being maximally mixed).

An important feature of the quantum causal influence is its nonlocality: a region A can have trivial influence on regions B, C while having nontrivial influence on their union $B \cup C$. Quantum error correction and quantum teleportation are both examples of such non-local causal influence. The non-locality of causal influence plays an essential role in holographic duality, where small disk-shape regions in the bulk have ordinary causal structure as prescribed by general relativity, while nonlocal regions have a different (and more exotic) causal structure required by the holographic principle. Specifically, any given bulk operator can be reconstructed on a big enough region of the boundary, which means (using our definition) the quantum causal influence of a bulk point on the boundary is nontrivial, even if the point is spacelike separated from the boundary from a Riemannian geometry point of view.

We also discussed how unconventional causal structures appear in the Horowitz-Maldacena final state proposal of the black hole singularity, where again the non-locality of quantum causal influence plays an essential role in reconciling the ordinary causal structure of the black hole geometry (between small disks) and the unitarity of time evolution. Additionally, we studied multiple probes of quantum causal influence, and discussed their relation to other quantum information quantities such as the quantum mutual information and Rényi entropies.

There are many open questions that can be studied with the quantum causal influence. For instance, it is interesting to ask whether there is a precise generalization of Cauchy surfaces defined in terms of the QCI. For instance, such a plausible quantum generalization of Cauchy surfaces is a foliation of a general tensor network into disjoint subsystems $C_1, C_2, ..., C_N$ such that C_i only has nontrivial causal influence with C_j if $j > i$. In addition, one should require that for each C_i, all

of its disjoint subregions are spacelike separated from one another other. In 4.5.1 we discussed an example of such quantum Cauchy surfaces in holographic tensor networks. In general systems, can Cauchy surfaces always be found? When Cauchy surfaces are defined, is it always possible to define a "quantum state" on each surface, as in the (semi-)classical setting?

Another open question is how to generalize the quantum causal influence to measure (the quantum generalization) of spacetime geometry. In a similar vein, there have previously been proposals relating spatial distances between local subsystems to their quantum mutual information [243, 206]. It would be interesting to investigate whether a combination of these ideas can lead to a generalization of quantum causal influence which probes a (quantum generalization of a) spacetime metric. An even more general question concerns whether quantum causal influence can be applied to spacetime tensor networks with fluctuating geometries, such as those proposed in [210].

Acknowledgements

We would like to thank C. Brukner, W. Donnelly, P. Hayden, D. Huse, J. Preskill, D. Ranard, R. Spekkens, and F. Wilczek for helpful discussions. JC is supported by the Fannie and John Hertz Foundation and the Stanford Graduate Fellowship program. XH is supported by the Stanford Graduate Fellowship program. XLQ is supported by the National Science Foundation under grant No. 1720504 and the David and Lucile Packard Foundation. ZY is supported by the David and Lucile Packard Foundation.

4.8 Appendix A: Quantum causal influence for non-unitary regions

Suppose we have a general tensor network given by $\{\{\mathcal{H}_i\}, |L\rangle, \rho_P\}$, and that R_1 is a subregion which is *not* a unitary region. This means that

$$\langle L|U_{R_1}\,\rho_P\,U_{R_1}^\dagger|L\rangle \neq \langle L|\rho_P|L\rangle \tag{4.90}$$

for some unitary U_{R_1}. This situation can occur even in some more modest examples, such as systems with post-selection.

In this context, it is natural to define quantum causal influence for non-unitary regions. We let

$$M'(U_{R_1} : O_{R_2}) := \frac{\langle L|(U_{R_1}\otimes O_{R_2})\rho_P\,(U_{R_1}^\dagger\otimes O_{R_2}^\dagger)|L\rangle}{\langle L|U_{R_1}\,\rho_P\,U_{R_1}^\dagger|L\rangle} \tag{4.91}$$

where R_1 is not a unitary region. Here, M has been furnished with a prime $'$ to distinguish it from the usual $M(U_{R_1} : O_{R_2})$. Then the corresponding mQCI for non-unitary regions is

$$\mathrm{CI}'(R_1 : R_2) = \sup_{U_{R_1}, O_{R_2}} \frac{1}{||O_{R_2}||_2^2}\left|M'(U_{R_1} : O_{R_2}) - \int dU_{R_1}\, M'(U_{R_1} : O_{R_2})\right| \tag{4.92}$$

and similarly, the corresponding aQCI for non-unitary regions is

$$\overline{\mathrm{CI}}'(R_1 : R_2) = \int dU_R \int_{||O_{R_2}||_2^2=1} dO_{R_2}\left|M'(U_{R_1} : O_{R_2}) - \int dU_{R_1}\, M'(U_{R_1} : O_{R_2})\right|^2 . \tag{4.93}$$

Notice that modified mQCI and the modified aQCI are also furnished with primes $'$ to distinguish them for their unmodified counterparts.

Note that if R_1 *is* a unitary region, then

$$\mathrm{CI}'(R_1 : R_2) = \frac{1}{\langle L|\rho_P|L\rangle}\,\mathrm{CI}(R_1 : R_2) \tag{4.94}$$

$$\overline{\mathrm{CI}}'(R_1 : R_2) = \frac{1}{\langle L|\rho_P|L\rangle}\,\overline{\mathrm{CI}}(R_1 : R_2)\,, \tag{4.95}$$

meaning the modified and unmodified mQCI and aQCI are related by a multiplicative constant in this case. Of course, if $\langle L|U_{R_1}\,\rho_P\,U_{R_1}^\dagger|L\rangle = 1$ for all U_{R_1}, then the multiplicative constant becomes one.

4.9 Appendix B: Review of superdensity operator formalism

Throughout the paper, we make use of the superdensity operator formalism to analyze spacetime correlation functions. We review superdensity operators here, and a full exposition can be found in [56].

A superdensity operator is a spacetime analog of a density operator, so first we begin by examining density operators. Consider a Hilbert space $\mathcal{H}$ of dimension d so that the space of density operators on $\mathcal{H}$ is denote by $\mathcal{S}(\mathcal{H})$. A density operator is denoted by ρ and is defined by:

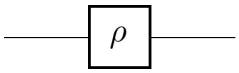
Definition (density operator): *A density operator ρ is a bilinear form*

$$\rho : \mathcal{H}^* \otimes \mathcal{H} \longrightarrow \mathbb{C}$$

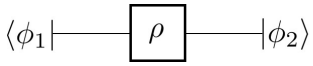
satisfying the conditions:

1. $\rho^\dagger = \rho$ (*Hermitian*)
2. $\rho \succeq 0$, *meaning* $\langle\phi|\rho|\phi\rangle \geq 0$ *for all* $|\phi\rangle$ (*positive semi-definite*)
3. $\mathrm{tr}(\rho) = 1$ (*unit trace*)

Since $\rho : \mathcal{H}^* \otimes \mathcal{H} \longrightarrow \mathbb{C}$, we can represent ρ by the tensor diagram

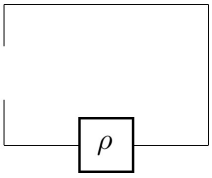


where

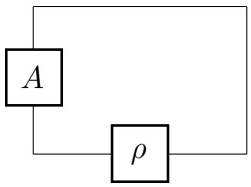


$$= \quad \langle\phi_1|\,\rho\,|\phi_2\rangle$$

Equivalently, we can think of ρ as a map from operators in $\mathcal{B}(\mathcal{H})$ to correlation functions (i.e., a map from $\mathcal{B}(\mathcal{H}) \to \mathbb{C}$) by re-writing the tensor diagram as

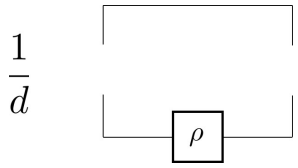


where similarly

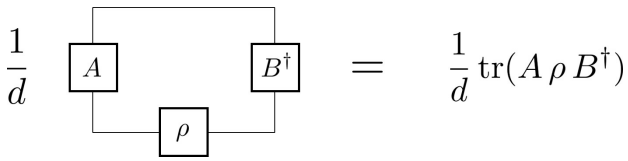


$$= \quad \mathrm{tr}(A\,\rho)$$

Now we introduce a new object which may at first appear peculiar, but will later appear natural. It is given diagramatically by



This object satisfies



and so is a bilinear form from $\mathcal{B}^*(\mathcal{H}) \otimes \mathcal{B}(\mathcal{H}) \to \mathbb{C}$.

This new object is clearly a repackaging of ρ, since it contains all of the same data. Now let us write this new object in non-diagrammatic notation, and call it ϱ_{super}. Consider the space of operators on $\mathcal{H}$, denoted by $\mathcal{B}(\mathcal{H})$. Let $\{X_i\}_{i=1}^{d^2}$ be an orthonormal basis of operators for $\mathcal{B}(\mathcal{H})$, so that $\text{tr}(X_i^\dagger X_j) = \delta_{ij}$. Since $\mathcal{B}(\mathcal{H})$ is itself a Hilbert space, we can write its basis in bra-ket notation as $\{|X_i\rangle\}_{i=1}^{d^2}$ where $\langle X_i|X_j\rangle := \text{tr}(X_i^\dagger X_j) = \delta_{ij}$. Then we can write ϱ_{super} in this basis as

$$\varrho_{\text{super}} = \frac{1}{d} \sum_{i,j=1}^{d^2} \text{tr}(X_i \, \rho \, X_j^\dagger) \, |X_i\rangle\langle X_j| \,. \tag{4.96}$$

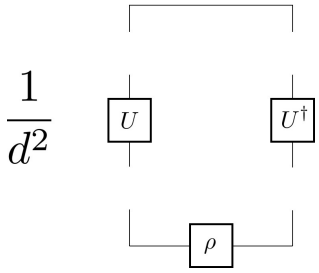
Then we have

$$\langle A| \, \varrho_{\text{super}} \, |B\rangle = \frac{1}{d} \, \text{tr}(A \, \rho \, B^\dagger) \tag{4.97}$$

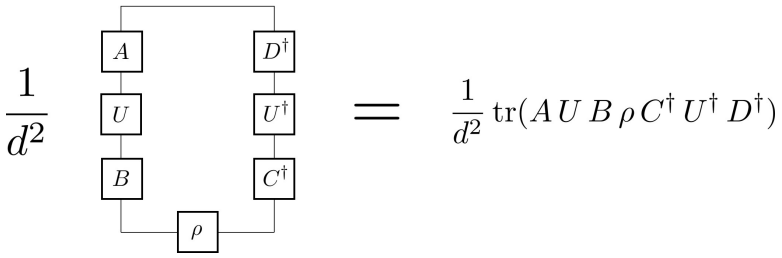
which matches the diagram above.

Several comments are in order. The object ϱ_{super} is our first example of a superdensity operator, which we will define shortly. While a standard density operator ρ is a map $\rho : \mathcal{H}^* \otimes \mathcal{H} \to \mathbb{C}$, the object ϱ_{super} is a map $\varrho_{\text{super}} : \mathcal{B}^*(\mathcal{H}) \otimes \mathcal{B}(\mathcal{H}) \to \mathbb{C}$. In fact, it is easy to check that ϱ_{super} is Hermitian, positive semi-definite, and has unit trace. Therefore, just as ρ is a density operator on $\mathcal{H}$, we have that ϱ_{super} is a density operator on $\mathcal{B}(\mathcal{H})$ (and hence a *super*density operator).

So far, we have merely repackaged ρ as the superdensity operator ϱ_{super}. Both objects capture the data of correlation functions of a system at a single time. But now suppose we want to capture the data of the correlation functions of a system at two times. Letting U be the unitary evolution between these two times, we can write down the new superdensity operator σ_{super}, namely



which satisfies



and can be written non-diagrammatically as

$$\sigma_{\text{super}} = \frac{1}{d^2} \sum_{i,j,k,\ell=1}^{d^2} \mathrm{tr}(X_i\,U\,X_j\,\rho\,X_k^\dagger\,U^\dagger\,X_\ell^\dagger)\,|X_j\rangle\langle X_k| \otimes |X_i\rangle\langle X_\ell|\,. \tag{4.98}$$

Here, σ_{super} maps operators at an initial time t_1 and operators at a final time t_2 to a correlation function. We can write this map as

$$\sigma_{\text{super}} : \big(\mathcal{B}^*(\mathcal{H}_{t_1}) \otimes \mathcal{B}(\mathcal{H}_{t_1})\big) \otimes \big(\mathcal{B}^*(\mathcal{H}_{t_2}) \otimes \mathcal{B}(\mathcal{H}_{t_2})\big) \longrightarrow \mathbb{C}\,, \tag{4.99}$$

or isomorphically

$$\sigma_{\text{super}} : \mathcal{B}^*(\mathcal{H}_{t_1} \otimes \mathcal{H}_{t_2}) \otimes \mathcal{B}(\mathcal{H}_{t_1} \otimes \mathcal{H}_{t_2}) \longrightarrow \mathbb{C}\,. \tag{4.100}$$

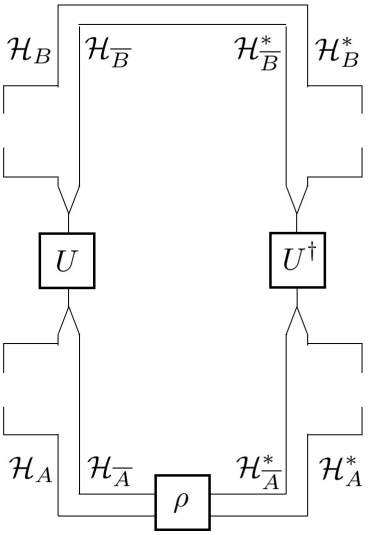
Indeed, σ_{super} is Hermitian, positive semi-definite, and has unit trace. Therefore, σ_{super} is a density operator on the operator space $\mathcal{B}(\mathcal{H}_{t_1} \otimes \mathcal{H}_{t_2})$. We refer to Hilbert spaces of the form $\bigotimes_t \mathcal{H}_t$, such as $\mathcal{H}_{t_1} \otimes \mathcal{H}_{t_2}$, as "history Hilbert spaces."

As illustrated above, σ_{super} contains the data of two-time correlation functions of a system, all packaged into a density operator on an appropriate operator space (for instance, $\mathcal{B}(\mathcal{H}_{t_1} \otimes \mathcal{H}_{t_2})$). The reason we package this data into a density operator is because we can immediately use many of the tools and techniques of quantum information theory, which are designed for generic density operators (although they are typically applied only to standard density operators). For instance, one can compute spacetime quantum entropies, spacetime quantum mutual information, and so on,

and the results are physically and mathematically meaningful (see [56] for an in-depth discussion of these points). We will remark on the quantum mutual information below.

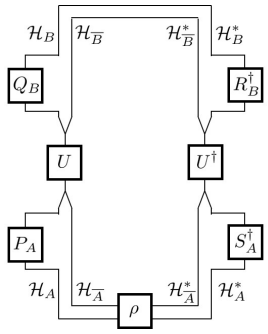
Of course, our construction above naturally generalizes to any number of times $t_1, t_2, ..., t_n$. The construction also generalizes to subsystems of the Hilbert space in the following way. Consider a Hilbert space $\mathcal{H}$ which has (possibly overlapping) subsystems $\mathcal{H}_A$ and $\mathcal{H}_B$ with dimensions d_A and d_B, respectively. We will consider, for concreteness, a two-time superdensity operator χ_{super}, given diagrammatically by

$\frac{1}{d_A d_B}$



satisfying

$\frac{1}{d_A d_B}$



$$= \frac{1}{d_A d_B} \operatorname{tr}\left((Q_B \otimes \mathbf{1}_{\overline{B}})\, U\, (P_A \otimes \mathbf{1}_{\overline{A}})\, \rho\, (S_A^\dagger \otimes \mathbf{1}_{\overline{A}})\, U^\dagger\, (S_B^\dagger \otimes \mathbf{1}_{\overline{B}}) \right)$$

and written in non-diagrammatic notation as

$$\chi_{\text{super}} = \frac{1}{d_A d_B} \sum_{i,j=1}^{d_A^2} \sum_{k,\ell=1}^{d_B^2} \operatorname{tr}\left(\left(X_k^B \otimes \mathbf{1}_{\overline{B}}\right)\, U\, \left(X_i^A \otimes \mathbf{1}_{\overline{A}}\right)\, \rho\, \left(X_j^{\dagger\, A} \otimes \mathbf{1}_{\overline{A}}\right)\, U^\dagger\, \left(X_\ell^{\dagger\, B} \otimes \mathbf{1}_{\overline{B}}\right) \right) \times |X_i^A\rangle\langle X_j^A| \otimes |X_k^B\rangle\langle X_\ell^B| \,. \quad (4.101)$$

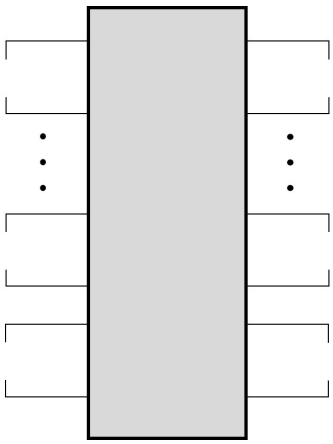
In this case, χ_{super} is a map from

$$\chi_{\text{super}} : \mathcal{B}^*(\mathcal{H}_{A,\,t_1} \otimes \mathcal{H}_{B,\,t_2}) \otimes \mathcal{B}(\mathcal{H}_{A,\,t_1} \otimes \mathcal{H}_{B,\,t_2}) \longrightarrow \mathbb{C}\,, \tag{4.102}$$

and is likewise Hermitian, positive semi-definite, and has unit trace. Then χ_{super} captures the data of two-time correlation functions with operators on the subsystem A at time t_1 and operators on the subsystem B at time t_2. This construction generalizes naturally to many times $t_1, t_2, ..., t_n$ and arbitrary subsystems at each time.

The superdensity operators we have considered so far have a particular form: an initial state followed by slots for operator insertions, followed by unitary evolution, followed by more slots for operators insertions, and so on until a final trace is taken. These kinds of superdensity operators can also be thought of as the quantum state of ancillary apparatus which couples to an evolving system in a certain manner (see [56] for details).

More generally, we might be agnostic to the internal structure of a superdensity operator ϱ, and notate it as



which is a bilinear map

$$\varrho : \mathcal{B}^*(\mathcal{H}_{\text{hist.}}) \otimes \mathcal{B}(\mathcal{H}_{\text{hist.}}) \longrightarrow \mathbb{C} \tag{4.103}$$

for some Hilbert space $\mathcal{H}_{\text{hist.}}$ that we designate as the history Hilbert space (in keeping with our previous terminology). We may require that ϱ is Hermitian, positive semi-definite, and has unit trace, so that it is formally a density operator (albeit on an operator space $\mathcal{B}(\mathcal{H}_{\text{hist.}})$). This brings us to the definition:

Definition (superdensity operator): *A superdensity operator ϱ is a bilinear form*

$$\varrho : \mathcal{B}^*(\mathcal{H}_{\text{hist.}}) \otimes \mathcal{B}(\mathcal{H}_{\text{hist.}}) \longrightarrow \mathbb{C}$$

satisfying the conditions:

1. $\varrho^\dagger = \varrho$ (*Hermitian*)
2. $\varrho \succeq 0$, *meaning* $\langle W|\varrho|W\rangle \geq 0$ *for all* $|W\rangle$ (*positive semi-definite*)
3. $\mathrm{tr}(\varrho) = 1$ (*unit trace*)

As mentioned above, measures of quantum information of density operators can be upgraded to be measures of spacetime quantum information of superdensity operators. These upgraded measures are meaningful [56]. For instance, recall the quantum mutual information bound [253]

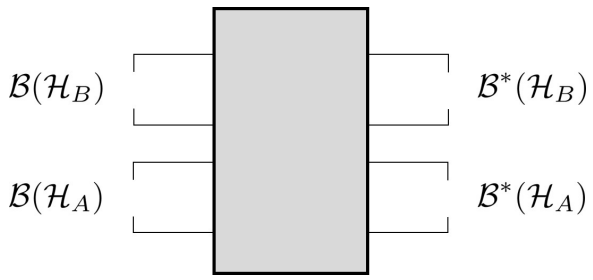
$$\frac{1}{2\,\|P_A\|_1^2\,\|Q_B\|_1^2}\left|\mathrm{tr}\left(\left(P_A\otimes Q_B\right)\rho\right)-\mathrm{tr}\left(P_A\,\rho\right)\mathrm{tr}\left(Q_B\,\rho\right)\right|^2 \leq I_\rho(A:B) \tag{4.104}$$

where $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \cdots$ and $I_\rho(A:B)$ is the quantum mutual information between A and B with respect to ρ. Here, A and B are arbitrary disjoint spatial subregions.

One can straightforwardly show [56] that the superdensity analog is

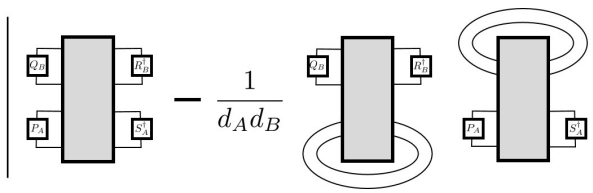
$$\begin{aligned}&\frac{1}{2\,\|P_A\|_2^2\,\|Q_B\|_2^2\,\|R_B\|_2^2\,\|S_A\|_2^2}\\&\quad\times\left|\left(\langle P_A|\otimes\langle Q_B|\right)\varrho_{\text{super}}^{AB}\left(|S_A\rangle\otimes|R_B\rangle\right)-\langle P_A|\,\mathrm{tr}_{\mathcal{B}(\mathcal{H}_B)}\left(\varrho_{\text{super}}^{AB}\right)|S_A\rangle\,\langle Q_B|\,\mathrm{tr}_{\mathcal{B}(\mathcal{H}_A)}\left(\varrho_{\text{super}}^{AB}\right)|R_B\rangle\right|^2\\&\qquad\leq I_{\varrho_{\text{super}}^{AB}}(A:B)\end{aligned} \tag{4.105}$$

where $\mathcal{B}(\mathcal{H}_{\text{hist.}}) = \mathcal{B}(\mathcal{H}_A)\otimes\mathcal{B}(\mathcal{H}_B)\otimes\cdots$ and $I_{\varrho_{\text{super}}^{AB}}(A:B)$ is the (spacetime) quantum mutual information between A and B with respect to $\varrho_{\text{super}}^{AB}$ which can be depicted by



Here, by contrast, A and B are arbitrary disjoint spacetime subregions. The spacetime quantum mutual information bound can be depicted diagrammatically by

$$\frac{1}{2\,\|P_A\|_2^2\,\|Q_B\|_2^2\,\|R_B\|_2^2\,\|S_A\|_2^2}$$



$$\leq I_{\varrho_{\text{super}}^{AB}}(A:B)$$

which we utilize in Section 4.6.4.

4.10 Appendix C: Classical analog of non-local causality

In this paper, we have been primarily focused on causal influence in quantum systems. Here, we will explore features of causal influence in classical systems, and in particular focus on non-local aspects of causal influence. We will compare and contrast with the quantum case, and find key differences.

In order to adapt our framework to the classical setting, we find it convenient to embed a classical system into a quantum system, and continue to use bra-ket notation and the operator formalism. First, we establish how to present a classical system in this notation. Suppose we have n qubits, and consider the canonical basis $\{|i_1 \cdots i_n\rangle\}_{i_1,\ldots,i_n=0}^{1}$ which picks out the z-basis for every qubit. We will refer to this basis as the classical basis, and write it more compactly using multi-index notation as $\{|I\rangle\}_{I\in\{0,1\}^n}$. We require that a classical density operator $\rho_{\text{classical}}$ is a convex combination of projectors onto classical basis elements, namely of the form

$$\rho_{\text{classical}} = \sum_{I\in\{0,1\}^n} p_I \, |I\rangle\langle I| \,, \qquad \sum_{I\in\{0,1\}^n} p_I = 1 \,, \qquad p_I \geq 0 \text{ for all } I. \tag{4.106}$$

In words, a classical density operator is a probabilistic (incoherent) mixture of classical states in which each qubit has a definite z-direction.

Now we construct operators which act on classical states. An arbitrary operator A has the form

$$A = \sum_{I\in\{0,1\}^n} a_I \, |f(I)\rangle\langle I| \,, \tag{4.107}$$

where f is an arbitrary function $f : \{0,1\}^n \to \{0,1\}^n$ and the a_I's are complex numbers. Notice that this operator maps pure classical states to pure classical states (up to a complex scalar prefactor) since $O|J\rangle = c_J|f(J)\rangle$. We can specialize to Hermitian operators B which have the form

$$B = \sum_{I\in\{0,1\}^n} b_I \, |f(I)\rangle\langle I| \,, \qquad f \circ f = \text{Identity} \,, \qquad b_I = b^*_{f(I)} \,. \tag{4.108}$$

Here, we see that $f : \{0,1\}^n \to \{0,1\}^n$ is its own inverse, meaning that $f \circ f$ is the identity map.

Now we turn to observables. In the classical context, observables C are Hermitian operators that satisfy the superselection rule $\langle I|O|J\rangle = 0$ if $I \neq J$, so that the eigenvectors cannot be superpositions of classical states. Thus, observables have the form

$$O = \sum_{I\in\{0,1\}^n} c_I \, |I\rangle\langle I| \tag{4.109}$$

where the c_I's are real numbers.

Finally, the classical analog of unitary operators are invertible operators satisfying $P^\dagger P = PP^\dagger = \mathbf{1}$. Comparing to Eqn. (4.107), we see that such a P must have the form

$$P = \sum_{I \in \{0,1\}^n} |f(I)\rangle\langle I| \,, \qquad f \text{ invertible.} \tag{4.110}$$

This means that P is a permutation operator on the classical basis elements. This is intuitive: the classical analog of unitary evolution can only interchange classical states.

Now we define the classical analog of causal influence for our n-qubit system. Analogous to Eqn. (4.12), we define the classical maximal influence by

$$\mathrm{CI}_{\text{classical}}(A:B) := \sup_{\substack{P_A \in \text{ permutations on } A \\ O_B \in \text{ classical operators on } B}} \frac{1}{||O_B||_2^2} \left| M\left(P_A : O_B\right) - \frac{1}{n!} \sum_{P_A \in \text{Perms}} M\left(P_A : O_B\right) \right| . \tag{4.111}$$

Having set up classical causal influence, we turn to an example.[9] We will consider a hallmark of classical cryptography: the one-time pad. Suppose we have two parties Alice and Bob, and that Alice has a secret message that she wishes to share with Bob. For concreteness, suppose that this secret message M comprises of an n-bit string. In the one-time pad protocol, Alice and Bob share in advance a secret key K, called the one-time pad, which is likewise an n-bit string that is unknown to anyone else. This secret key K has been sampled from a uniform distribution on all n-bit strings and must be discarded the protocol is completed (i.e., only used "one time"). Suppose Alice's messages is $(x_1, x_2, ..., x_n)$ with $x_i \in \{0,1\}$, and the secret key is $(y_1, y_2, ..., y_n)$ with $y_i \in \{0,1\}$. Then Alice produces an encrypted message E, whose ith bit is the sum, modulo 2, of the ith bits of M and K. The encrypted message E would be

$$((x_1 \oplus y_1), (x_2 \oplus y_2), ..., (x_n \oplus y_n)) \,, \tag{4.112}$$

where here $\oplus$ denotes summation modulo 2. This encrypted message is then sent to Bob. Bob decodes the message by taking its ith bit, and adding it modulo 2 to the ith bit of the secret key. The result is

$$\begin{aligned} &((x_1 \oplus y_1 \oplus y_1), (x_2 \oplus y_2 \oplus y_2), ..., (x_n \oplus y_n \oplus y_n)) \\ &= (x_1, x_2, ..., x_n) \,, \end{aligned} \tag{4.113}$$

which is exactly Alice's original message M. The secret key K (i.e., the one-time pad) cannot be used in subsequent instantiations of the protocol since an eavesdropper can glean information about encrypted messages by looking for patterns, although we will not discuss this in detail here.

Let us express the encoding step of this protocol in terms of a superdensity operator. Consider

[9] We thank Robert Spekkens for suggesting this example.

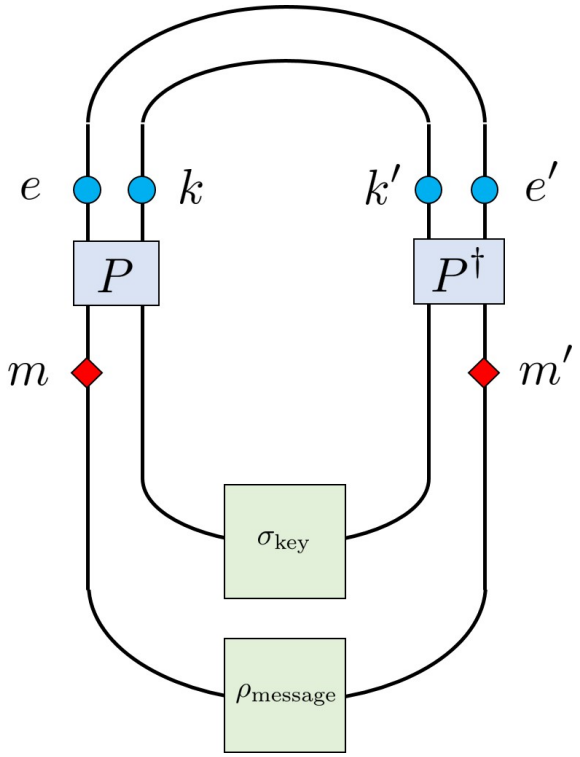


Figure 4.26: A diagram for the one-time pad. Here, ρ_{message} is the state of the message, σ_{key} is the state of the secret key, and P is encrypts the message using the secret key, as described in Eqn. (4.114) and the surrounding text.

the diagram in Figure 4.26 below. Let $\rho_{\text{message}} = |M\rangle\langle M|$, which is a classical state corresponding to the secret message. Let σ_{key} be the uniform distribution over classical states, namely the maximally mixed state $\sigma_{\text{key}} = \frac{1}{2^n}\sum_{J\in\{0,1\}^n} |J\rangle\langle J| = \frac{1}{2^n}\mathbf{1}$. We also let P map

$$P|I\rangle \otimes |J\rangle = |I \oplus J\rangle \otimes |J\rangle\,, \tag{4.114}$$

where $I \oplus J$ represents bitwise addition modulo 2 as per Eqn. (4.107). Then we have

$$P(\rho_{\text{message}} \otimes \sigma_{\text{key}})P^\dagger = \frac{1}{2^n}\sum_{J\in\{0,1\}^n} |M \oplus J\rangle\langle M \oplus J| \otimes |J\rangle\langle J|\,. \tag{4.115}$$

Now let us consider the classical causal influence between m (a place where an operator insertion affects the message) and e (a place where an operator insertion probes the encrypted message). Since

$$\text{tr}_{\text{key}}\left(P(\rho_{\text{message}} \otimes \sigma_{\text{key}})P^\dagger\right) = \frac{1}{2^n}\,\mathbf{1}\,, \tag{4.116}$$

it follows that

$$\mathrm{CI}_{\mathrm{classical}}(m : e) = 0\,. \tag{4.117}$$

This is intuitive – it means that manipulating the message at m does not affect the encrypted message at e, and hence no information from the message is contained in e alone. Thus, if an eavesdropper was positioned at e and could tamper with the encrypted message, the secret message could not be discovered.

Similarly, we can consider the classical causal influence between m and k (a place where an operator insertion probes the secret key). Since

$$\mathrm{tr}_{\text{encrypted message}}\left(P(\rho_{\text{message}} \otimes \sigma_{\text{key}})P^\dagger\right) = \frac{1}{2^n}\,\mathbf{1}\,, \tag{4.118}$$

we find

$$\mathrm{CI}_{\mathrm{classical}}(m : k) = 0\,. \tag{4.119}$$

This is not surprising at all, since the initial message is not correlated with the secret key.

However, if we consider the classical causal influence between m and $e \cup k$, we find

$$\mathrm{CI}_{\mathrm{classical}}(m : e \cup k) > 0\,. \tag{4.120}$$

The result again is intuitive, since given access to both the encrypted message and the secret key, one can recover the initial message. This is an example of classical non-local causal influence: even though m does not influence either e and k, it influences $e \cup k$.

This example appears superficially similar to examples of non-local causal influence earlier in the paper, such as the quantum erasure code example in Section 4.4.1 above. However, there are key differences. In our classical example, we treated the state of the key as a uniform distribution over all n-bit strings. But in an actual instantiation of the protocol, a *particular* key K is chosen, and so $\sigma_{\text{key}} = |K\rangle\langle K|$ would be a pure state. In this case, we would find $\mathrm{CI}_{\mathrm{classical}}(m : k) = 0$, $\mathrm{CI}_{\mathrm{classical}}(m : e) > 0$ and $\mathrm{CI}_{\mathrm{classical}}(m : e \cup k) > 0$, which is *not* an example of non-local causal influence.

So why did we choose $\sigma_{\text{key}} = \mathbf{1}/2^n$? We did this because in the context of the protocol, a putative eavesdropper has a uniform prior on the state of the key, and so to her it is *as if* the key was in a maximally mixed state. But this is a reflection of the eavesdropper's particular knowledge, and not the state of the universe in which she lives.

If the classical universe of the protocol starts in a pure state, it will remain in a pure state for all time, and so it would instead be correct to use $\sigma_{\text{key}} = |K\rangle\langle K|$ for some particular K. In such a universe, there can be no non-local causal influence. If the universe was, in fact, at least partially in a mixed state, then we could harness some of the randomness to produce something like $\sigma_{\text{key}} = \mathbf{1}/2^n$.

Now we summarize the key point. In the classical setting, if the global state of the system is

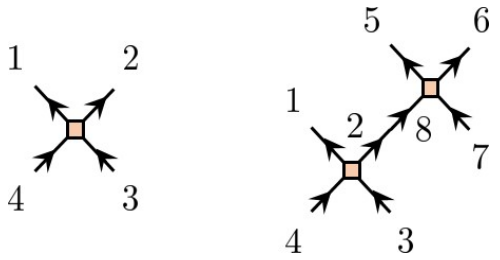


Figure 4.27: Ordering of links in two simple geometries: one rank-four tensor and two rank-four tensors with a pair of links contracted.

pure (i.e., not a probabilistic mixture), then the state of any subsystem is likewise pure. This is emphatically not the case for a quantum system due to entanglement, and so subsystems of a pure quantum state are often mixed states. If a classical universe starts in a pure classical state which remains pure and classical for all time, then there cannot be non-local causal influence with respect to subsystems. However, if a quantum universe starts in a pure quantum state which is pure for all time, then there can be non-local causal influence with respect to subsystems.

4.11 Appendix D: Numerics for stabilizer tensor networks

Here we review stabilizer tensor networks, and explain how we implement numerical calculations of these networks as discussed in Section 4.6.

To begin, stabilizer tensor networks are tensor networks comprised of connected unit stabilizer codes. Each unit stabilizer code is a tensor defined as the state fixed by a set of operators (stabilizers). Pictorially, a tensor can be represented as a vertex, and there is a Hilbert space on each link. The basic units we consider here are rank-four qutrit codes, i.e., there is a three-dimensional Hilbert space associated with each link and each vertex is degree four. The space of operators on each three-dimensional Hilbert space has a complex basis $X^n Z^m$ where $n, m = 0, 1, 2$, and

$$X = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i2\pi/3} & 0 \\ 0 & 0 & e^{i4\pi/3} \end{pmatrix}, \tag{4.121}$$

in a preferred basis $\{|0\rangle, |1\rangle, |2\rangle\}$ of the Hilbert space. Note that $XZ = \exp(i2\pi/3)ZX$, and so the basis operators $X^n Z^m$ all commute up to phases. Stabilizer operators are products of such basis operators, for example, $X \otimes I \otimes X \otimes I$, where operators on different links are separated by $\otimes$ and links are ordered as in Figure 4.27.

A more convenient notation for stabilizer operators would be vectors with elements in $\mathbb{F}_3$, i.e., the field of three elements. For example, stabilizer operators for the rank-four swap code can be

written as (denote $\overline{X} = X^2 = X^{-1}$ and $\overline{Z} = Z^2 = Z^{-1}$)

$$\left(\begin{array}{c|cccccccc} 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 2 \end{array}\right) \Leftrightarrow \left(\begin{array}{c} X \otimes I \otimes X \otimes I \\ I \otimes X \otimes I \otimes X \\ Z \otimes I \otimes \overline{Z} \otimes I \\ I \otimes Z \otimes I \otimes \overline{Z} \end{array}\right), \tag{4.122}$$

that is,

$$\left(\begin{array}{c|cccccccc} k & n_1 & m_1 & n_2 & m_2 & n_3 & m_3 & n_4 & m_4 \end{array} \right) \Leftrightarrow e^{i2\pi k/3} X^{n_1} Z^{m_1} \otimes X^{n_2} Z^{m_2} \otimes X^{n_3} Z^{m_3} \otimes X^{n_4} Z^{m_4}. \tag{4.123}$$

Indeed, it is easy to verify that the code

$$\sum_{i,j \in \mathbb{F}_3} |i\rangle \otimes |j\rangle \otimes |i\rangle \otimes |j\rangle \tag{4.124}$$

is (up to a multiplicative constant) the only state fixed by these four stabilizers given by the rows of Eqn. (4.122). If we regard this state as a unitary gate from links 3, 4 to 1, 2, it merely transports the state from link 3 to 1, and from link 4 to 2, hence is called a "swap" gate. The dynamics of multiple catenated and layered swap gates simply propagates qutrits along diagonal lines in the stabilizer tensor network, and so clearly corresponds to integrable time evolution.

Another code that we use is the $[[4, 0, 3]]$ perfect code where the state is (note that division is in $\mathbb{F}_3$)

$$\sum_{i,j \in \mathbb{F}_3} |i\rangle \otimes |j\rangle \otimes |(i-j)/2\rangle \otimes |(i+j)/2\rangle, \tag{4.125}$$

corresponding to a set of stabilizers

$$\left(\begin{array}{c|cccccccc} 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \end{array}\right) \Leftrightarrow \left(\begin{array}{c} Z \otimes Z \otimes I \otimes Z \\ Z \otimes \overline{Z} \otimes Z \otimes I \\ X \otimes X \otimes I \otimes X \\ X \otimes \overline{X} \otimes X \otimes I \end{array}\right). \tag{4.126}$$

Of course the full set of stabilizer operators of this code should contain products of these operators as well, so the choice of four generating operators is not unique.

Now we proceed to finding stabilizers for networks composed of simple rank-four tensors. As an example, consider contracting two swap codes (identifying links 2 and 8 as in Figure 4.27). Taking the product of operators corresponds to addition in the vector notation, so a general stabilizer (up

to phase factors) of two swap codes takes the form

$$\begin{pmatrix} a_1 & a_2 & a_3 & a_4 & b_1 & b_2 & b_3 & b_4 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 2 \end{pmatrix}, \tag{4.127}$$

where we have temporarily suppressed the prefactor column for simplicity. The stabilizers on contracted links should cancel to give an operator acting on the remaining links only. Specifically, if the stabilizer on link 2 is $X^n Z^m$, then the stabilizer on link 8 must be $X^n Z^{-m}$. To find such solutions, only columns 3, 4 (link 2) and columns 15, 16 (link 8) in the matrix are relevant. The algebraic equation in $\mathbb{F}_3$ is thus

$$\begin{pmatrix} a_1 & a_2 & a_3 & a_4 & b_1 & b_2 & b_3 & b_4 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & 1 \end{pmatrix} = 0, \tag{4.128}$$

and the solution is $a_2 = b_2$, $a_4 = b_4$ and $a_i, b_i \in \mathbb{F}_3$ for $i = 1, 2, 3, 4$, i.e., the row space of

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}. \tag{4.129}$$

Hence a generating set of stabilizers is the product of (4.129) with (4.127) (with columns corresponding

to contracted links dropped)

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 2 \end{pmatrix} \Leftrightarrow \begin{pmatrix} X \otimes X \otimes I \otimes I \otimes I \otimes I \\ I \otimes I \otimes X \otimes I \otimes X \otimes I \\ Z \otimes \overline{Z} \otimes I \otimes I \otimes I \otimes I \\ I \otimes I \otimes \overline{Z} \otimes I \otimes Z \otimes I \\ I \otimes I \otimes I \otimes X \otimes I \otimes X \\ I \otimes I \otimes I \otimes Z \otimes I \otimes \overline{Z} \end{pmatrix}, \tag{4.130}$$

which are indeed stabilizers for

$$\sum_{i,j,k \in \mathbb{F}_3} |i\rangle \otimes |i\rangle \otimes |j\rangle \otimes |k\rangle \otimes |j\rangle \otimes |k\rangle. \tag{4.131}$$

Intuitively, this code transports states from link 3 to 1, 4 to 6 and 7 to 5.

For general codes, phase factors must be taken into account when multiplying operators. Addition rules for phases are modified due to the non-commutativity of X and Z operators. For each link,

$$X^n Z^m \times X^{n'} Z^{m'} = e^{-i2\pi mn'/3} X^{n+n'} Z^{m+m'}, \tag{4.132}$$

that is,

$$\left(\begin{array}{c|cc} k & n & m \end{array} \right) + \left(\begin{array}{c|cc} k' & n' & m' \end{array} \right) = \left(\begin{array}{c|cc} k + k' - mn' & n + n' & m + m' \end{array} \right). \tag{4.133}$$

And the total phase is a sum of contributions from each link i:

$$\left(\begin{array}{c|cc} k & n_i & m_i \end{array} \right) + \left(\begin{array}{c|cc} k' & n'_i & m'_i \end{array} \right) = \left(\begin{array}{c|cc} k + k' - \sum_i m_i n'_i & n_i + n'_i & m_i + m'_i \end{array} \right). \tag{4.134}$$

Then determining stabilizers of the network is reduced to a linear algebra problem that can be solved in time polynomial in the network size. More specifically, the algorithm consists of three steps:

1. List the stabilizers of all constituent tensors;
2. Solve the linear equations imposed by requiring that operators on contracted links cancel;
3. Use the solution to the linear equations to find combinations of the stabilizers in step 1 that are the identity on the contracted links (taking into account the phase additions).

Given stabilizers $O_1, \ldots, O_n$, the state fixed by all stabilizers is then the eigenstate of $O_1 + \cdots + O_n$ with eigenvalue n because the spectrum of each operator O_i only contains values $\exp(i2\pi k/3)$, $k = 0, 1, 2$. The superdensity operator of stabilizer tensor networks with few-vertex insertions (as

shown in Figure 4.13) is then itself a stabilizer state which can be computed up to a prefactor in polynomial time. The prefactor can be fixed by requiring the trace of the superdensity operator to be one. Causal influence is evaluated according to Eqn. (4.75) using the superdensity operator, which produces Figure 4.25.

Chapter 5

Deep Quantum Geometry of Matrices

This chapter is essentially the same as

- Han, Xizhi, and Sean A. Hartnoll. "Deep quantum geometry of matrices." *Physical Review X* 10.1 (2020): 011069.

Abstract

We employ machine learning techniques to provide accurate variational wavefunctions for matrix quantum mechanics, with multiple bosonic and fermionic matrices. Variational quantum Monte Carlo is implemented with deep generative flows to search for gauge invariant low energy states. The ground state, and also long-lived metastable states, of an SU(N) matrix quantum mechanics with three bosonic matrices, as well as its supersymmetric 'mini-BMN' extension, are studied as a function of coupling and N. Known semiclassical fuzzy sphere states are recovered, and the collapse of these geometries in more strongly quantum regimes is probed using the variational wavefunction. We then describe a factorization of the quantum mechanical Hilbert space that corresponds to a spatial partition of the emergent geometry. Under this partition, the fuzzy sphere states show a boundary-law entanglement entropy in the large N limit.

5.1 Introduction

A quantitative, first principles understanding of the emergence of spacetime from non-geometric microscopic degrees of freedom remains among the key challenges in quantum gravity. Holographic duality has provided a firm foundation for attacking this problem; we now know that supersymmetric large N matrix theories can lead to emergent geometry [164, 202]. What remains is the technical

challenge of solving these strongly quantum mechanical systems and extracting the emergent spacetime dynamics from their quantum states. Recent years have seen significant progress in numerical studies of large N matrix quantum mechanics at nonzero temperature. Using Monte Carlo simulations, quantitatively correct features of emergent black hole geometries have been obtained, e.g. [7, 49, 25]. To grapple with questions such as the emergence of local spacetime physics, and its associated short distance entanglement [33, 235], new and inherently quantum mechanical tools are needed.

Variational wavefunctions can capture essential aspects of low energy physics. However, the design of accurate many-body wavefunction ansatze has typically required significant physical insight. For example, the power of tensor network states, such as Matrix Product States, hinges upon an understanding of entanglement in local systems [198, 192]. We are faced, in contrast, with models where there is an emergent locality that is not manifest in the microscopic interactions. This locality cannot be used *a priori*; it must be uncovered. Facing a similar challenge of extracting the most relevant variables in high-dimensional data, deep learning has demonstrated remarkable success [105, 140, 81], in tasks ranging from image classification [137] to game playing [233]. These successes, and others, have motivated tackling many-body physics problems with the machine learning toolbox [59]. For example, there has been much interest and progress in applications of Restricted Boltzmann Machines to characterize states of spin systems [44, 65, 78, 80].

In this work we solve for low-energy states of quantum mechanical Hamiltonians with both bosons and fermions, using generative flows (normalizing flows [66, 125, 67] and masked autoregressive flows [79, 130, 194] in particular) and variational quantum Monte Carlo. Compared with spin systems, the problem we are trying to solve contains continuous degrees of freedom and gauge symmetry, and there is no explicit spatial locality. Recent works have applied generative models to physics problems [45, 90, 255] and have aimed to understand holographic geometry, broadly conceived, with machine learning [258, 100, 110]. We will use generative flows to characterize emergent geometry in large N multimatrix quantum mechanics. As we have noted above, such models form the microscopic basis of established holographic dualities.

We will focus on quantum mechanical models with three bosonic large N matrices. These are among the simplest models with the core structure that is common to holographic theories. The bosonic part of the Hamiltonian takes the form

$$H_B = \mathrm{tr}\left(\frac{1}{2}\Pi^i\Pi^i - \frac{1}{4}[X^i, X^j][X^i, X^j] + \frac{1}{2}\nu^2 X^i X^i + i\nu\epsilon^{ijk}X^i X^j X^k\right). \tag{5.1}$$

Here the X^i are N by N traceless Hermitian matrices, with $i = 1, 2, 3$. The Π^i are conjugate momenta and ν is a mass deformation parameter. The potential energy in (5.1) is a total square: $V(X) = \frac{1}{4}\mathrm{tr}\left[\left(\nu\epsilon^{ijk}X^k + i[X^i, X^j]\right)^2\right]$. The supersymmetric extension of this model [12], discussed below, can be thought of as a simplified version of the BMN matrix quantum mechanics [23]. We refer to the supersymmetric model as 'mini-BMN', following [10]. For the low energy physics we will

be exploring, the large N planar diagram expansion in this model is controlled by the dimensionless coupling $\lambda \equiv N/\nu^3$. Here λ can be understood as the usual dimensionful 't Hooft coupling of a large N quantum mechanics at an energy scale set by the mass term (cf. [119]).

The mass deformation in the Hamiltonian (5.1) inhibits the spatial spread of wavefunctions — which will be helpful for numerics — and leads to minima of the potential at

$$[X^i, X^j] = i\nu\epsilon^{ijk}X^k \,. \tag{5.2}$$

In particular, one can have $X^i = \nu J^i$ with the J^i being, for example, the N dimensional irreducible representation of the $\mathfrak{su}(2)$ algebra. This set of matrices defines a 'fuzzy sphere' [160]. There are two important features of this solution. Firstly, in the large N limit the noncommutative algebra generated by the X^i approaches the commutative algebra of functions on a smooth two dimensional sphere [107, 62]. Secondly, the large ν limit is a semiclassical limit in which the classical fuzzy sphere solution accurately describes the quantum state. In this semiclassical limit, the low energy excitations above the fuzzy sphere state are obtained from classical harmonic perturbations of the matrices about the fuzzy sphere [121]. See also [60] for an analogous study of the large-mass BMN theory. At large N and ν, these excitations describe fields propagating on an emergent spatial geometry.

By using variational Monte Carlo with generative flows we will obtain a fully quantum mechanical description of this emergent space. This, in itself, is excessive given that the physics of the fuzzy sphere is accessible to semiclassical computations. Our variational wavefunctions will quantitatively reproduce the semiclassical results in the large ν limit, thereby providing a solid starting point for extending the variational method across the entire N and ν phase diagram. Exploring the parameter space, we find that the fuzzy sphere collapses upon moving into the small ν, quantum regime. We will consider two different 'sectors' of the model, with different fermion number R. The first will be purely bosonic states, with $R = 0$. The second will have a $R = N^2 - N$. In this latter sector, the fuzzy sphere state is supersymmetric at large positive ν, so we refer to this as the 'supersymmetric sector'. In the bosonic sector of the model the fuzzy sphere is a metastable state, and collapses in a first order large N transition at $\nu \sim \nu_c \approx 4$. See Figs. 5.2 and 5.3 below. In the supersymmetric sector of the model, where the fuzzy sphere is stable, the collapse is found to be more gradual. See Figs. 5.6 and 5.7. In Fig. 5.8 we start to explore the small ν limit of the supersymmetric sector.

Beyond the energetics of the fuzzy sphere state, we will define a factorization of the microscopic quantum mechanical Hilbert space that leads to a boundary-law entanglement entropy at large ν. See (5.48) below. This factorization at once captures the emergent local dynamics of fields on the fuzzy sphere and also reveals a microscopic cutoff to this dynamics at a scale set by N. The nature of the emergent fields and their cutoff can be usefully discussed in string theory realizations of the model. In string-theoretic constructions, fuzzy spheres arise from the polarization of D branes in background fields [175, 3, 168, 176]. A matrix quantum mechanics theory such as (5.1) describes

N 'D0 branes' — see [12] and the discussion section below for a more precise characterization of the string theory embedding of mini-BMN theory — and the maximal fuzzy sphere corresponds to a configuration in which the D0 branes polarize into a single spherical D2 brane. There is no gravity associated to this emergent space, the emergent fields describe the low energy worldvolume dynamics of the D2 brane. In this case, the emergent fields are a Maxwell field and a single scalar field corresponding to transverse fluctuations of the brane. In the final section of the paper we will discuss how richer, gravitating states may arise in the opposite small ν limit of the model.

5.2 The mini-BMN model

The mini-BMN Hamiltonian is [12]

$$H = H_B + \mathrm{tr}\left(\lambda^\dagger \sigma^k [X^k, \lambda] + \frac{3}{2}\nu \lambda^\dagger \lambda\right) - \frac{3}{2}\nu(N^2-1)\,. \tag{5.3}$$

The bosonic part H_B is given in (5.1). The σ^k are Pauli matrices. The λ are matrices of two-component SO(3) spinors. It can be useful to write the matrices in terms of the $\mathfrak{su}(N)$ generators T_A, with $A = 1, 2, \ldots, N^2-1$, which obey $[T_A, T_B] = i f_{ABC} T_C$ and are Hermitian and orthonormal (with respect to the Killing form). That is, $X^i = X^i_A T^A$ and $\lambda^\alpha = \lambda^\alpha_A T^A$.[1] The full Hamiltonian can then be written

$$\begin{aligned} H = &-\frac{1}{2}\frac{\partial^2}{(\partial X^i_A)^2} + \frac{1}{4}\left(f_{ABC} X^i_B X^j_C\right)^2 + \frac{1}{2}\nu^2 \left(X^i_A\right)^2 - \frac{1}{2}\nu f_{ABC}\epsilon^{ijk} X^i_A X^j_B X^k_C \\ &+ i f_{ABC} \lambda^{\alpha\dagger}_A X^k_B \sigma^{k\beta}_\alpha \lambda_{C\beta} + \frac{3}{2}\nu \lambda^{\alpha\dagger}_A \lambda_{A\alpha} - \frac{3}{2}\nu(N^2-1), \end{aligned} \tag{5.4}$$

where $\lambda^{\alpha\dagger}_A \equiv (\lambda_{A\alpha})^\dagger$ and $\{\lambda^{\alpha\dagger}_A, \lambda_{B\beta}\} = \delta_{AB}\delta^\alpha_\beta$ are complex fermion creation and annihilation operators. This Hamiltonian is seen to have four supercharges

$$Q_\alpha = \left(-i\frac{\partial}{\partial X^i_A} + i\nu X^i_A - \frac{i}{2} f_{ABC}\epsilon_{ijk} X^j_B X^k_C\right)\sigma^{i\beta}_\alpha \lambda_{A\beta}, \quad \bar{Q}^\alpha = (Q_\alpha)^\dagger\,, \tag{5.5}$$

that obey

$$\{Q_\alpha, \bar{Q}^\alpha\} = 4H. \tag{5.6}$$

States that are invariant under all supercharges therefore have vanishing energy.

Matrix quantum mechanics theories arising from microscopic string theory constructions are typically gauged. This means that physical states must be invariant under the SU(N) symmetry.

[1] The ijk and ABC indices are freely raised and lowered. Lower $\alpha\beta$ indices are for spinors transforming in the **2** representation of SO(3), while upper indices are for $\bar{\mathbf{2}}$. We will not raise or lower spinor indices.

In particular, physical state are annihilated by the generators

$$G_A = -if_{ABC}\left(X^i_B\frac{\partial}{\partial X^i_C} + \lambda^{\alpha\dagger}_B\lambda_{C\alpha}\right). \tag{5.7}$$

5.2.1 Representation of the fermion wavefunction

The mini-BMN wavefunction can be represented as a function from bosonic matrix coordinates to fermionic states $\psi(X) = f(X)|M(X)\rangle$. Here X denotes the three bosonic traceless Hermitian matrices. The function $f(X) \geq 0$ is the norm of the wavefunction at X while $|M(X)\rangle$ is a normalized state of matrix fermions. A fermionic state with definite fermion number R is parametrized by a complex tensor $M^{ra}_{A\alpha}$ such that

$$|M\rangle \equiv \sum_{r=1}^{D}\prod_{a=1}^{R}\Big(\sum_{\alpha=1}^{2}\sum_{A=1}^{N^2-1}M^{ra}_{A\alpha}\lambda^{\alpha\dagger}_A\Big)|0\rangle, \tag{5.8}$$

where $|0\rangle$ is the state with all fermionic modes unoccupied.

The definition (5.8) is parsed as follows: for any fixed r and a, $\eta^{ra\dagger} = \sum_{\alpha A} M^{ra}_{A\alpha}\lambda^{\alpha\dagger}_A$ is the creation operator for the matrix fermionic modes, where A runs over some orthonormal basis of the $\mathfrak{su}(N)$ Lie algebra and $\alpha = 1, 2$ for two fermionic matrices. Then $\prod_a \eta^{ra\dagger}|0\rangle$ is a state of multiple free fermions created by $\eta^\dagger$. The final summation over r in (5.8) is a decomposition of a general fermionic state into a sum of free fermion states. Such a representation is seen to be completely general (but not unique) if we have the number of free fermion states D sufficiently large.

For purely bosonic models, $|M(X)\rangle$ is simply the phase of the wavefunction.

5.2.2 Gauge invariance and gauge fixing

The generators (5.7) correspond to the following action of an element $U \in G = \mathrm{SU}(N)$ on the wavefunction:

$$(U\psi)(X) = f(U^{-1}XU)|(UMU^{-1})(U^{-1}XU)\rangle, \tag{5.9}$$

that is, the group acts by matrix conjugation. The wavefunction is required to be invariant under the group action, i.e. $U\psi = \psi$ for any $U \in G$.

Gauge invariance allows us to evaluate the wavefunction using a representative for each orbit of the gauge group. Let $\widetilde{X}$ be the representative in the gauge orbit of X. Gauge invariance of the wavefunction implies that there must exist functions $\widetilde{f}$ and $\widetilde{M}$ such that

$$f(X) = \widetilde{f}(\widetilde{X}), \quad |M(X)\rangle = |U\widetilde{M}(\widetilde{X})U^{-1}\rangle \text{ where } X = U\widetilde{X}U^{-1}. \tag{5.10}$$

The functions $\widetilde{f}$ and $\widetilde{M}$ take gauge representatives as inputs, or may be thought as gauge invariant functions. The wavefunction we use will be in the form (5.10). The functions $\widetilde{f}$ and $\widetilde{M}$ will be

parametrized by neural networks, as we describe in the following section 5.3.

We proceed to describe the gauge fixing we use to select the representative for each orbit, as well as the measure factor associated with this choice. The SU(N) gauge representative $\widetilde{X}$ will be such that

1. $X^i = U\widetilde{X}^i U^{-1}$ for $i = 1, 2, 3$ and some unitary matrix U.
2. $\widetilde{X}^1$ is diagonal and $\widetilde{X}^1_{11} \leq \widetilde{X}^1_{22} \leq \ldots \leq \widetilde{X}^1_{NN}$.
3. $\widetilde{X}^2_{i(i+1)}$ is purely imaginary with the imaginary part positive for $i = 1, 2, \ldots, N-1$.

The third condition is needed to fix the $\mathrm{U}(1)^{N-1}$ residual gauge freedom after diagonalizing X^1. The representative $\widetilde{X}$ is well-defined except on a subspace of measure zero where the matrices are degenerate. Then $\widetilde{X}$ can be represented as a vector in $\mathbb{R}^{2(N^2-1)}$ with a positivity constraint on some components. The change of variables from X to $\widetilde{X}$ leads to a measure factor given by the volume of the gauge orbit:

$$d^{3(N^2-1)}X = \Delta(\widetilde{X})\, d^{2(N^2-1)}\widetilde{X}\,, \tag{5.11}$$

with

$$\Delta(\widetilde{X}) \propto \prod_{i\neq j=1}^{N} \left|\widetilde{X}^1_{ii} - \widetilde{X}^1_{jj}\right| \prod_{i=1}^{N-1} \left|\widetilde{X}^2_{i(i+1)}\right|. \tag{5.12}$$

Keeping track of this measure (apart from an overall prefactor) will be important for proper sampling in the Monte Carlo algorithm. The derivation of (5.12) is shown in Appendix 5.7.

5.3 Architecture design for matrix quantum mechanics

In this work we propose a variational Monte Carlo method with importance sampling to approximate the ground state of matrix quantum mechanics theories, leading to an upper bound on the ground state energy. The importance sampling is implemented with generative flows. The basic workflow is sketched as follows:

1. Start with a wavefunction ψ_θ with variational parameters θ. In our case θ will characterize neural networks.
2. Write the expectation value of the Hamiltonian to be minimized as

$$E_\theta = \langle\psi_\theta|H|\psi_\theta\rangle = \int dX\, |\psi_\theta(X)|^2 H_X[\psi_\theta] = \mathbb{E}_{X\sim|\psi_\theta|^2}[H_X[\psi_\theta]]\,. \tag{5.13}$$

   In the mini-BMN case X denotes three traceless Hermitian matrices (indices omitted) and $H_X[\psi_\theta]$ is the energy density at X. Notationally $\mathbb{E}_{X\sim p(X)}$ is the expectation value, with the random variable X drawn from the probability distribution $p(X)$.

3. Generate random samples according to the wavefunction probabilities $X \sim p_\theta(X) = |\psi_\theta(X)|^2$, and evaluate their energy densities $H_X[\psi_\theta]$. The variational energy (5.13) can then be estimated as the average of energy densities of the samples.

4. Update the parameters θ (via stochastic gradient descent) to minimize E_θ:

$$\theta_{t+1} = \theta_t - \alpha \nabla_{\theta_t} E_{\theta_t}, \tag{5.14}$$

where $t = 1, 2, \ldots$ denotes the steps of training and the parameter $\alpha > 0$ sets the learning rate. The gradient of energy is estimated from Monte Carlo samples:

$$\nabla_\theta E_\theta = \mathbb{E}_{X\sim p_\theta}[\nabla_\theta H_X[\psi_\theta]] + \mathbb{E}_{X\sim p_\theta}[\nabla_\theta (\ln p_\theta(X))\,(H_X[\psi_\theta] - E_\theta)]. \tag{5.15}$$

The method is applicable even if the probabilities are available only up to an unknown normalization factor.

5. Repeat steps 3 and 4 until E_θ converges. Observables of physical interest are evaluated with respect to the optimal parameters after training.

In the following we discuss details of parametrizing and sampling from gauge invariant wavefunctions with fermions. Technicalities concerning the evaluation of $H_X[\psi_\theta]$ are spelled out in Appendix 5.8. More details concerning the training are given in Appendix 5.10. Benchmarks are presented at the end of this section.

5.3.1 Parametrizing and sampling the gauge invariant wavefunction

We first describe how gauge invariance is incorporated into the variational Monte Carlo algorithm. As just discussed, an important step is to sample according to $X \sim |\psi(X)|^2$. From (5.10), for a gauge invariant wavefunction $|\psi(X)|^2 = |\widetilde{f}(\widetilde{X})|^2$. However, in sampling $\widetilde{X}$ we must keep track of the measure factor $\Delta(\widetilde{X})$ in (5.12). This is done as follows:

1. Sample $\widetilde{X}$ according to $p(\widetilde{X}) = \Delta(\widetilde{X})|\widetilde{f}(\widetilde{X})|^2$.

2. Generate Haar random elements $U \in \mathrm{SU}(N)$.

3. Output samples $X = U\widetilde{X}U^{-1}$.

The correctness of this procedure is shown in Appendix 5.7.

Conversely at the evaluation stage, $\psi(X)$ can be computed in the following steps for gauge invariant wavefunctions (5.10):

1. Gauge fix $X = U\widetilde{X}U^{-1}$ as discussed in the last section.

2. Compute $\widetilde{M}(\widetilde{X})$ and $\widetilde{f}(\widetilde{X})$. Details of the structure of $\widetilde{M}$ and $\widetilde{f}$ will be discussed below.

3. Return $\psi(X) = \widetilde{f}(\widetilde{X})|U\widetilde{M}(\widetilde{X})U^{-1}\rangle$ according to (5.10).

We now describe the implementation of $\widetilde{M}$ and $\widetilde{f}$ as neural networks. The basic building block, a multilayer fully-connected (also called dense) neural network, is an elemental architecture capable of parametrizing complicated functions efficiently [81]. The neural network defines a function $F : x \mapsto y$ mapping an input vector x to an output vector y via a sequence of affine and nonlinear transformations:

$$F = A_\theta^m \circ \tanh \circ A_\theta^{m-1} \circ \tanh \circ \cdots \circ \tanh \circ A_\theta^1 . \tag{5.16}$$

Here $A_\theta^1(x) = M_\theta^1 x + b_\theta^1$ is an affine transformation, where the weights M_θ^1 and the biases b_θ^1 are trainable parameters. The hyperbolic tangent nonlinearity then acts elementwise on $A_\theta^1(x)$.[2] Similar mappings are applied m times, allowing M_θ^i and b_θ^i to be different for different layers i, to produce the output vector y. The mapping $F : x \mapsto y$ is nonlinear and capable of approximating any square integrable function if the number of layers and the dimensions of the affine transformations are sufficiently large [156].

The function $\widetilde{M}(\widetilde{X})$ is implemented as such a multilayer fully-connected neural network, mapping from vectorized $\widetilde{X}$ to $\widetilde{M}$ in (5.8), i.e., $\mathbb{R}^{2(N^2-1)} \to \mathbb{R}^{DR\,2(N^2-1)}$. The implementation of $\widetilde{f}(\widetilde{X})$ is more interesting, as both evaluating $\widetilde{f}(\widetilde{X})$ and sampling from the distribution $p(\widetilde{X}) = \Delta(\widetilde{X})|\widetilde{f}(\widetilde{X})|^2$ are necessary for the Monte Carlo algorithm. Generative flows are powerful tools to efficiently parameterize and sample from complicated probability distributions. The function

$$\widetilde{f}(\widetilde{X}) = \sqrt{p(\widetilde{X})/\Delta(\widetilde{X})}, \tag{5.17}$$

so we can focus on sampling and evaluating $p(\widetilde{X})$, which will be implemented by generative flows.

Two generative flow architectures are implemented for comparison: a normalizing flow and a masked autoregressive flow. The normalizing flow starts with a product of simple univariate probability distributions $p(x) = p_1(x_1) \dots p_M(x_M)$, where the p_i can be different. Values of x sampled from this distribution are passed through an invertible multilayer dense network as in (5.16). The probability distribution of the output y is then

$$q(y) = p(x) \left| \det \frac{Dy}{Dx} \right|^{-1} = p(F^{-1}(y)) |\det DF|^{-1}. \tag{5.18}$$

The masked autoregressive flow generates samples progressively. It requires an ordering of the components of the input, say $x_1, x_2, \dots, x_M$. Each component is drawn from a parametrized distribution $p_i(x_i; F_i(x_1, \dots, x_{i-1}))$, where the parameter depends only on previous components. Thus x_1 is sampled independently and for other components, the dependence F_i is given by (5.16).

[2] We experimented with different activation functions; the final result is not sensitive to this choice.

The overall probability is the product

$$q(x) = \prod_{i=1}^{M} p_i(x_i; F_i(x_1, \ldots, x_{i-1})). \tag{5.19}$$

When $p_i(x_i)$ are chosen as normal distributions, both flows are able to represent any multivariate normal distribution exactly. Features of the wavefunction (such as polynomial or exponential tails) can be probed by experimenting with different base distributions $p_i(x_i)$. Choices of the base distributions and performances of the two flows are assessed in the following benchmark subsection and also in Appendix 5.10. We will use both types of flow in the numerical results of section 5.4.

5.3.2 Benchmarking the architecture

In [10] the Schrödinger equation for the $N = 2$ mini-BMN model was solved numerically. Comparison with the results in that paper will allow us to benchmark our architecture, before moving to larger values of N. In [10] the Schrödinger equation is solved in sectors with a fixed fermion number

$$R = \sum_{A\alpha} \lambda_A^{\alpha\dagger} \lambda_{A\alpha}, \qquad [R, H] = 0, \tag{5.20}$$

and total SO(3) angular momentum $j = 0, 1/2$. We do not constrain j, but do fix the number of fermions in the variational wavefunction.

The variational energies obtained from our machine learning architecture with $R = 0$ and $R = 2$ are shown as a function of ν in Fig. 5.1. We take negative ν to compare with the results given in [10], which uses an opposite sign convention.[3] The masked autoregressive flow yields better (lower) variational energies. These energies are seen to be close to the $j = 0$ results obtained in [10]. The variational results seem to be asymptotically accurate as $|\nu| \to \infty$, while remaining a reasonably good approximation at small ν. Small ν is an intrinsically more difficult regime, as the potential develops flat directions (visualized in [10]) and hence the wavefunction is more complicated, possibly with long tails. In the 'supersymmetric' $R = 2$ sector, where quantum mechanical effects at small ν are expected to be strongest, further significant improvement at the smallest values of ν is seen with deeper autoregressive networks and more flexible base distributions, as we describe shortly. Analogous improvements in these regimes will also be seen at larger N in Sec. 5.4.3 and Appendix 5.10.

In Fig. 5.1 the base distributions $p_i(x_i)$, introduced in the previous subsection, are chosen to be a mixture of s generalized normal distributions:

$$p_i(x_i) = \sum_{r=1}^{s} k_r^i \frac{\beta_r^i}{2\alpha_r^i \Gamma(1/\beta_r^i)} e^{-(|x_i - \mu_r^i|/\alpha_r^i)^{\beta_r^i}}, \quad \sum_{r=1}^{s} k_r^i = 1\,. \tag{5.21}$$

[3] There is a particle-hole symmetry of the Hamiltonian (5.4) via $\nu \to -\nu$, $\lambda \to \lambda^\dagger$, $\lambda^\dagger \to \lambda$ and $X \to -X$.

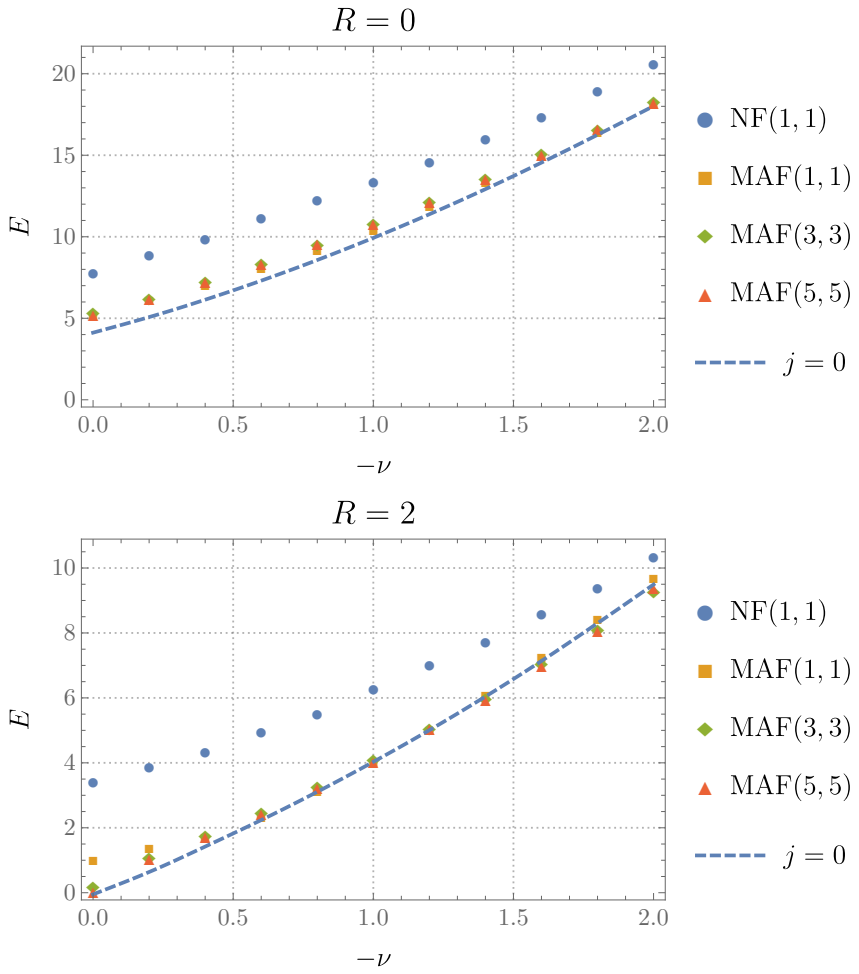


Figure 5.1: **Benchmarking the architecture:** Variational ground state energies for the mini-BMN model with $N = 2$ and fermion numbers $R = 0$ and $R = 2$ (shown as dots) compared to the exact ground state energy in the $j = 0$ sector, obtained in [10] (shown as the dashed curve). Uncertainties are at or below the scale of the markers; in particular the variational energies slightly below the dashed line are within numerical error of the line. NF stands for normalizing flows and MAF for masked autoregressive flows. As described in the main text, the numbers in the brackets are firstly the number of layers in the neural networks, and secondly the number of generalized normal distributions in each base mixed distribution.

Here the k_r^i are positive weights for each generalized normal distribution in the mixture. In (5.21) the k_r^i, α_r^i, β_r^i and μ_r^i are learnable (i.e. variational) parameters. For autoregressive flows these parameters further depend on x_j, with $1 \leq j < i$, according to (5.16).

Due to the gauge fixing conditions 2 and 3 in section 5.2.2, some components x_i are constrained to be positive. In the normalization flow this is implemented by an additional map $x_i \mapsto \exp(x_i)$. For the autoregressive flows we have a more refined control over the base distributions; in this case, for components x_i that must be positive, we draw from Gamma distributions instead:

$$p_i(x_i > 0) = \sum_{r=1}^{s} k_r^i \frac{(\beta_r^i)^{\alpha_r^i}}{\Gamma(\alpha_r^i)} (x_i)^{\alpha_r^i - 1} e^{-\beta_r^i x_i}, \quad \sum_{r=1}^{s} k_r^i = 1. \tag{5.22}$$

Where again the k_r^i, α_r^i and β_r^i depend on x_j, with $1 \leq j < i$, according to (5.16).

In Fig. 5.1 we have shown mixtures with $s = 1, 3, 5$ distributions. The number of layers in (5.16) has been increased with s to search for potential improvements in the space of variational wavefunctions. As noted, the only improvement within the autoregressive flows in going beyond one layer and one generalized normal distribution is seen at the smallest values of ν with $R = 2$. On the other hand, the gap between the variational energies of the two types of flows in Fig. 5.1 suggests that the wavefunction is complicated in this regime, so that the more sophisticated MAF architecture shows an advantage. The recursive nature of the MAF flows means that they are already 'deep' with only a single layer. The complexity of the small ν wavefunction should be contrasted with the fuzzy sphere phase at large positive ν discussed in the following section 5.4 and shown in e.g. Figs. 5.2 and 5.3 below. The wavefunction in this semiclassical regime is almost Gaussian, and indeed the NF(1, 1) and MAF(1, 1) flows give similar energies when initialized near fuzzy sphere configurations. The NF architecture in fact gives slightly lower energies in this regime, so we have used normalizing flows in Figs. 5.2 and 5.3 for the fuzzy sphere.

The numerics above and below are performed with $D = 4$ in (5.8), so that the fermionic wavefunction $|M(X)\rangle$ is a sum of four free fermion states for each value of the bosonic coordinates X. In Appendix 5.10 we see that increasing D above one lowers the variational energy at small ν, indicating that the fermionic states are not Hartree-Fock in this regime.

5.4 The emergence of geometry

5.4.1 Numerical results, bosonic sector

The architecture described above gives a variational wavefunction for low energy states of the mini-BMN model. With the wavefunction in hand, we can evaluate observables. We will start with the purely bosonic sector of the model (i.e. $R = 0$). Then we will add fermions. An important difference between the bosonic and supersymmetric cases will be that the semiclassical fuzzy sphere state is

metastable in the bosonic theory but stable in the supersymmetric theory.

Figure 5.2 shows the expectation value of the radius

$$r = \sqrt{\frac{1}{N}\mathrm{tr}(X_1^2 + X_2^2 + X_3^2)}\,, \tag{5.23}$$

for runs initialized close to a fuzzy sphere configuration (solid) and close to zero (open). For large ν a fuzzy sphere state with large radius is found, in addition to a 'collapsed' state without significant spatial extent. Below $\nu_c \approx 4$, the fuzzy sphere state ceases to exist. The nature of the transition at

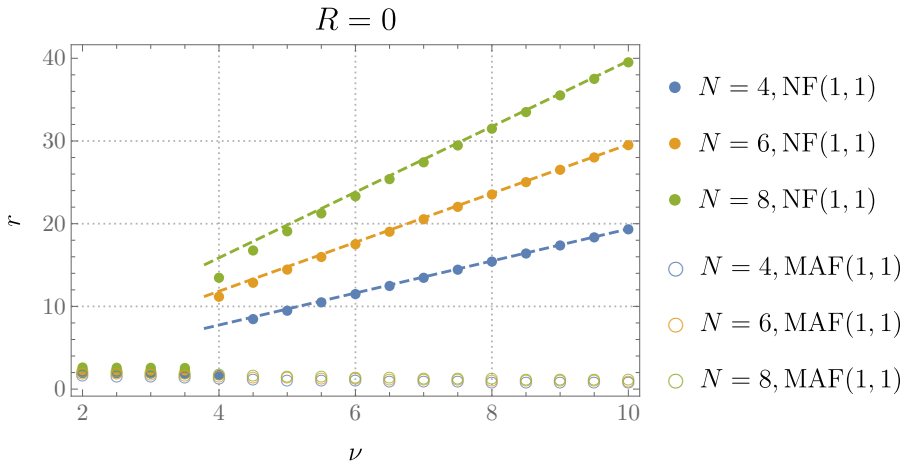


Figure 5.2: Expectation value of the radius in the zero fermion sector of the mini-BMN model, for different N and ν. The dashed lines are the semiclassical values (5.26). Solid dots are initialized near the fuzzy sphere configuration, and the open markers are initialized near zero. We have used normalizing and autoregressive flows, respectively, as these produce more accurate variational wavefunctions in the two different regimes.

ν_c can be understood from the variational energy of the states, plotted in Figure 5.3. The bosonic semiclassical fuzzy sphere state is seen to be metastable at large ν, as the collapsed state has lower energy. For $\nu < \nu_c$ the fuzzy sphere is no longer even metastable. We will gain a semiclassical understanding of this transition in section 5.4.2 shortly.

Figures 5.2 and 5.3 show that the radius and energy of the fuzzy sphere state are accurately described by semiclassical formulae (derived in the following section) for all $\nu > \nu_c$. In particular this means that E/N^3 and r/N are rapidly converging towards their large N values. Figure 5.4 further shows that the probability distribution for the radius r becomes strongly peaked about its semiclassical expectation value at large ν.

Analogous behavior to that shown in Figures 5.2 and 5.3 has previously been seen in classical Monte Carlo simulations of a thermal analogue of our quantum transition [14, 48, 64]. These papers study the thermal partition function of models similar to (5.1) in the classical limit, i.e. without

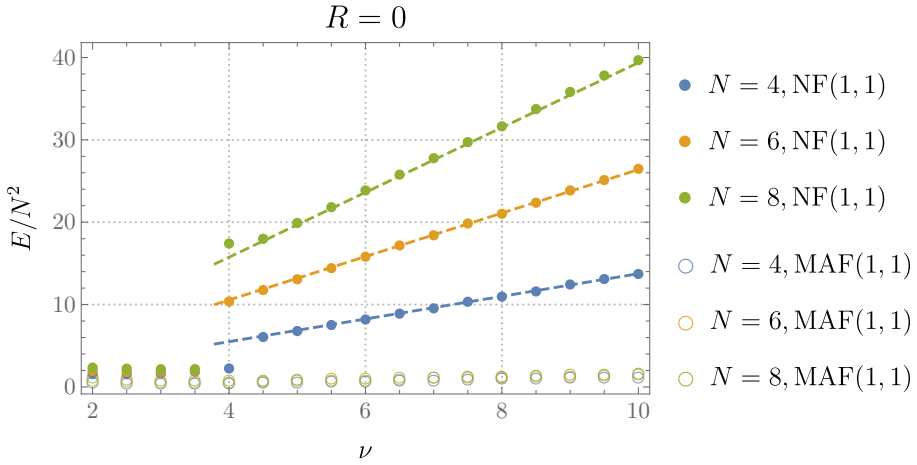


Figure 5.3: Variational energies in the zero fermion sector of the mini-BMN model, for different N and ν. The dashed lines are semiclassical values: $E = -\frac{3}{2}\nu(N^2-1) + \Delta E|_{\text{bos}}$, with $\Delta E|_{\text{bos}}$ given in (5.30). As in Fig. 5.2, solid dots are initialized near the fuzzy sphere configuration, and the open markers are initialized near zero.

the Π^2 kinetic energy term. The fuzzy geometry emerges in a first order phase transition as a low temperature phase in these models. We will see that in our quantum mechanical context the geometric phase is associated with the presence of a specific boundary-law entanglement.

5.4.2 Semiclassical analysis of the fuzzy sphere

The results above describe the emergence of a (metastable) geometric fuzzy sphere state at $\nu > \nu_c$. In this section we recall that in the $\nu \to \infty$ limit the fluctuations of the geometry are classical fields. For finite $\nu > \nu_c$ the background geometry is well-defined at large N, but fluctuations will be described by an interacting (noncommutative) quantum field theory.

In the large ν limit, the wavefunction can be described semiclassically [121, 60]. We will now briefly review this limit, with details given in the Appendix 5.9. These results provide a further useful check on the numerics, and will guide our discussion of entanglement in the following section 5.5.

The minima of the classical potential occur at:

$$[X^i, X^j] = i\nu\epsilon^{ijk}X^k \,. \tag{5.24}$$

These are supersymmetric solutions of the classical theory, annihilated by the supercharges (5.5) in the classical limit, and therefore have vanishing energy. The solutions of equations (5.24) are

$$X^i = \nu J^i \,, \tag{5.25}$$

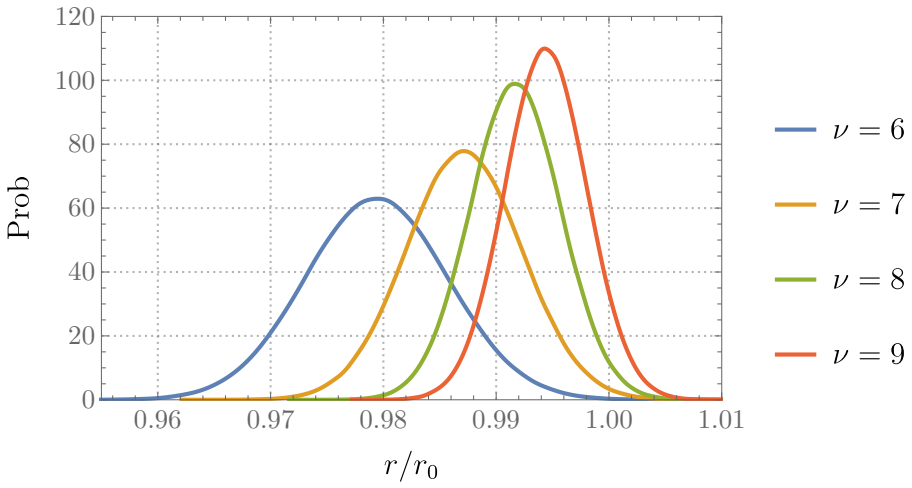


Figure 5.4: Probability distribution, from the variational wavefunction, for the radius in the fuzzy sphere phase for $N = 8$ and different ν. The horizontal axis is rescaled by the semiclassical value of the radius r_0, given in (5.26) below. The width of the distribution in units of the classical radius becomes smaller as ν is increased.

where the J^i are representations of the $\mathfrak{su}(2)$ algebra, $[J^i, J^j] = i\epsilon^{ijk}J^k$. We will be interested here in maximal, N-dimensional irreducible representations. (Reducible representations can also be studied, corresponding to multiple polarized D branes.)

The $\mathfrak{su}(2)$ Casimir operator suggests a notion of 'radius' given by

$$r^2 = \frac{1}{N}\sum_{i=1}^{3} \mathrm{tr}(X^i)^2 = \frac{\nu^2(N^2-1)}{4}. \tag{5.26}$$

Indeed, the algebra generated by the X^i matrices tends towards the algebra of functions on a sphere as $N \to \infty$ [107, 62]. At finite N, a basis for this space of matrices is provided by the matrix spherical harmonics $\hat{Y}_{jm}$. These obey

$$\sum_{i=1}^{3}[J^i,[J^i,\hat{Y}_{jm}]] = j(j+1)\hat{Y}_{jm}, \qquad [J^3,\hat{Y}_{jm}] = m\hat{Y}_{jm}\,. \tag{5.27}$$

We construct the $\hat{Y}_{jm}$ explicitly in Appendix 5.9. The j index is restricted to $0 \le j \le j_{\max} = N-1$. The space of matrices therefore defines a regularized or 'fuzzy' sphere [160].

Matrix spherical harmonics are useful for parametrizing fluctuations about the classical state (5.25). Writing

$$X^i = \nu J^i + \sum_{jm} y^i_{jm}\hat{Y}_{jm}\,, \tag{5.28}$$

the classical equations of motion can be perturbed about the fuzzy sphere background to give linear

equations for the parameters y^i_{jm}. The solutions of these equations define the classical normal modes. We find the normal modes in Appendix 5.9, proceeding as in [121, 60]. The normal mode frequencies are found to be $\nu\omega$ with

$$\begin{aligned} \omega^2 &= 0 \qquad \text{multiplicity } N^2-1\,, \\ \omega^2 &= j^2 \qquad \text{multiplicity } 2(j-1)+1\,, \\ \omega^2 &= (j+1)^2 \qquad \text{multiplicity } 2(j+1)+1\,. \end{aligned} \tag{5.29}$$

Recall that $1 \le j \le j_{\max} = N-1$. The three different sets of frequencies in (5.29) correspond to the group theoretic $\mathfrak{su}(2)$ decomposition $j \otimes 1 = (j-1) \oplus j \oplus (j+1)$. Here j is the 'orbital' angular momentum and the 1 is due to the vector nature of the X^i. We will give a field theoretic interpretation of these modes shortly. The modes give the following semiclassical contribution to the energy of the fuzzy sphere state

$$\Delta E|_{\rm bos} = \frac{|\nu|}{2}\sum|\omega| = \frac{4N^3+5N-9}{6}|\nu|\,. \tag{5.30}$$

This energy is shown in Figure 5.3. The scaling as N^3 arises because there are N^2 oscillators, with maximal frequency of order N. This semiclassical contribution will be cancelled out in the supersymmetric sector studied in section 5.4.3 below.

The normal modes (5.29) can be understood by mapping the matrix quantum mechanics Hamiltonian onto a noncommutative gauge theory. The analogous mapping for the classical model has been discussed in [118]. We carry out this map in Appendix 5.9. The original Hamiltonian (5.1) becomes the following noncommutative U(1) gauge theory on a unit spatial S^2 (setting the sphere radius to one in the field theory description will connect easily to the quantized modes in (5.29)):

$$H = \nu\int d\Omega\,\left(\frac{1}{2}(\pi^i)^2+\frac{1}{4}(f^{ij})^2\right)+\text{const}\,. \tag{5.31}$$

The noncommutative star product $\star$ is defined in the Appendix and

$$f^{ij} \equiv i\left(L^i a^j - L^j a^i\right)+\epsilon^{ijk}a^k + i\sqrt{\frac{4\pi}{N\nu^3}}[a^i,a^j]_\star\,, \tag{5.32}$$

where the derivatives generate rotations on the sphere $L^i = -i\epsilon_{ijk}x^j\partial_k$ and $[f,g]_\star \equiv f\star g - g\star f$. In (5.31) and (5.32) the vector potential a^i can be decomposed into two components tangential to the sphere, that become the two dimensional gauge field, and a component transverse to the sphere, that becomes a scalar field. This decomposition is described in Appendix 5.9. The normal modes (5.29) are coupled fluctuations of the gauge field and the transverse scalar field. The zero modes in (5.29) are pure gauge modes, given in (5.33) below. In (5.32) the effective coupling controlling quantum field theoretic interactions is seen to be $1/(N\nu)^{3/2}$. The extra $1/N$ arises because the commutator

$[a^i, a^j]_\star$ vanishes as $N \to \infty$, see Appendix 5.9. Corrections to the Gaussian fuzzy sphere state are therefore controlled by a different coupling than that of the 't Hooft expansion (recall $\lambda = N/\nu^3$).

The SU(N) gauge symmetry generators (5.7) are realized in an interesting way in the non-commutative field theory description. We see in Appendix 5.9 that upon mapping to non-commutative fields, the gauge transformations become

$$\delta a^i = -iL^i y - \sqrt{\frac{4\pi}{N\nu^3}}(\boldsymbol{n} \times \nabla y \cdot \nabla) a^i \,. \tag{5.33}$$

Here $\boldsymbol{n}$ is the normal vector and $y(\theta, \phi)$ a local field on the sphere. The first term in (5.33) is the usual U(1) transformation. The second term describes a coordinate transformation with infinitesimal displacement $\boldsymbol{n} \times \nabla y$. Indeed, it is known that non-commutative gauge theories mix internal and spacetime symmetries, which in this case are area-preserving diffeomorphisms of the sphere [193, 153]. The emergent U(1) non-commutative gauge theory thereby realizes the large N limit of the microscopic SU(N) gauge symmetry, as area-preserving diffeomorphisms [107, 62].

The fluctuation modes about the fuzzy sphere background allow a one-loop quantum effective potential for the radius to be computed in Appendix 5.9. The potential at $N \to \infty$ is shown in Fig. 5.5. At large ν the effective potential shows a metastable minimum at $r \sim N\nu/2$. For $\nu < \nu_{c,N=\infty}^{\text{1-loop}}$ this minimum ceases to exist. The large N, one-loop analysis therefore qualitatively reproduces the behavior seen in Figs. 5.2 and 5.3. The quantitative disagreement is mainly due to finite N corrections. The transition is only sharp as $N \to \infty$.

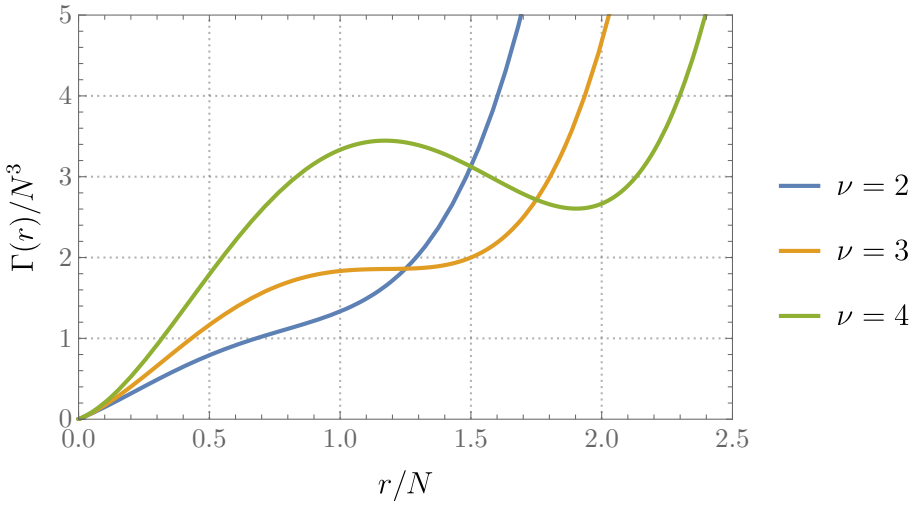


Figure 5.5: One-loop effective potential $\Gamma(r)$ for the radius of the bosonic ($R = 0$) fuzzy sphere as $N \to \infty$. The fuzzy sphere is only metastable when $\nu > \nu_{c,N=\infty}^{\text{1-loop}} \approx 3.03$, see Appendix 5.9.

5.4.3 Numerical results, supersymmetric sector

We now consider states with fermion number $R = N^2 - N$. The fuzzy sphere background is now supersymmetric at large positive ν [12]. The contribution of the fermions to the ground state energy is seen in Appendix 5.9 to cancel the bosonic contribution (5.30) at one loop:

$$-\frac{3}{2}\nu(N^2-1) + \Delta E|_{\text{fer}} + \Delta E|_{\text{bos}} = 0\,. \tag{5.34}$$

In Figure 5.6 the variational upper bound on the energy of the fuzzy sphere state remains close to zero for all values of ν. Figure 5.7 shows the radius as a function of ν. Probing the smallest values of ν requires a more powerful wavefunction ansatz than those of Figs. 5.6 and 5.7. We will consider that regime shortly.

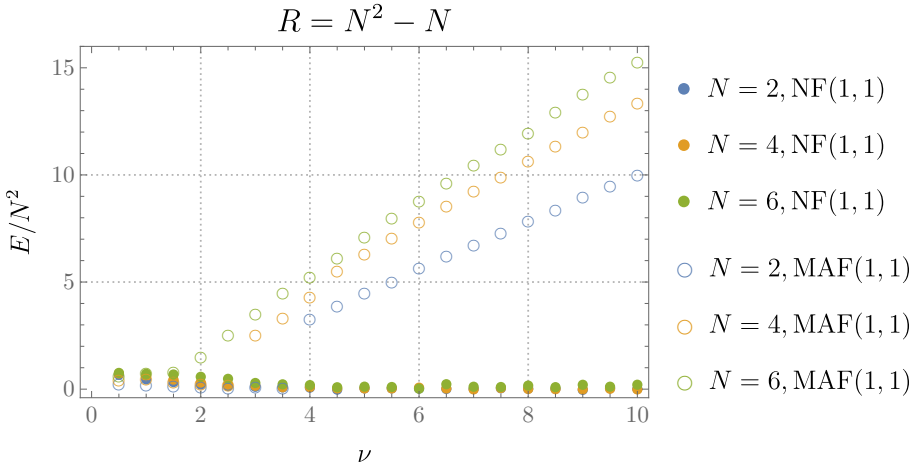


Figure 5.6: Variational energies in the SUSY sector of the mini-BMN model, for different N and ν. Solid dots are initialized near the fuzzy sphere configuration, and the open markers are initialized near zero. We are using normalizing and autoregressive flows, respectively, as these produce more accurate variational wavefunctions in the two different regimes.

In contrast to the states with zero fermion number in Figure 5.3, here the fuzzy sphere is seen to be the stable ground state at large ν. However, the fuzzy sphere appears to merge with the collapsed state below a value of ν that decreases with N. This is physically plausible: while the classical fuzzy sphere radius $r^2 \sim \nu^2 N^2$ decreases at small ν, quantum fluctuations of the collapsed state are expected to grow in space as $\nu \to 0$. This is because the flat directions in the classical potential of the $\nu = 0$ theory, given by commuting matrices, are not lifted in the presence of supersymmetry [63]. Eventually, the fuzzy sphere should be subsumed into these quantum fluctuations. This smoother large N evolution towards small ν (relative to the bosonic sector) is mirrored in the thermal behavior of classical supersymmetric models [6, 257].

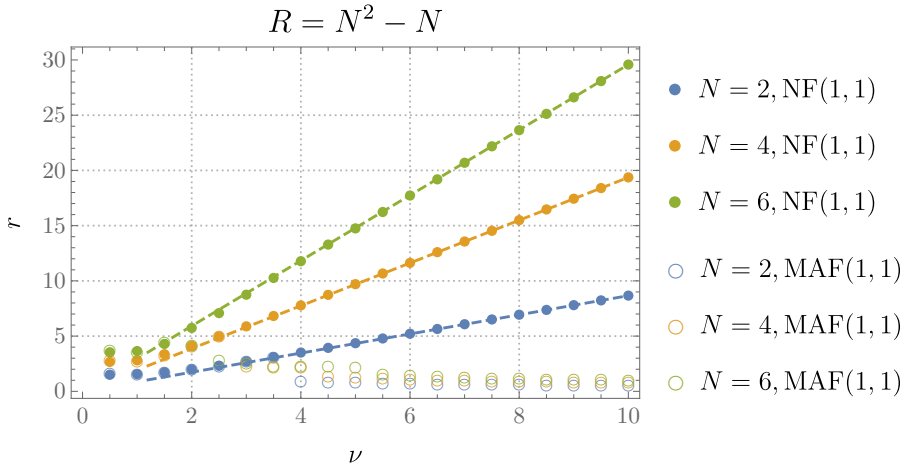


Figure 5.7: Expectation value of radius in the SUSY sector of the mini-BMN model, for different N and ν. Solid dots are initialized near the fuzzy sphere configuration, and the open markers are initialized near zero. The dashed lines are the semiclassical values (5.26).

Indeed, exploring the small ν region with more precision we observe a physically expected feature. In Fig. 5.8 we see that as ν decreases towards zero, the radius not only ceases to follow the semiclassical decreasing behavior, but turns around and starts to increase. The variance in the distribution of the radius is also seen to increase towards small ν, revealing the quantum mechanical nature of this regime. These behaviors (non-monotonicity of radius and increasing variance) are expected — and proven for $N = 2$ — because the flat directions of the classical potential at $\nu = 0$ mean that the extent of the wavefunction is set by purely quantum mechanical effects in this limit.

The small ν regime here is furthermore an opportunity to test the versatility of our variational ansatz away from semiclassical regimes. In Appendix 5.10 we see that for small ν MAFs achieve much lower energies than NFs. Increasing the number of distributions in the mixture and the number D of free fermions states in (5.8) further lowers the energy. These facts mirror the behavior we found in our $N = 2$ benchmarking in Sec. 5.3.2 at small ν, increasing our confidence in the ability of the network to capture this regime for large N also. The error in a variational ansatz is, as always, not controlled and therefore further exploration of this regime is warranted before very strong conclusions can be drawn. We plan to revisit this regime in future work, to search for the possible presence of emergent 'throat' geometries as we discuss in Sec. 5.6 below.

5.5 Entanglement on the fuzzy sphere

In this section we will see that the large ν fuzzy sphere state discussed above contains boundary-law entanglement. To compute the entanglement, one must first define a factorization of the Hilbert

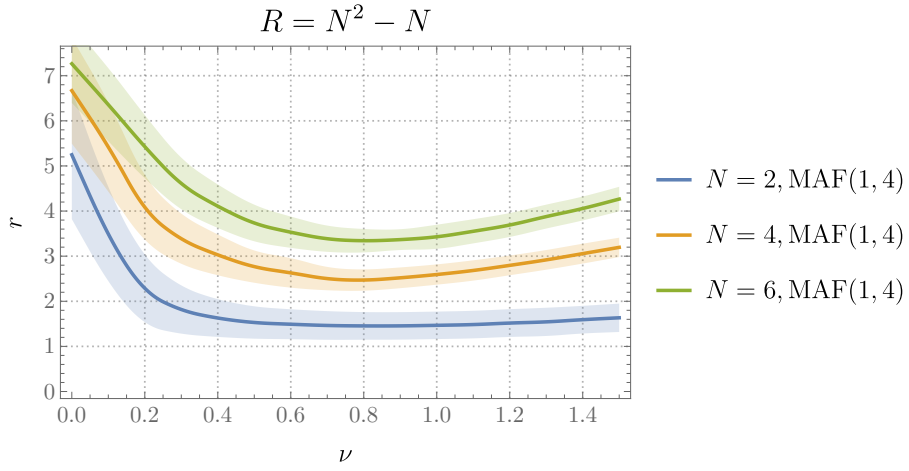


Figure 5.8: Distribution of radius for different N and small ν. Bands show the standard deviation of the quantum mechanical distribution of $r = \sqrt{\frac{1}{N}\sum \mathrm{tr} X_i^2}$, not to be confused with numerical uncertainty of the average. Recall that the numbers in the brackets are firstly the number of layers in the neural networks, and secondly the number of generalized normal distributions in each base mixed distribution.

space. For our emergent space at finite N and ν the geometry is both fuzzy and fluctuating, and hence lacks a canonical spatial partition. The fuzziness of the sphere is captured by a toy model of a free field on a sphere with an angular momentum cutoff. Recall from the previous section 5.4 that the noncommutative nature of the fuzzy sphere amounts to an angular momentum cutoff $j_{\max} = N - 1$. We will start, then, by defining a partition of the space of functions with such a cutoff.

5.5.1 Free field with an angular momentum cutoff

Consider a free massive complex scalar field $\varphi(\theta, \phi)$ on a unit two-sphere with the following Hamiltonian:

$$H = \int_{S^2} d\Omega \left[|\pi|^2 + |\nabla\varphi|^2 + \mu^2 |\varphi|^2 \right]. \tag{5.35}$$

Here π is the field conjugate to φ. We impose a cutoff $j \leq j_{\max}$ on the angular momentum, rending the quantum mechanical problem well-defined. The fields can therefore be decomposed into a sum of spherical harmonic modes:

$$\varphi(\theta, \phi) = \sum_{0 \leq j \leq j_{\max}}^{|m| \leq j} a_{jm} Y_{jm}(\theta, \phi). \tag{5.36}$$

The 'wavefunctional' of the quantum field $\varphi(\theta,\phi)$ is then a mapping from coefficients a_{jm} to complex amplitudes. The ground state wavefunctional of the Hamiltonian (5.35) is

$$\psi(a_{jm}) \propto e^{-\sum_{jm}\sqrt{j(j+1)+\mu^2}|a_{jm}|^2}. \tag{5.37}$$

To calculate entanglement for quantum states a factorization of the Hilbert space $\mathcal{H}=\mathcal{H}_1\otimes\mathcal{H}_2$ is prescribed. To motivate the construction of such a factorization in the fuzzy sphere case, we now review a general framework of defining entanglement in (factorizable) quantum field theories. In quantum mechanics, a quantum state is a function from the configuration space Q to complex numbers, and the Hilbert space of all quantum states is commonly the square integrable functions $\mathcal{H}=L^2(Q)$. In quantum field theories, the space Q is furthermore a linear space of functions on some geometric manifold M, and thus an orthogonal decomposition $Q=Q_1\oplus Q_2$ induces a factorization of $\mathcal{H}=L^2(Q_1)\otimes L^2(Q_2)$, which can be exploited to define entanglement.

To define entanglement it then suffices to find an orthogonal decomposition of the space of fields on the fuzzy sphere. Without an angular momentum cutoff, i.e. with $j_{\max}\to\infty$, there is a natural choice for any region A on the sphere, which sets Q_1 to be all functions supported on A, and Q_2 all functions supported on $\bar{A}$, the complement of A. Any function f on M can be uniquely written as a sum of $f_1\in Q_1$ and $f_2\in Q_2$, where $f_1=f\chi_A$ and $f_2=f(1-\chi_A)$. Here χ_A is the function on the sphere that is 1 on A and 0 otherwise. Note that the map of multiplication by χ_A, $f\mapsto f\chi_A$, acts as the projection $Q_1\oplus Q_2\to Q_1$. Conversely, given any orthogonal projection operator $P:Q\to Q$, we can decompose $Q=\mathrm{im}P\oplus\ker P$.

When the cutoff $j_{\max}$ is finite, multiplication by χ_A will generally take the function out of the subspace of functions with $j\le j_{\max}$. However, we can still do our best to approximate the projector P_A^∞ of multiplication by χ_A, as defined in the previous paragraph, with a projector $P_A^{j_{\max}}$ that lives in the subspace with $j\le j_{\max}$. Formally let $Q^{j_{\max}}$ be the space of functions on the sphere spanned by $Y_{jm}(\theta,\phi)$ with $j\le j_{\max}$. Define the orthogonal projector $P_A^{j_{\max}}:Q^{j_{\max}}\to Q^{j_{\max}}$ to minimize the distance $\|P_A^{j_{\max}}-P_A^\infty\|$. The projector $P_A^{j_{\max}}$ annihilates all functions in the orthogonal complement of $Q^{j_{\max}}$, when viewed as an operator acting on Q^∞. It is convenient to choose $\|\cdot\|$ to be the Frobenius norm, and in Appendix 5.11 an explicit formula for $P_A^{j_{\max}}$ is obtained.

The projector $P_A^{j_{\max}}$ then defines a factorization of the Hilbert space $L^2(Q^{j_{\max}})=L^2(\mathrm{im}\,P_A^{j_{\max}})\otimes L^2(\ker P_A^{j_{\max}})$ for any region A, and entanglement can be evaluated in the usual way. In particular, the second Rényi entropy of a pure state $|\psi\rangle$ on a region A is

$$\begin{aligned} S_2(\rho_A) &= -\ln\int dx_A dx_{\bar{A}} dx'_A dx'_{\bar{A}}\,\psi(x_A+x_{\bar{A}})\psi^*(x'_A+x_{\bar{A}})\psi(x'_A+x'_{\bar{A}})\psi^*(x_A+x'_{\bar{A}}) \\ &= -\ln\int dx dx'\,\psi(x)\psi^*(Px'+(I-P)x)\psi(x')\psi^*(Px+(I-P)x'), \end{aligned} \tag{5.38}$$

where $x_A=Px$ and $x_{\bar{A}}=(I-P)x$ are integrated over $\mathrm{im}\,P$ and $\ker P$, for $P=P_A^{j_{\max}}$, and x_A and

$x_{\bar{A}}$ can be more compactly combined into a field x with $j \leq j_{\max}$. Note that the various x's in (5.38) denote functions on the sphere.

The projector $P_A^{j_{\max}}$ is found to have two important geometric features:

1. The trace of the projector, which counts the number of modes in a region, is proportional to the size of the region. Specifically, at large $j_{\max}$, $\mathrm{tr} P_A^{j_{\max}} \propto j_{\max}^2 |A|$ as is seen numerically in Fig. 5.9 and understood analytically in Appendix 5.11.

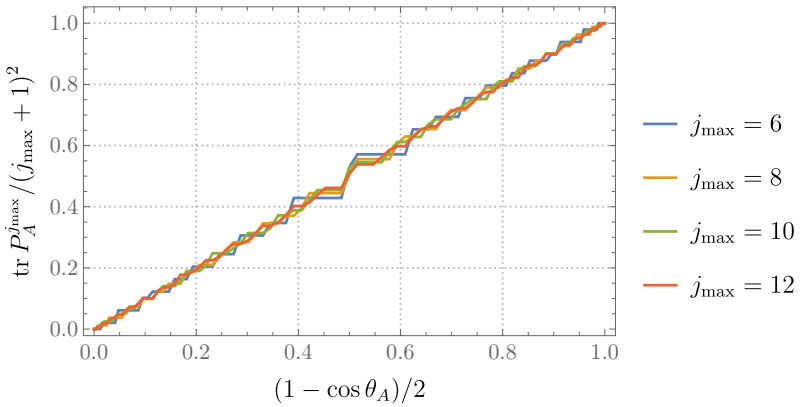


Figure 5.9: Trace of the projector versus fractional area of the region (a spherical cap with polar angle θ_A), with different angular momentum cutoffs $j_{\max}$. A linear proportionality is observed at large $j_{\max}$. The discreteness in the plot arises because the finite $j_{\max}$ space of functions cannot resolve all angles.

2. The second Rényi entropy defined by the projector follows a boundary law. At large $j_{\max}$, with the mass fixed to $\mu = 1$, the entropy $S_2 \approx 0.03\, j_{\max} |\partial A|$ as is seen numerically in Fig. 5.10 and understood analytically in Appendix 5.11.

This boundary entanglement law in Fig. 5.10 is of course precisely the expected entanglement in the ground state of a local quantum field [33, 235]. As the cutoff $j_{\max}$ is removed, the entanglement grows unboundedly.

The partition we have just defined can now be adapted to the fluctuations about the large ν fuzzy sphere state in the matrix quantum mechanics model. We do this in the following subsection. Intuitively, we would like to replace the $j(j+1)+\mu^2$ spectrum of the free field in the wavefunction (5.37) with the matrix mechanics modes (5.29). Recall that the matrix modes are cut off at angular momentum $j_{\max} = N - 1$.

5.5.2 Fuzzy sphere in the mini-BMN model

Now we address two additional subtleties that arise when adapting the free field ideas above to the mini-BMN fuzzy sphere. Firstly, the mini-BMN theory is an SU(N) gauge theory. It is known that

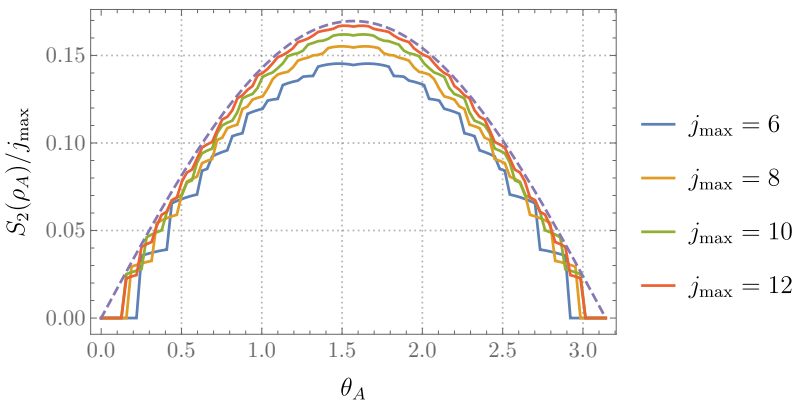


Figure 5.10: The second Rényi entropy for a complex scalar free field (with mass $\mu = 1$) versus the polar angle θ_A of a spherical cap. The entropy with different cutoffs $j_{\max}$ is shown. At large $j_{\max}$ the curve approaches the boundary law $0.03 \times 2\pi \sin\theta_A$, shown as a dashed line. Discreteness in the plot is again due to the finite $j_{\max}$ space of functions.

entanglement in gauge theories may depend upon the choice of gauge-invariant algebras associated to spatial regions [46]. Different prescriptions correspond to different boundary or gauge conditions [150]. However for a fuzzy geometry, the boundaries of regions and gauge edge modes are not sharply defined. To introduce the fewest additional degrees of freedom, we choose to factorize the physical Hilbert space, instead of an extended one [68, 69], to evaluate entanglement in the mini-BMN model. This is similar to the 'balanced center' procedure in [46], where edge modes are absent.[4]

Secondly, the emergent fields include fluctuations of the geometry itself. The factorization that we have discussed in the previous subsection is tailored to a region on the sphere, and does not need to approximate a spatial region in other geometries. The partition is even less meaningful in non-geometric regions of the Hilbert space. The variational wavefunction we have constructed can be used to compute entanglement for any given factorization of the Hilbert space, but it is unclear that preferred factorizations exist away from geometric limits. In this work we will focus on the entanglement in the $\nu \to \infty$ limit where the fields are infinitesimal, and hence do not backreact on the spherical geometry. In this limit the factorization is precisely — up to issues of gauge invariance — that of the free-field case discussed in the previous subsection.

The matrices corresponding to the infinitesimal fields on the fuzzy sphere are, cf. (5.28),

$$A^i = X^i - \nu J^i, \tag{5.39}$$

which should be thought of as living in the tangent space at $X^i = \nu J^i$. At large ν the wavefunction is

[4]It should, nonetheless, be possible to identify meaningful SU(N) 'edge modes' that would reproduce the edge mode contribution of the emergent Maxwell field. This is an especially interesting question in the light of the fact that the microscopic SU(N) gauge symmetry also acts as an area-preserving diffeomorphism on the emergent fields in (5.33). This is left for future work.

strongly supported on the classical configuration and hence in this limit the infinitesimal description is accurate. Gauge transformations then act as

$$A^i \to A^i + i\epsilon[Y, \nu J^i] + \dots, \tag{5.40}$$

where ϵ is infinitesimal and Y is an arbitrary Hermitian matrix. The $\epsilon[Y, A^i]$ term is omitted in (5.40) as it is of higher order. Gauge invariance of the state is manifested as

$$\psi(\nu J^i + A^i) = \psi(\nu J^i + A^i + i\epsilon[Y, \nu J^i]). \tag{5.41}$$

Physical states are wavefunctions on gauge orbits $[A^i]$, the set of infinitesimal matrices differing from A^i by a gauge transformation (5.40). Similarly to the discussion of free fields above, a partition of the space of gauge orbits is specified by a projector P. We will now explain how this projector is constructed. Given a projector P' acting on infinitesimal matrices A^i, a projector acting on gauge orbits can be defined as

$$P([A^i]) = [P'(A^i)]. \tag{5.42}$$

However, for P to be well-defined, P' must preserve gauge directions:

$$P'(A^i + i\epsilon[Y, \nu J^i]) = P'(A^i) + i\epsilon[Y', \nu J^i], \tag{5.43}$$

for any A^i, Y and some Y' dependent on Y. Let V be the subspace of gauge directions:

$$V = \{i[Y, J^i] : Y \text{ is Hermitian}\}, \tag{5.44}$$

then (5.43) is equivalent to the requirement that $P'(V) \subset V$. The strategy for finding the projector P is to solve for the projector P' that minimizes $\|P' - \chi_A\|$ subject to the constraint that (5.43) is satisfied. Then P is defined via P' as in (5.42).

The problem of minimizing $\|P' - \chi_A\|$ for orthogonal projectors P' such that $P'(V) \subset V$ is exactly solvable as follows. The condition that $P'(V) \subset V$ is equivalent to imposing that $P' = P_V \oplus P_{V_\perp}$, where P_V is some projector in the subspace V and $P_{V_\perp}$ in its orthogonal complement $V_\perp$. And $\|P' - \chi_A\|$ is minimized if and only if $\|P_V - \chi_A|_V\|$ and $\|P_{V_\perp} - \chi_A|_{V_\perp}\|$ are both minimized. Via the correspondence between matrix spherical harmonics $\hat{Y}_{jm}$ and spherical harmonic functions $Y_{jm}(\theta, \phi)$ in Appendix 5.9, both of these minimizations become the same problem as in the free field case, with a detailed solution in Appendix 5.11.

The second Rényi entropy, in terms of gauge orbits, is evaluated similarly to (5.38):

$$S_2(\rho_A) = -\ln \int d[A]d[A']\,\Delta([A])\Delta([A']) \\ \times \psi_{\rm inv}([A])\psi^*_{\rm inv}(P[A'] + (I-P)[A])\psi_{\rm inv}([A'])\psi^*_{\rm inv}(P[A] + (I-P)[A']), \tag{5.45}$$

where Δ are measure factors for gauge orbits and $\psi_{\rm inv}([A]) = \psi(\nu J + A)$. Recall that ψ is gauge invariant according to (5.41). The formula (5.45) as displayed does not involve any gauge choice. However, there are some gauges where evaluating (5.45) is particularly convenient. The gauge we choose for this purpose, which is different from that in section 5.2.2, is that $A \in V_\perp$, i.e., the fields are perpendicular to gauge directions. In this gauge measure factors are trivial and the projector is simply $P_{V_\perp}$ that minimizes $\|P_{V_\perp} - \chi_A|_{V_\perp}\|$:

$$S_2(\rho_A) = -\ln \int_{V_\perp} dA dA' \\ \times \psi_\perp(A)\psi^*_\perp(P_{V_\perp}A' + (I-P_{V_\perp})A)\psi_\perp(A')\psi^*_\perp(P_{V_\perp}A + (I-P_{V_\perp})A'), \tag{5.46}$$

where $\psi_\perp(A)$ is defined as $\psi(\nu J + A)$ for $A \in V_\perp$.[5]

The bosonic fuzzy sphere wavefunction can be written in the $\nu \to \infty$ limit as follows. As in (5.28), the perturbations can be decomposed as $A^i = \sum_a \delta x_a \sum_{jm} y^i_{jma} \hat{Y}_{jm}$, where the y^i_{jma} diagonalize the potential energy at quadratic order in A so that $V = \frac{\nu^2}{2}\sum_a \omega_a^2 (\delta x_a)^2 + \cdots$ (see Appendix 5.9). The wavefunction is then, analogously to (5.37),

$$\psi_\perp(A) \propto e^{-\frac{|\nu|}{2}\sum_a |\omega_a| (\delta x_a)^2}. \tag{5.47}$$

The frequencies are given by (5.29), excluding the pure gauge zero modes. Using this wavefunction, the Rényi entropy (5.46) can be computed exactly and is shown as a solid line in Fig. 5.11. As $N \to \infty$ these curves approach a boundary law

$$S_2(\rho_A) \approx 0.03\, N\, |\partial A|\ . \tag{5.48}$$

Here $|\partial A| = 2\pi \sin\theta_A$ is again the circumference of the spherical cap A (in units where the sphere has radius one, consistent with the field theoretic description in (5.31)). The result (5.48) is the same as that of the toy model in Fig. 5.10, with $j_{\rm max}$ now set by the microscopic matrix dynamics to be $N-1$.[6] This regulated boundary-law entanglement underpins the emergent locality on the fuzzy

[5] We can find a gauge transformation $U \in {\rm SU}(N)$ mapping any matrices X^i into this perpendicular gauge as follows. We are looking for $\widetilde{X}^i = UX^iU^{-1}$, such that $\widetilde{X}^i - \nu J^i \in V_\perp$. This means that $\sum_i {\rm tr}\left([Y,J^i]^\dagger(\widetilde{X}^i - \nu J^i)\right) = 0$ for any Hermitian matrix Y. Equivalently, $\sum_i {\rm tr}\left(J^i[Y,\widetilde{X}^i]\right) = 0$ for any Y. This is achieved by numerically finding the U that maximizes the overlap $\sum_i {\rm tr}\left(J^iUX^iU^{-1}\right)$.

[6] A (simpler) instance of entanglement revealing the inherent graininess of a spacetime built from matrices is two

sphere at large N and ν. Recall from the discussion around (5.31) that there are only two emergent fields on the sphere: a Maxwell field and a scalar field. The perpendicular gauge choice we have made translates into the Coulomb gauge for the emergent Maxwell field, cf. the discussion around (5.33) above. The factor of N in (5.48) is due to the microscopic cutoff at a scale $L_{\rm fuzz} \sim L_{\rm sph}/N$.

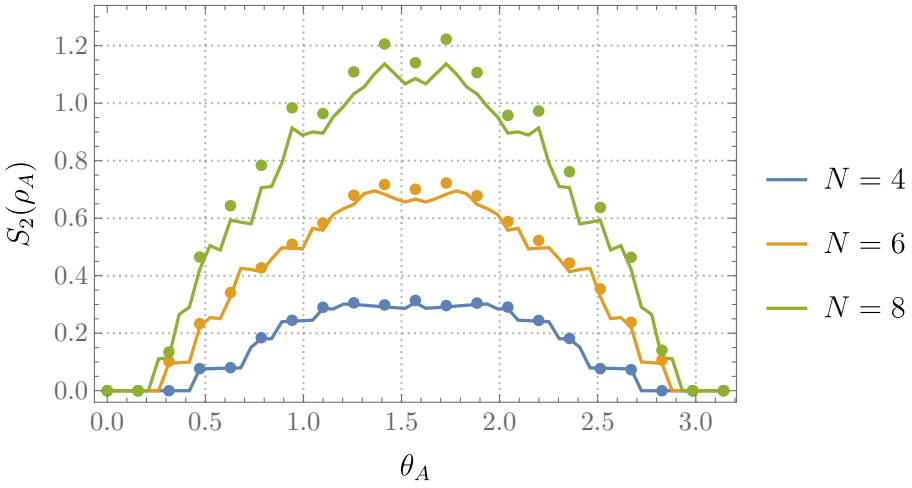


Figure 5.11: The second Rényi entropy for a spherical cap on the matrix theory fuzzy sphere versus the polar angle θ_A of the cap. Solid curves are exact values at $\nu = \infty$ and dots are numerical values from variational wavefunctions at $\nu = 10$ for different N. The wavefunctions are NF(1, 1) in the zero fermion sector as shown in Figs. 5.2 and 5.3.

Previous works on the entanglement of a free field on a fuzzy sphere involved similar wavefunctions but a different factorization of the Hilbert space, which was inspired instead by coherent states [71, 126, 187, 51]. Those results did not always produce boundary-law entanglement. Here we see that the UV/IR mixing in noncommutative field theories does not preclude a partition of the large N and large ν Hilbert space with a boundary-law entanglement.

We can also evaluate the entropy (5.46) using the large ν variational wavefunctions, without assuming the asymptotic form (5.47). The results are shown as dots in Fig. 5.11. However, we stress that only the $\nu \to \infty$ limit has a clear physical meaning, where fluctuations are infinitesimal. The variational results are close to the exact values in Fig. 5.11, showing that the neural network ansatz captures the entanglement structure of these matrix wavefunctions.

The results in this section are for the bosonic fuzzy sphere. The projection we have introduced in order to partition the space of matrices can be extended in a similar, but more involved, way to factorize the fermionic Hilbert space.

dimensional string theory [57, 99].

5.6 Discussion

We have seen that neural network variational wavefunctions capture in detail the physics of a semiclassical spherical geometry that emerges in the mini-BMN model (5.3) at large ν. Away from the semiclassical limit, the spherical geometry either abruptly or gradually collapses towards a new state. In Fig. 5.8 we saw that in the 'supersymmetric' sector this new state was characterized by an increase in both the expectation value and quantum mechanical variance of the radius as $\nu \to 0$. To understand the physics of this process, and to start thinking about the nature of the collapsed state as $\nu \to 0$, it is helpful to consider the string theoretic embedding of the model.

The mini-BMN model can be realized in string theory as the description of N D-particles in an AdS_4 spacetime. Let us review some aspects of this realization [12]. The parameter

$$\frac{1}{\nu^3} \sim g_s \left(\frac{L_{AdS}}{L_s}\right)^3 . \tag{5.49}$$

Here L_{AdS} is the AdS radius, L_s is the string length and g_s is the string coupling. The proportionality in (5.49) depends on the volume, in units of the string length, of internal cycles wrapped by the branes in the compactification down to AdS_4. In particular, the mass of a single D-particle goes like $1/g_s$ times the wrapped internal volume. The strength of the gravitational backreaction of N coincident D-particles is then controlled by $G_N \cdot N/g_s$. Here $G_N \sim g_s^2$ is the four dimensional Newton constant, where we have suppressed a factor of the volume of the compactification manifold. Therefore, if we keep the AdS radius fixed in string units, gravitational backreation becomes important when $g_s N \sim N/\nu^3 \gtrsim 1$. Up to factors of the volume of compactification cycles, this is equivalent to the statement that the dimensionless 't Hooft coupling $\lambda = N/\nu^3$, introduced below (5.1), becomes large.

For $N/\nu^3 \lesssim 1$, then, the D-particles can be treated as light probes on the background AdS spacetime. The fuzzy sphere configuration describes a polarization of the D-particles into spherical 'dual giant gravitons'. From the string theory perspective, this polarization is driven by the 4-form flux $\Omega \sim 1/L_{AdS}$ supporting the background AdS_4 spacetime. Together with the discussion in the previous paragraph on the strength of the gravitational interaction, we can write the heuristic relation $N/\nu^3 \sim$ gravity/flux. At large ν the flux wins out and semiclassical fuzzy spheres can exist, but at small ν gravitational forces cause the spheres to collapse. The entanglement and emergent locality that we have described in this paper is that of the polarized spheres, whose excitations are described by the usual gauge fields and transverse scalar fields of string theoretic D-branes.

For $N/\nu^3 \gg 1$ it is possible that the strongly interacting, collapsed D-particles will develop a geometric 'throat', in the spirit of the canonical holographic correspondence [164]. It is not well-understood when such a throat would be captured by the mini-BMN matrix quantum mechanics. The variational wavefunctions that we have developed here provide a new window into this problem. In particular, we hope to investigate the small ν collapsed state in more detail in the future,

with the objective of revealing any entanglement associated to emergent local dynamics in the throat spacetime. If the emergent dynamics includes gravity, there are two potentially interesting complications. Firstly, the entanglement of bulk fields may be entwined with entanglement due to the 'stringy' degrees of freedom that seem to be manifested in the Bekenstein-Hawking entropy of black holes as well as in the Ryu-Takayanagi formula [237, 75, 26, 73]. Secondly, and perhaps relatedly, it may become crucial to understand the 'edge mode' contribution to the entanglement, that we have avoided in our discussion here [70, 94].

More generally, the methods we have developed will be applicable to a wide range of quantum problems of interest in the holographic correspondence. The benefit of the variational neural network approach is direct access to properties of the zero temperature quantum mechanical state. Optimizing the numerical methods and variational ansatz further, and with more computational power, it should not be difficult to work with larger values of N. In addition to understanding the emergence of spacetime from first principles, it should also be possible to study, for example, the microstates and dynamics of quantum black holes.

Acknowledgements

It is a pleasure to thank Frederik Denef and Xiaoliang Qi for helpful discussions, Aitor Lewkowycz, Raghu Mahajan and Edward Mazenc for comments on the draft, and Tarek Anous for sharing his code with us. We also thank Zhaoheng Guo and Yang Song for collaboration on a related project. SAH is partially funded by DOE award de-sc0018134. XH is supported by a Stanford Graduate Fellowship. Computational work was performed on the Sherlock cluster at Stanford University, with the TensorFlow code for the project available online.

5.7 Appendix A: Geometry of the gauge

Gauge invariant sampling

In the procedure of sampling bosonic matrices X according to the wavefunction probability distribution $|\psi(X)|^2 = |f(X)|^2$, it is asserted in the main text that $X \sim |f(X)|^2$ if we let $X = U\widetilde{X}U^{-1}$ where U is a Haar random element in SU(N) and the representative of the gauge orbit $\widetilde{X} \sim \Delta(\widetilde{X})|\widetilde{f}(\widetilde{X})|^2$. A proof of this assertion, along with a more precise definition of the gauge orbit measure Δ, is presented here.

To simplify notation, denote $\widetilde{X} \sim \widetilde{p}(\widetilde{X})$. If the random variable $X = U\widetilde{X}U^{-1}$, it follows the probability distribution

$$p(X = X_0) = \int dU d\widetilde{X}\, \widetilde{p}(\widetilde{X})\delta(U\widetilde{X}U^{-1} = X_0), \tag{5.50}$$

where the integral over SU(N) is with respect to the normalized Haar measure, and δ is the Dirac delta distribution. For almost any X_0, there is a unique gauge representative $\widetilde{X}_0$, with a discrete set of $U_i \in \mathrm{SU}(N)$ ($i = 1, 2, \ldots, N$), such that $U_i\widetilde{X}_0U_i^{-1} = X_0$. These unitaries differ by an overall phase (powers of $\exp(i2\pi/N)$). Hence

$$p(X = X_0) = \widetilde{p}(\widetilde{X}_0)\sum_{i=1}^{N}|J^{-1}(\widetilde{X}_0, U_i)|, \tag{5.51}$$

where J is the Jacobian determinant of the map $(\widetilde{X}, U) \mapsto U\widetilde{X}U^{-1}$. As will be seen in the next subsection, $J(\widetilde{X}, U) = J(\widetilde{X})$ does not depend on the unitary U. So if we assign

$$\Delta(\widetilde{X}) = N^{-1}|J(\widetilde{X})|, \tag{5.52}$$

and note $\widetilde{p}(\widetilde{X}) = \Delta(\widetilde{X})|\widetilde{f}(\widetilde{X})|^2$,

$$p(X = X_0) = N^{-1}|J(\widetilde{X}_0)||\widetilde{f}(\widetilde{X}_0)|^2\sum_{i=1}^{N}|J^{-1}(\widetilde{X}_0)| = |\widetilde{f}(\widetilde{X}_0)|^2 = |f(X_0)|^2, \tag{5.53}$$

for a gauge invariant wavefunction (5.10). This is the desired result.

Derivation of the gauge orbit measure

From (5.52), the gauge orbit measure Δ is given by the Jacobian determinant J of the map $X : (\widetilde{X}, U) \mapsto U\widetilde{X}U^{-1}$. Recall that for a general mapping F between smooth manifolds of equal dimension $S \to T$, the Jacobian determinant can be written in terms of the pullback of the volume

form

$$F^*(\omega_T) = J\omega_S, \tag{5.54}$$

where ω_S and ω_T are volume forms on S and T. That is, J is the ratio of the volume element after and before the mapping. If x_i and y_i are two orthonormal coordinate systems at $x \in S$ and $y = F(x) \in T$, in terms of the wedge product,

$$\omega_S = \bigwedge_i dx_i, \quad \omega_T = \bigwedge_i dy_i, \quad F^*(dy_i) = \sum_j \frac{\partial y_i}{\partial x_j} dx_j \,. \tag{5.55}$$

Therefore equation (5.54) can be expressed more explicitly as

$$\bigwedge_i \sum_j \frac{\partial y_i}{\partial x_j} dx_j = J \bigwedge_i dx_i \quad \Leftrightarrow \quad J = \det \frac{\partial y_i}{\partial x_j}. \tag{5.56}$$

We would like to show firstly that $J(\widetilde{X}, U)$ does not depend on U. Note that the map $X : (\widetilde{X}, U) \mapsto U\widetilde{X}U^{-1}$ is equivariant with respect to the following actions of $G = \mathrm{SU}(N)$: for any $U' \in G$, in the base space $U' \cdot (\widetilde{X}, U) = (\widetilde{X}, U'U)$, and in the target space $U' \cdot X = U'XU'^{-1}$. And the two actions preserve the volume forms, because the Haar measure is left invariant and the metric $\mathrm{tr} dX^\dagger dX$ is invariant under matrix conjugation. Hence the Jacobian $J(\widetilde{X}, U) = J(\widetilde{X})$ is independent of U.

We will obtain the Jacobian by explicitly computing the pullback of the volume form at X. As the Jacobian does not depend on U, it is convenient to evaluate it at $U = I$. To further simplify the computation, we shall complexify the cotangent spaces, which does not change the Jacobian determinant. The $\mathfrak{su}(N)$ real Lie algebra is complexified to $\mathfrak{sl}(N)$, and the following basis $\{D_i, E_{ij}\}$ of $\mathfrak{sl}(N)$ is employed. The basis is orthonormal with respect to the matrix inner product $\mathrm{tr} X^\dagger Y$:

1. For $1 \le i \le N-1$, D_i is a diagonal matrix with $(D_i)_{jj} = 1/\sqrt{i(i+1)}$ for $1 \le j \le i$, $(D_i)_{jj} = -(j-1)/\sqrt{i(i+1)}$ for $j = i+1$ and $(D_i)_{jj} = 0$ for $j > i+1$.

2. For $1 \le i, j \le N$ and $i \ne j$, E_{ij} is the matrix that has only one nonzero entry $(E_{ij})_{ij} = 1$.

A general element in the complexified cotangent space of $\widetilde{X}$ is (with the gauge choice defined in the main text)

$$\begin{aligned} d\widetilde{X}^1 &= \sum_{i=1}^{N-1} D_i d\widetilde{c}_i^1, \quad d\widetilde{X}^3 = \sum_{i=1}^{N-1} D_i d\widetilde{c}_i^3 + \sum_{1 \le i \ne j \le N} E_{ij} d\widetilde{e}_{ij}^3, \\ d\widetilde{X}^2 &= \sum_{i=1}^{N-1} D_i d\widetilde{c}_i^2 + \sum_{i=1}^{N-1} \frac{1}{\sqrt{2}} \left(E_{i(i+1)} - E_{(i+1)i}\right) d\widetilde{e}_{i(i+1)}^2 + \sum_{1 \le i \ne j \le N}^{|i-j| \ne 1} E_{ij} d\widetilde{e}_{ij}^2, \end{aligned} \tag{5.57}$$

where the superscript $i = 1, 2, 3$ denotes three bosonic matrices. The equations (5.57) thus define a

basis $\{d\widetilde{c}^1_i, d\widetilde{c}^2_i, d\widetilde{e}^2_{i(i+1)}, d\widetilde{e}^2_{ij}, d\widetilde{c}^3_i, d\widetilde{e}^3_{ij}\}$ of the complexified cotangent space of $\widetilde{X}$.

The complexified cotangent space of SU(N) at $U = I$ is isomorphic to the Lie algebra $\mathfrak{sl}(N)$, so that (introducing basis forms dc_i, de_{ij}):

$$-idU = \sum_{i=1}^{N-1} D_i dc_i + \sum_{1\le i\neq j\le N} E_{ij} de_{ij}. \tag{5.58}$$

The differential of the map $X : (\widetilde{X}, U) \mapsto U\widetilde{X}U^{-1}$ at $U = I$ is

$$dX = [dU, \widetilde{X}] + d\widetilde{X}, \tag{5.59}$$

and the cotangent space of X is complexified to three copies of $\mathfrak{sl}(N)$, so that (introducing basis forms dc^k_i, de^k_{ij}):

$$dX^k = \sum_{i=1}^{N-1} D_i dc^k_i + \sum_{1\le i\neq j\le N} E_{ij} de^k_{ij}. \tag{5.60}$$

Substituting (5.57) and (5.58) into (5.59), recalling that $\widetilde{X}^1$ is diagonal, and equating the expressions for dX^1 we have

$$dc^1_i = d\widetilde{c}^1_i, \quad de^1_{ij} = i\left(\widetilde{X}^1_{jj} - \widetilde{X}^1_{ii}\right) de_{ij}. \tag{5.61}$$

Equating the expressions for dX^2 gives, with terms that drop out of the final result omitted:

$$\begin{aligned}
dc^2_i &= d\widetilde{c}^2_i + (\text{terms with } de), \quad de^2_{ij} = d\widetilde{e}^2_{ij} + (\text{terms with } dc, de),\\
de^2_{i(i+1)} &= +i\widetilde{X}^2_{i(i+1)}\sqrt{\frac{i+1}{i}}dc_i + \frac{1}{\sqrt{2}}d\widetilde{e}^2_{i(i+1)} + (\text{terms with } dc_{i-1}, de),\\
de^2_{(i+1)i} &= -i\widetilde{X}^2_{(i+1)i}\sqrt{\frac{i+1}{i}}dc_i - \frac{1}{\sqrt{2}}d\widetilde{e}^2_{i(i+1)} + (\text{terms with } dc_{i-1}, de),
\end{aligned} \tag{5.62}$$

where the expression for de^2_{ij} holds for $|i-j| \neq 1$ and the prefactor i in the expressions for $de^2_{i(i+1)}$ and $de^2_{(i+1)i}$ is the imaginary unit. Subscripts are omitted if that term with any subscript is unimportant, e.g., de means linear combinations of de_{ij} for $1 \le i \neq j \le N$. Similarly

$$dc^3_i = d\widetilde{c}^3_i + (\text{terms with } de), \quad de^3_{ij} = d\widetilde{e}^3_{ij} + (\text{terms with } dc, de). \tag{5.63}$$

The Jacobian determinant J is evaluated as, schematically,

$$dc^1_i \wedge de^1_{ij} \wedge dc^2_i \wedge de^2_{ij} \wedge dc^3_i \wedge de^3_{ij} = J\, d\widetilde{c}^1_i \wedge d\widetilde{c}^2_i \wedge d\widetilde{e}^2_{i(i+1)} \wedge d\widetilde{e}^2_{ij} \wedge d\widetilde{c}^3_i \wedge d\widetilde{e}^3_{ij} \wedge dc_i \wedge de_{ij}, \tag{5.64}$$

where de^1_{ij} denotes $\bigwedge_{ij} de^1_{ij}$ for $1 \le i \neq j \le N$, for example. Substitution of (5.61), (5.62) and (5.63) into the left-hand side of (5.64) yields a sum of wedge products of differentials. The wedge product

is nonzero only if each factor on the right-hand side of (5.64) appears exactly once. Now observe that de_{ij} already appears in de^1_{ij} in (5.61), hence all de_{ij} terms in other factors can be safely ignored.

With the de_{ij} ignored, $de^2_{i(i+1)} \wedge de^2_{(i+1)i}$ is proportional to $dc_i \wedge d\widetilde{e}^2_{i(i+1)}$ for $i = 1$, because for any differential da, $da \wedge da = 0$. Then remaining factors of dc_1 and $d\widetilde{e}^2_{12}$ can be ignored. Next, for $i = 2$, $de^2_{i(i+1)} \wedge de^2_{(i+1)i}$ must be proportional to $dc_i \wedge d\widetilde{e}^2_{i(i+1)}$ as well, up to terms that can be ignored. In the end we have (note that $\widetilde{X}^2_{i(i+1)} = -\widetilde{X}^2_{(i+1)i}$ is purely imaginary)

$$\bigwedge_{i=1}^{N-1} de^2_{i(i+1)} \wedge de^2_{(i+1)i} = \sqrt{2^{N-1}N} \bigwedge_{i=1}^{N-1} \mathrm{Im}\, \widetilde{X}^2_{i(i+1)} dc_i \wedge d\widetilde{e}^2_{i(i+1)} + (\text{terms with } de). \tag{5.65}$$

Now terms with dc_i can be ignored as well as they appear in (5.65). With the dc_i and de_{ij} ignored, dc^1_i, dc^2_i, de^2_{ij} for $|i-j| \neq 1$, dc^3_i and de^3_{ij} on the left-hand side of (5.64) can be replaced by $d\widetilde{c}^1_i$, $d\widetilde{c}^2_i$, $d\widetilde{e}^2_{ij}$, $d\widetilde{c}^3_i$ and $d\widetilde{e}^3_{ij}$, respectively, in the light of (5.61), (5.62) and (5.63). The Jacobian is then a product of the factors in (5.61) and (5.65). Thus overall the gauge orbit measure is

$$\Delta \propto |J| \propto \prod_{i\neq j=1}^{N} \left|\widetilde{X}^1_{ii} - \widetilde{X}^1_{jj}\right| \prod_{i=1}^{N-1} \left|\widetilde{X}^2_{i(i+1)}\right|. \tag{5.66}$$

5.8 Appendix B: Evaluation of observables

The physical observables that we are interested in fall into roughly three categories: *(i)* bosonic potentials; *(ii)* fermionic bilinears; *(iii)* casimirs of Lie group actions. Efficient numerical recipes for evaluating these observables via Monte Carlo simulation are discussed in this Appendix. Monte Carlo requires that the integrals are written as the average over samples $\mathbb{E}_{X\sim|f|^2}[\cdot]$.

Bosonic potentials are real functions of bosonic matrix coordinates $V(X)$, and they are straightforward to evaluate:

$$\langle\psi|\hat{V}_1|\psi\rangle \equiv \int dX\, |f(X)|^2 V(X) = \mathbb{E}_{X\sim|f|^2}[V(X)]. \tag{5.67}$$

Fermionic bilinears and casimirs are more elaborate to compute. The final results are (5.80) and (5.85) with detailed derivations presented below.

Fermionic bilinears

Expectation values of fermionic bilinears $B(\lambda^\dagger, \lambda, X)$ are

$$\langle\psi|\hat{V}_2|\psi\rangle \equiv \int dX\, |f(X)|^2 \langle M(X)|B(\lambda^\dagger, \lambda, X)|M(X)\rangle. \tag{5.68}$$

The problem is thus essentially to evaluate fermionic bilinears in the fermionic state $|M(X)\rangle$, which can furthermore be reduced to calcuating

$$\langle M^r|B(\lambda^\dagger,\lambda,X)|M^s\rangle, \tag{5.69}$$

where $|M^r\rangle$ is the free fermion state

$$|M^r\rangle \equiv \prod_{a=1}^{R}\Big(\sum_{\alpha=1}^{2}\sum_{A=1}^{N^2-1} M^{ra}_{A\alpha}\lambda_A^{\alpha\dagger}\Big)|0\rangle. \tag{5.70}$$

The question is more generally formulated as follows: let M be a complex matrix of size $R\times P$ and denote its corresponding free fermion state as

$$|M\rangle = \prod_{a=1}^{R}\Big(\sum_{p=1}^{P} M_{ap}\lambda_p^\dagger\Big)|0\rangle, \tag{5.71}$$

then what are the matrix elements $\langle M'|B(\lambda^\dagger,\lambda,X)|M\rangle$? The starting point is the Slater determinant:

$$\langle M'|M\rangle = \det(MM'^\dagger), \tag{5.72}$$

and note that

$$\langle M'|\lambda_p^\dagger\lambda_q|M\rangle = \delta_{pq}\langle M'|M\rangle - \langle M'|\lambda_q\lambda_p^\dagger|M\rangle, \tag{5.73}$$

where the first term on the right-hand side can be evaluated from (5.72). The second term in (5.73) can be read as the overlap between free fermion states $\lambda_q^\dagger|M'\rangle$ and $\lambda_p^\dagger|M\rangle$ and thus (5.72) is again applicable:

$$\begin{aligned} s^2\langle M'|\lambda_q\lambda_p^\dagger|M\rangle &= \det\begin{pmatrix} s^2\delta_{pq} & sM'^\dagger_{p:} \\ sM_{:q} & MM'^\dagger \end{pmatrix} \\ &= \det\begin{pmatrix} 1 & sM'^\dagger_{p:} \\ sM_{:q} & MM'^\dagger \end{pmatrix} + (s^2\delta_{pq}-1)\det(MM'^\dagger) \\ &= \det(MM'^\dagger - s^2M_{:q}M'^\dagger_{p:}) + (s^2\delta_{pq}-1)\det(MM'^\dagger). \end{aligned} \tag{5.74}$$

A dummy variable s is introduced for later convenience. Using (5.74) in (5.73)

$$s^2\langle M'|\lambda_p^\dagger\lambda_q|M\rangle = \det(MM'^\dagger) - \det(MM'^\dagger - s^2M_{:q}M'^\dagger_{p:}). \tag{5.75}$$

Differentiate both sides with respect to s^2 to obtain a more compact expression:

$$\langle M'|\lambda_p^\dagger\lambda_q|M\rangle = \mathrm{tr}\left[\mathrm{adj}(MM'^\dagger)M_{:q}M'^\dagger_{p:}\right], \tag{5.76}$$

where $\mathrm{adj}\, A = (\det A)A^{-1}$ is the adjucate of A. For an arbitrary bilinear W,

$$\sum_{pq}\langle M'|\lambda_p^\dagger W_{qp}\lambda_q|M\rangle = \det(MM'^\dagger)\mathrm{tr}\left[(MM'^\dagger)^{-1}MWM'^\dagger\right]. \tag{5.77}$$

Back to the original problem of calculating (5.69). Equation (5.77) is applicable if we regard the index p in (5.71) as running over both the indices α and A in (5.70). Define the overlap matrix

$$(O^{rs})^{ab} \equiv \sum_{\alpha=1}^{2}\sum_{A=1}^{N^2-1}(M_{A\alpha}^{ra})^* M_{A\alpha}^{sb}, \tag{5.78}$$

then

$$\langle M^r|B(\lambda^\dagger,\lambda,X)|M^s\rangle = \sum_{ab=1}^{R}(\mathrm{adj}\, O^{rs})_{ba}B(M^{ra\dagger},M^{sb},X), \tag{5.79}$$

where the fermionic operators in the bilinear are replaced by complex matrices so that the expression is a complex number. Finally summing over r and s,

$$\langle\psi|\hat{V}_2|\psi\rangle = \mathbb{E}_{X\sim|f|^2}\left[\sum_{rs=1}^{D}\sum_{ab=1}^{R}(\mathrm{adj}\, O^{rs}(X))_{ba}B(M^{ra\dagger}(X),M^{sb}(X),X)\right]. \tag{5.80}$$

Casimirs

The observables discussed above do not involve derivatives. Derivatives show up in kinetic terms, for example, and can be understood in a geometric way. For an action of a Lie group G on the wavefunction ψ, a casimir term can be defined as

$$\langle\psi|\hat{V}_3|\psi\rangle \equiv \sum_A\int dX\,\langle d_A\psi(X)|d_A\psi(X)\rangle, \tag{5.81}$$

where the summation is over an orthonormal basis of the Lie algebra and

$$|d_A\psi(X)\rangle \equiv \frac{d}{ds}(e^{isT_A}\psi)(X)\Big|_{s=0}. \tag{5.82}$$

As an example, consider the group of translations of bosonic coordinates $X \to X+\delta X$ that acts on the wavefunction as

$$(e^{isT_A}\psi)(X) = \psi(X-sT_A), \quad \frac{d}{ds}(e^{isT_A}\psi)(X)\Big|_{s=0} = -\sum_{ij}T_{Aij}\frac{\partial\psi}{\partial X_{ij}}, \tag{5.83}$$

and thus in this case

$$\langle\psi|\hat{V}_3|\psi\rangle = \sum_{Aiji'j'} \int dX\, T^*_{Ai'j'} T_{Aij} \Big\langle \frac{\partial\psi}{\partial X_{i'j'}} \Big| \frac{\partial\psi}{\partial X_{ij}} \Big\rangle = \sum_{ij} \int dX \Big\langle \frac{\partial\psi}{\partial X_{ij}} \Big| \frac{\partial\psi}{\partial X_{ij}} \Big\rangle, \tag{5.84}$$

which is the usual kinetic term. If $G = \mathrm{SU}(N)$ with the adjoint action on matrices, the observable (5.81) is the casimir of the gauge group, and if $G = \mathrm{SO}(3)$ in the mini-BMN model, the observable measures the angular momentum quantum number of the state.

The summation and the integral in (5.81) are estimated from Monte Carlo samples as:

$$\langle\psi|\hat{V}_3|\psi\rangle = \mathbb{E}_{|T_A|^2=\dim G, X\sim|f|^2}\Big[|f(X)|^{-2}\langle d_A\psi(X)|d_A\psi(X)\rangle\Big], \tag{5.85}$$

where $f = |\psi|$, $|T_A|^2 = \dim G$ means that the expectation value averages over all Lie algebra elements T_A with norm $\sqrt{\dim G}$.

5.9 Appendix C: Semiclassical analysis of the fuzzy sphere

Correspondence between matrices and fields on the emergent sphere

A mapping from any N-by-N complex matrix A to a function $f_A(\theta,\phi)$ is constructed as follows. The construction is motivated by the following principles: *(i)* the map $A \mapsto f_A(\theta,\phi)$ should be linear; *(ii)* the map should preserve the inner products:

$$\frac{1}{N}\mathrm{tr}(A^\dagger A') = \frac{1}{4\pi}\int d\Omega\, f^*_A(\theta,\phi) f_{A'}(\theta,\phi). \tag{5.86}$$

Here $\int d\Omega$ is the integral over a 4π solid angle; *(iii)* the map should preserve the $\mathfrak{su}(2)$ action:

$$f_{[J^i,A]}(\theta,\phi) = (L^i f_A)(\theta,\phi). \tag{5.87}$$

As in the main text, the J^i are generators of the N dimensional irreducible representation of $\mathfrak{su}(2)$ and the L^i are generators for rotations of functions on a sphere:

$$L^i = -i\epsilon_{ijk}x^j\frac{\partial}{\partial x^k}, \tag{5.88}$$

and $(x^1,x^2,x^3) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$.

Requirements *(i)* and *(ii)* can be accomplished by mapping an orthonormal basis of matrices to an orthonormal basis of functions on the sphere. In the light of *(iii)*, we choose spherical harmonics

$Y_{jm}(\theta,\phi)$ ($j \geq 0$, $|m| \leq j$) as the basis of functions:

$$\sum_{i=1}^{3} L^i L^i Y_{jm} = j(j+1)Y_{jm}, \quad L^3 Y_{jm} = mY_{jm}, \tag{5.89}$$

and they are orthonormal with respect to the inner product in (5.86):

$$\frac{1}{4\pi}\int d\Omega\, Y^*_{jm}(\theta,\phi)Y_{j'm'}(\theta,\phi) = \delta_{jj'}\delta_{mm'}. \tag{5.90}$$

To construct matrix counterparts of spherical harmonics $\hat{Y}_{jm}$, we note that

$$Y_{j(m+1)} = \frac{L^+ Y_{jm}}{\sqrt{(j-m)(j+1+m)}}, \tag{5.91}$$

where $L^\pm = L^1 \pm iL^2$, so *(iii)* requires (denote $J^\pm = J^1 \pm iJ^2$)

$$\hat{Y}_{j(m+1)} = \frac{[J^+, \hat{Y}_{jm}]}{\sqrt{(j-m)(j+1+m)}}, \tag{5.92}$$

which fixes all the matrices $\hat{Y}_{jm}$ given $\hat{Y}_{j(-j)}$. The $\mathfrak{su}(2)$ representation further requires that $L^- Y_{j(-j)} = 0$ and $L^+ Y_{jj} = 0$, which translates to the matrix side as $[J^-, Y_{j(-j)}] = 0$ and $[J^+, Y_{jj}] = 0$. Thus for some normalizing factor C,

$$\hat{Y}_{j(-j)} = C(J^-)^j. \tag{5.93}$$

The matrix J^- is nilpotent with order N: $(J^-)^N = 0$. Therefore the matrices in (5.93) are restricted to $j \leq N-1$. For $j \leq N-1$, the numerical factor C is chosen such that

$$\frac{1}{N}\mathrm{tr}\hat{Y}^\dagger_{j(-j)}\hat{Y}_{j(-j)} = 1. \tag{5.94}$$

The sign of C is not fixed by the three requirements, and we pick $C > 0$ in correspondence with spherical harmonics $Y_{j(-j)} \propto (x^1 - ix^2)^j$.

It is straightforward to verify that

$$\sum_{i=1}^{3}[J^i,[J^i,\hat{Y}_{jm}]] = j(j+1)\hat{Y}_{jm}, \quad [J^3,\hat{Y}_{jm}] = m\hat{Y}_{jm}, \tag{5.95}$$

given the $\mathfrak{su}(2)$ algebra and eqs. (5.92) and (5.93). Hence the matrices $\hat{Y}_{jm}$ form an eigenbasis of adjoint actions of J^3 and the casimir $(J^i)^2$, and are therefore orthogonal. They are normalized as well because of (5.94). The map $A \mapsto f_A(\theta,\phi)$ is then defined on the basis as $\hat{Y}_{jm} \mapsto Y_{jm}(\theta,\phi)$, fulfilling the requirements *(i)* to *(iii)*.

Under the correspondence $\hat{Y}_{jm} \mapsto Y_{jm}(\theta,\phi)$, N-by-N matrices describe fields on a sphere with angular momentum cutoff $j_{\max} = N-1$. Furthermore (5.86) connects matrix observables and averages of fields on the emergent sphere. For instance, the classical fuzzy sphere solution sets $X^i = \nu J^i$, and we would like to interpret $f_{X^i}(\theta,\phi)$ as coordinates x^i of the point on the sphere at angle (θ,ϕ). Thus according to (5.86), the radius of the emergent sphere (for irreducible representation J^i) is

$$\begin{aligned} r^2 &= \frac{1}{4\pi}\sum_{i=1}^{3}\int d\Omega\, f_{X^i}(\theta,\phi)^2 = \frac{1}{N}\sum_{i=1}^{3}\mathrm{tr}(X^i)^2 \\ &= \frac{\nu^2}{N}\sum_{i=1}^{3}\mathrm{tr}(J^i)^2 = \frac{\nu^2(N^2-1)}{4}. \end{aligned} \tag{5.96}$$

Noncommutative gauge theory on the fuzzy sphere

In the last subsection we have discussed the correspondence between matrix degrees of freedom and fields on the fuzzy sphere. Given that correspondence the matrix Hamiltonian (5.4) can be cast into a quantum field theory on the sphere. The caveat is that the fields on the sphere are not commutative, due to the noncommutative nature of matrix multiplication.

To be more precise, we define the 'star product' of the fields as induced from their corresponding matrix multiplications:

$$(f\star g)(\theta,\phi) \equiv \frac{1}{N}\sum_{jm}\mathrm{tr}\left(\hat{Y}^\dagger_{jm}\hat{f}\hat{g}\right)Y_{jm}(\theta,\phi), \tag{5.97}$$

where $\hat{f}$ and $\hat{g}$ are the matrix counterparts of functions $f(\theta,\phi)$ and $g(\theta,\phi)$ via the correspondence between matrix spherical harmonics and spherical harmonics on the sphere: $\hat{Y}_{jm} \leftrightarrow Y_{jm}(\theta,\phi)$. The prefactor is a result of the normalization (5.94).

The star product is associative but noncommutative. In particular, the commutator of scalar functions may not vanish. For example,

$$\begin{aligned} [Y_{j_1m_1}, Y_{j_2m_2}]_\star(\theta,\phi) &= \frac{1}{N}\sum_{jm}\mathrm{tr}\left(\hat{Y}^\dagger_{jm}[\hat{Y}_{j_1m_1},\hat{Y}_{j_2m_2}]\right)Y_{jm}(\theta,\phi) \\ &\equiv \sum_{jm} f^{jm}_{j_1m_1j_2m_2}Y_{jm}(\theta,\phi), \end{aligned} \tag{5.98}$$

where $[\cdot,\cdot]_\star$ is the commutator with the star product for multiplication. The structure constants f in (5.98) are known to vanish as $1/N$ as $N\to\infty$ (see, e.g., the Appendix of [118]). The usual commutative product is recovered at $N=\infty$.

To repackage matrix degrees of freedom into emergent fields, expand the bosonic matrices around their classical values:

$$X^i = \nu J^i + A^i, \tag{5.99}$$

where the A^i are Hermitian matrices parametrizing fluctuations around the fuzzy sphere. Our rewriting of the Hamiltonian will be exact in A. The corresponding emergent fields $\widetilde{a}^i(\theta,\phi)$ are as follows:

$$\widetilde{a}^i(\theta,\phi) = \sum_{jm} a^i_{jm} Y_{jm}(\theta,\phi), \quad \text{if } A^i = \sum_{jm} a^i_{jm} \hat{Y}_{jm}. \tag{5.100}$$

The conjugate momenta to the A^i are

$$\Pi^i_A = -\frac{i}{N} \sum_{jm} \hat{Y}^\dagger_{jm} \frac{\partial}{\partial a^i_{jm}}\,, \tag{5.101}$$

obeying the canonical commutation relations $[A^i_{ab}, (\Pi^j_A)_{cd}] = i\delta^{ij}\delta_{ad}\delta_{bc}$. We will also want to introduce the momenta

$$\widetilde{\pi}^i(\theta,\phi) = -\frac{i}{4\pi} \sum_{jm} Y^*_{jm}(\theta,\phi) \frac{\partial}{\partial a^i_{jm}}, \tag{5.102}$$

which obey

$$[\widetilde{a}^i(\theta,\phi), \widetilde{\pi}^k(\theta',\phi')] = \frac{i\delta^{ik}}{4\pi} \sum_{jm} Y_{jm}(\theta,\phi) Y^*_{jm}(\theta',\phi'). \tag{5.103}$$

The $\widetilde{\pi}^i$ therefore become the usual conjugate momenta when $j_{\max} = \infty$, where the summation in (5.103) becomes $4\pi\delta(\cos\theta - \cos\theta')\delta(\phi - \phi')$. Hermiticity of the matrices A^i and Π^i_A is manifested as reality of the fields $\widetilde{a}^i$ and $\widetilde{\pi}^i$.

Substituting (5.100), (5.101) and (5.102) into the matrix Hamiltonian, the kinetic terms are

$$\frac{1}{2}\mathrm{tr}\left(\Pi^i\Pi^i\right) = \frac{1}{2}\mathrm{tr}\left(\Pi^i_A\Pi^i_A\right) = -\frac{1}{2N}\sum_{ijm}\frac{\partial^2}{(\partial a^i_{jm})^2} = \frac{2\pi}{N}\int d\Omega\,(\widetilde{\pi}^i(\theta,\phi))^2. \tag{5.104}$$

The bosonic potential in (5.1) can be written as a square:

$$V(X) = \frac{1}{4}\mathrm{tr}\left(i[X^i,X^j] + \nu\epsilon^{ijk}X^k\right)^2 \equiv \frac{\nu^2}{4}\mathrm{tr}\left(F^{ij}\right)^2, \tag{5.105}$$

and substituting (5.99) into (5.105):

$$F^{ij} = i\left([J^i,A^j] - [J^j,A^i]\right) + i\nu^{-1}[A^i,A^j] + \epsilon^{ijk}A^k. \tag{5.106}$$

The corresponding field is (recall (5.87) and (5.97))

$$\widetilde{f}^{ij}(\theta,\phi) = i\left(L^i\widetilde{a}^j - L^j\widetilde{a}^i\right) + \epsilon^{ijk}\widetilde{a}^k + i\nu^{-1}[\widetilde{a}^i,\widetilde{a}^j]_\star, \tag{5.107}$$

and the potential can now be written

$$V(X) = \frac{N\nu^2}{4} \int \frac{d\Omega}{4\pi} (\widetilde{f}^{ij}(\theta,\phi))^2. \tag{5.108}$$

The fermionic potential in (5.3) is, in terms of A^i,

$$\nu \mathrm{tr} \left(\lambda^\dagger \sigma^k [J^k + \nu^{-1} A^k, \lambda] + \frac{3}{2} \lambda^\dagger \lambda \right) - \frac{3}{2} \nu (N^2 - 1) \,. \tag{5.109}$$

Let $\widetilde{\psi}(\theta,\phi)$ be the fermionc field corresponding to λ, then (5.109) is recast into

$$\frac{N\nu}{4\pi} \int d\Omega \, \left(-i\widetilde{\psi}^\dagger \sigma^k D^k \widetilde{\psi} + \frac{3}{2} \widetilde{\psi}^\dagger \widetilde{\psi} \right) + \text{const}, \tag{5.110}$$

where $D^k \widetilde{\psi} \equiv iL^k \widetilde{\psi} + i\nu^{-1} [\widetilde{a}^k, \widetilde{\psi}]_\star$.

Collect all three parts (5.104), (5.108) and (5.110), and rescale the fields

$$\widetilde{a}^i = \sqrt{\frac{4\pi}{N\nu}} a^i, \quad \widetilde{\pi}^i = \sqrt{\frac{N\nu}{4\pi}} \pi^i, \quad \widetilde{\psi} = \sqrt{\frac{4\pi}{N}} \psi. \tag{5.111}$$

The Hamiltonian for the emergent fields, which is equivalent to (5.4) for matrices, is then

$$H = \nu \int d\Omega \, \left(\frac{1}{2} (\pi^i)^2 + \frac{1}{4} (f^{ij})^2 - i\psi^\dagger \sigma^k D^k \psi + \frac{3}{2} \psi^\dagger \psi \right) + \text{const}, \tag{5.112}$$

where

$$\begin{aligned} f^{ij} &\equiv i \left(L^i a^j - L^j a^i \right) + \epsilon^{ijk} a^k + i \sqrt{\frac{4\pi}{N\nu^3}} [a^i, a^j]_\star, \\ D^k \psi &\equiv iL^k \psi + i \sqrt{\frac{4\pi}{N\nu^3}} [a^k, \psi]_\star. \end{aligned} \tag{5.113}$$

The SU(N) gauge symmetry of the matrices leads to the noncommutative U(1) gauge symmetry of (5.113). Under an infinitesimal SU(N) gauge transformation parametrized by a Hermitian matrix Y, $\delta X^i = i[Y, X^i]$, $\delta \lambda^\alpha = i[Y, \lambda^\alpha]$, and thus by (5.99),

$$\delta A^i = -i[\nu J^i, Y] + i[Y, A^i]. \tag{5.114}$$

Let $\widetilde{y}(\theta,\phi)$ be the field corresponding to the matrix Y, then the gauge transformation of the noncommutative fields is ($\boldsymbol{n}$ is the radial vector and fields should be considered as defined on the unit sphere)

$$\delta \widetilde{a}^i = -i\nu L^i \widetilde{y} - (\boldsymbol{n} \times \nabla \widetilde{y} \cdot \nabla) \widetilde{a}^i, \quad \delta \widetilde{\psi}^\alpha = -(\boldsymbol{n} \times \nabla \widetilde{y} \cdot \nabla) \widetilde{\psi}^\alpha. \tag{5.115}$$

Recall the rescaling (5.111) and let $\widetilde{y} = y\sqrt{4\pi/N\nu^3}$,

$$\delta a^i = -iL^i y - \sqrt{\frac{4\pi}{N\nu^3}}(\boldsymbol{n}\times\nabla y\cdot\nabla)a^i, \quad \delta\psi^\alpha = -\sqrt{\frac{4\pi}{N\nu^3}}(\boldsymbol{n}\times\nabla y\cdot\nabla)\psi^\alpha. \tag{5.116}$$

The first term in δa^i is the usual U(1) transformation. The second term, which can be obtained from the algebra in (5.98), describes a coordinate transformation with infinitesimal displacement $\boldsymbol{n}\times\nabla y$ [62]. Indeed, it is known that non-commutative gauge theories mix internal and spacetime symmetries, which in this case are area-preserving diffeomorphisms of the sphere [193, 153]. The coordinate transformation in (5.116) is area-preserving because $\nabla\cdot(\boldsymbol{n}\times\nabla y) = 0$.

In the commutative limit $\nu\to\infty$, the gauge field is decoupled from the fermions and the theory contains a U(1) gauge field on the sphere, with a real massive scalar and a massive Dirac fermion. To see more explicitly the field content of (5.112) in this limit, note that $\boldsymbol{L} = -i\boldsymbol{n}\times\nabla$ and $f^{ij} = \epsilon^{ijk}\left((\boldsymbol{n}\times\nabla)\times\boldsymbol{a}+\boldsymbol{a}\right)^k$ when $\nu\to\infty$ ($\boldsymbol{a}$ is the three-dimensional vector notation for a^i). We then obtain

$$\frac{1}{4}(f^{ij})^2 = \frac{1}{2}\left|(\boldsymbol{n}\times\nabla)\times\boldsymbol{a}+\boldsymbol{a}\right|^2. \tag{5.117}$$

The scalar field φ is the radial component of the gauge field, and we denote the U(1) gauge field on the sphere as $\boldsymbol{b}$:

$$\varphi = \boldsymbol{a}\cdot\boldsymbol{n}, \quad \boldsymbol{b} = \boldsymbol{a}\times\boldsymbol{n}. \tag{5.118}$$

The U(1) curvature f of the gauge field $\boldsymbol{b}$ defined on the sphere is

$$f = \boldsymbol{n}\cdot(\nabla\times\boldsymbol{b}) = 2\boldsymbol{n}\cdot\boldsymbol{a} - \nabla\cdot\boldsymbol{a}, \tag{5.119}$$

and we have (after some vector calculus manipulations)

$$(\boldsymbol{n}\times\nabla)\times\boldsymbol{a}+\boldsymbol{a} = f\boldsymbol{n} + \nabla(\boldsymbol{n}\cdot\boldsymbol{a}) - \boldsymbol{n}(\boldsymbol{n}\cdot\boldsymbol{a}) = (f-\varphi)\boldsymbol{n} + \nabla\varphi. \tag{5.120}$$

Substituting (5.120) into (5.117), the commutative gauge theory can be rewritten as

$$H = \nu\int d\Omega\left(\frac{1}{2}(\pi^a)^2 + \frac{1}{2}\pi^2 + \frac{1}{2}(f-\varphi)^2 + \frac{1}{2}(\nabla\varphi)^2 - i\psi^\dagger(\boldsymbol{\sigma}\times\boldsymbol{n})\cdot\nabla\psi + \frac{3}{2}\psi^\dagger\psi\right), \tag{5.121}$$

where π^a and π are the conjugate variables of $\boldsymbol{b}$ and φ, respectively, and $\boldsymbol{\sigma}$ is the vector of Pauli matrices. The fields in (5.121) should be thought as living on the unit sphere.

Fluctuation spectrum around the classical fuzzy sphere

The classical energy at the fuzzy sphere vanishes due to supersymmetry. In the following we analyze the spectrum of bosonic quadratic fluctuations near the fuzzy sphere configuration, and the spectrum of fermions, as the next order in a semiclassical expansion. The semiclassical correction to energy

at this level is shown to be zero as well.

The bosonic potential in (5.1) can be written as a square:

$$V(X) = \frac{1}{2}\mathrm{tr}\left(\nu X^i + i\epsilon^i{}_{jk}X^jX^k\right)^2, \tag{5.122}$$

and quadratic fluctuations around a classical solution are given by

$$\begin{aligned}\delta V(X) &= \frac{1}{2}\mathrm{tr}\left(\nu\delta X^i + i\epsilon^i{}_{jk}[X^j,\delta X^k]\right)^2 \\ &\equiv \sum_a \frac{1}{2}\nu^2\omega_a^2(\delta x_a)^2,\end{aligned} \tag{5.123}$$

where $\delta X^i = \sum_a \delta x_a Y_a^i$ and Y_a^i are the normalized eigen-matrices:

$$Y_a^i + i\epsilon^i{}_{jk}[J^j, Y_a^k] = \omega_a Y_a^i, \quad \sum_{i=1}^3 \mathrm{tr}[(Y_a^i)^\dagger Y_b^i] = \delta_{ab}. \tag{5.124}$$

Here we specialized to the background solution $X^j = \nu J^j$.

To solve the eigenvalue equation in (5.124), expand Y^i (subscript a omitted) into a sum of matrix spherical harmonics $Y^i = \sum_{jm} y^i_{jm}\hat{Y}_{jm}$, and note

$$\begin{aligned}\sum_{i=1}^3 [J^i,[J^i,\hat{Y}_{jm}]] = j(j+1)\hat{Y}_{jm}, \quad [J^+,\hat{Y}_{jm}] = \sqrt{(j-m)(j+m+1)}\hat{Y}_{j(m+1)}, \\ [J^3,\hat{Y}_{jm}] = m\hat{Y}_{jm}, \quad [J^-,\hat{Y}_{jm}] = \sqrt{(j+m)(j-m+1)}\hat{Y}_{j(m-1)}.\end{aligned} \tag{5.125}$$

For convenience introduce the $\pm$ basis: $y^\pm = y^1 \pm iy^2$ and the indices must be raised with $g^{+-} = g^{-+} = 2$ and $g^{33} = 1$ (other entries are zero). In this basis $\epsilon_{+-3} = i/2$. Then (5.124) can be cast into equations for the coefficients y^3_{jm} and $y^\pm_{jm}$:

$$y^3_{jm} + \frac{1}{2}\sqrt{(j+m+1)(j-m)}y^+_{j(m+1)} - \frac{1}{2}\sqrt{(j-m+1)(j+m)}y^-_{j(m-1)} = \omega y^3_{jm}, \tag{5.126}$$

$$(\omega \pm m)y^\pm_{j(m\pm1)} = \pm\sqrt{(j\pm m+1)(j\mp m)}y^3_{jm}. \tag{5.127}$$

Equations (5.126) and (5.127) consist of three linear equations with three variables y^3_{jm}, $y^+_{j(m+1)}$ and $y^-_{j(m-1)}$. For there to be nonzero solutions, the determinant must be zero:

$$\omega(\omega+j)(\omega-j-1) = 0. \tag{5.128}$$

Hence for $0 < j < N$, $|m| < j$, the eigenvalues are $\omega = 0, -j, j+1$. The edge cases $|m| = j, j+1$ should be treated separately due to the additional constraint $y^\pm_{jm} = 0$ if $|m| > j$. The eigenvalue

equation at $m = \pm j$ is instead $\omega(\omega - j - 1) = 0$, and for $m = \pm(j+1)$ it is $\omega - j - 1 = 0$.

The multiplicity of the eigenvalue $\omega = 0$ is $N^2 - 1$, which accounts for the SU(N) gauge degrees of freedom. The other eigenvalues are $\omega = -j$ for $1 \leq j \leq N-1$ with multiplicity $2j-1$ and $\omega = j+1$ for $1 \leq j \leq N-1$ with multiplicity $2j+3$. The ground state energy of the bosonic oscillators (5.123) is therefore

$$\frac{|\nu|}{2}\sum_a |\omega_a| = \frac{|\nu|}{2}\sum_{j=1}^{N-1}[j(2j-1) + (j+1)(2j+3)] = \frac{4N^3+5N-9}{6}|\nu|. \tag{5.129}$$

The spectrum of the fermionic bilinear is found similarly:

$$(\sigma^k)^\alpha_\beta[J^k, \lambda^\beta] + \frac{3}{2}\lambda^\alpha = \omega\lambda^\alpha. \tag{5.130}$$

Expand $\lambda^\alpha = \sum_{jm} y^\alpha_{jm}\hat{Y}_{jm}$ (note now $\alpha = \pm$ labels $\sigma^3 = \pm 1$ basis). The equations are

$$\left(\omega - m - \frac{3}{2}\right) y^+_{jm} = \sqrt{(j+m+1)(j-m)}\, y^-_{j(m+1)}, \tag{5.131}$$

$$\left(\omega + m - \frac{1}{2}\right) y^-_{j(m+1)} = \sqrt{(j+m+1)(j-m)}\, y^+_{jm}. \tag{5.132}$$

The eigenvalue equations (5.131) and (5.132) have nontrivial solutions when

$$\left(\omega - j - \frac{3}{2}\right)\left(\omega + j - \frac{1}{2}\right) = 0, \tag{5.133}$$

so that for $0 < j < N$ and $-j \leq m < j$ there are eigenvalues $\omega = j + 3/2$ and $\omega = -j + 1/2$. For $m = j$ or $m = -j-1$ the eigenvalue equation is instead $\omega - j - 3/2 = 0$, as $y^-_{j(j+1)} = y^+_{j(-j-1)} = 0$ is imposed.

So the eigenvalues for $0 < j < N$ are $\omega = j + 3/2$ with multiplicity $2j+2$ and $\omega = -j + 1/2$ with multiplicity $2j$. For $\nu > 0$ the $\omega = -j + 1/2$ modes are occupied with a total number of fermions:

$$\sum_{j=1}^{N-1}(2j) = N^2 - N. \tag{5.134}$$

And the fermionic energy for $\nu > 0$ at this order is

$$\nu\sum_{j=1}^{N-1}\left(-j + \frac{1}{2}\right)(2j) - \frac{3}{2}\nu(N^2-1) = -\frac{4N^3+5N-9}{6}\nu. \tag{5.135}$$

For $\nu < 0$ the $\omega = j + 3/2$ modes are occupied instead and the number of fermions is

$$\sum_{j=1}^{N-1}(2j+2) = N^2 + N - 2. \tag{5.136}$$

We see that supersymmetry requires different number of occupied fermions in the case of $\nu > 0$ and $\nu < 0$. The fermionic energy for $\nu < 0$ is

$$\nu \sum_{j=1}^{N-1}\left(j+\frac{3}{2}\right)(2j+2) - \frac{3}{2}\nu(N^2-1) = \frac{4N^3+5N-9}{6}\nu. \tag{5.137}$$

In either case (5.135) or (5.137) the energy is $-(4N^3 + 5N - 9)|\nu|/6$, which exactly cancels the bosonic contribution (5.129). Hence the semiclassical correction to the fuzzy sphere energy is zero at this order, for the specific number of fermions (5.134) or (5.136).

One-loop effective potential and the estimate of ν_c

In the main text we observe a first-order phase transition near $\nu_c \approx 4$ when the bosonic fuzzy sphere phase becomes unstable. Here we give an estimate of ν_c from the bosonic one-loop effective potential for the radius, at $N = \infty$.

We start with the bosonic potential (5.122) with matrix sources S_i:

$$V(X; S_i) = \frac{1}{2}\mathrm{tr}\left(\nu X^i + i\epsilon^i{}_{jk}X^jX^k\right)^2 + \mathrm{tr}S_iX^i, \tag{5.138}$$

where the sources $S_i(\phi)$ are such that the local energy minimum is at $X^i = \phi J^i$. The parameter $\phi > 0$ is proportional to the radius:

$$r = \frac{\phi}{2}\sqrt{N^2-1}. \tag{5.139}$$

The classical contribution to the energy (5.138) at $X^i = \phi J^i$ is

$$E_0(S_i(\phi)) = \frac{N(N^2-1)}{8}(\nu-\phi)^2\phi^2 + \mathrm{tr}S_i(\phi)\phi J^i. \tag{5.140}$$

Quadratic fluctuations of (5.138) around the local minimum give:

$$\delta V(X) = \frac{1}{2}\mathrm{tr}\left(\nu\delta X^i + i\phi\epsilon^i{}_{jk}[J^j, \delta X^k]\right)^2 + i\epsilon^i{}_{jk}(\nu-\phi)\phi\mathrm{tr}\left(J^i\delta X^j\delta X^k\right). \tag{5.141}$$

The norm of the spin matrices J^i scales as N, and hence to leading order in N:

$$\delta V(X) = \frac{1}{2}\mathrm{tr}\left(i\phi\epsilon^i{}_{jk}[J^j, \delta X^k]\right)^2 + \dots. \tag{5.142}$$

Diagonalizing this leading order piece as we did in the last subsection, the nonzero mode frequencies are now $\omega = -(j+1)\phi$ for $0 < j < N$ with multiplicity $2j-1$ and $\omega = j\phi$ for $0 < j < N$ with multiplicity $2j+3$. So, the one-loop quantum correction to the ground state energy is

$$\frac{1}{2}\sum_a |\omega_a| = \frac{1}{2}\sum_{j=1}^{N-1} \left[|-(j+1)\phi|\,(2j-1) + |j\phi|\,(2j+3)\right] + \ldots = \frac{2}{3}\phi N^3 + \ldots. \tag{5.143}$$

The one-loop effective potential $\Gamma(\phi) = E_0(S_i(\phi)) + \frac{1}{2}\sum_a |\omega_a| - \mathrm{tr} S_i(\phi)\phi J^i$ is then

$$N^{-3}\Gamma(\phi;\nu) = \frac{1}{8}(\nu-\phi)^2\phi^2 + \frac{2}{3}\phi + \ldots, \tag{5.144}$$

where omitted terms are higher order in N^{-1}. The critical value of ν is estimated as when the second order derivative of $\Gamma(\phi)$ at the fuzzy sphere solution vanishes:

$$\Gamma'(\phi;\nu_c) = \Gamma''(\phi;\nu_c) = 0, \quad \Rightarrow \quad \nu_c \approx 3.03,\ \phi \approx 2.39. \tag{5.145}$$

It is clear in (5.144) that, at large N, the leading quantum correction to the classical solution is suppressed by ν^{-3}. This shows that the large ν limit rapidly becomes classical. The critical ν_c estimated above is at $N = \infty$, where the transistion is sharp.

5.10 Appendix D: Training and tuning

Training of the model is divided into three epochs, each of which consists of 5000 iterations. The learning rate is set to be 10^{-3} for iterations from 1 to 5000, 2×10^{-4} from 5001 to 10000 and 4×10^{-5} from 10001 to 15000. In each iteration the energy is evaluated from a batch of 10^3 random samples, and while the Monte Carlo energy fluctuates among iterations, its average value converges. Some typical training histories are shown in Fig. 5.12.

The final energy of the trained variational wavefunction is evaluated from 5 million samples, with Monte Carlo uncertainties shown as error bars in Figs. 5.13, 5.14, 5.15 and 5.16. In these figures we compare performance of various architectures and observe that

- MAF obtains lower energies for small ν and NF has lower energies at larger ν.
- The result does not significantly depend on the initialization for small ν.
- In the supersymmetric sector the variational energy is close to zero (compared to a typical energy scale, say the bosonic energies).
- Consistent improvement is observed in MAFs if we increase the number of distributions in the mixture or D as in the fermionic wavefunction. However, increasing the number of layers in neural networks does not improve the results.

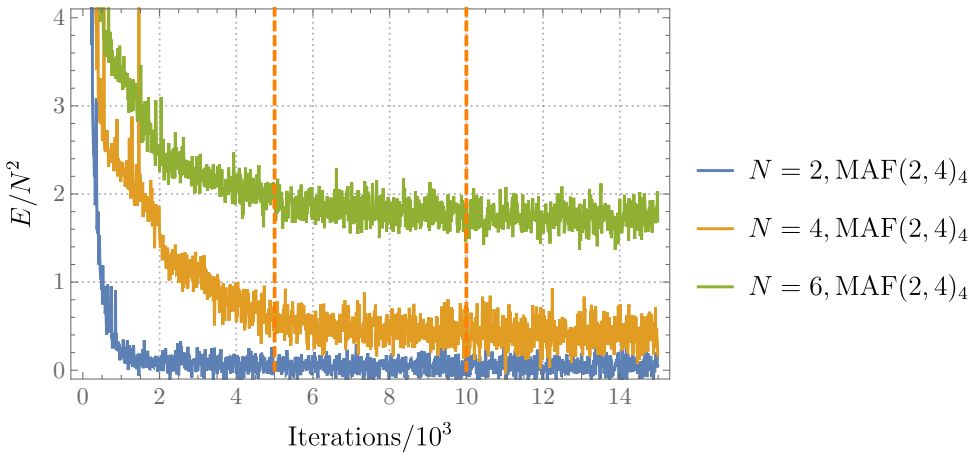


Figure 5.12: The variational energy as a function of training iterations for $N = 2, 4, 6$, with $\nu = 2$ and architecture MAF(2, 4) — the subscript is $D = 4$ as in (5.8). The dashed lines separate the three phases.

5.11 Appendix E: Entanglement of free fields on a sphere

Solution for the projector

We wish to solve the following optimization problem: find an orthogonal projection operator P such that $\|P - Q\|$ is minimal given another Hermitian operator Q. We will now do this in the case that $\|\cdot\|$ is the Frobenius norm. In this case, diagonalize $Q = UQ'U^\dagger$ such that Q' is diagonal with diagonal elements nonincreasing. Then $\|P - Q\|$ is minimized if and only if $\|P' - Q'\|$ is minimized and $P = UP'U^\dagger$.

Firstly we search for P' that minimizes $\|P' - Q'\|$ in the subspace of projectors with fixed rank r. It is equivalent to maximizing $\mathrm{tr}(P'Q')$ by definition of the Frobenius norm. Let $F(V) = \mathrm{tr}(VP'V^\dagger Q')$ for unitary V. If P' maximizes $\mathrm{tr}(P'Q')$, $dF = 0$ at $V = I$ for any dV in the Lie algebra of the unitary group:

$$dF = \mathrm{tr}P'[Q', dV] = 0. \tag{5.146}$$

If Q' is diagonal with distinct eigenvalues, (5.146) implies that P' should be diagonal as well. Then the P' that maximizes $\mathrm{tr}(P'Q')$ should be such that $(P')_{ii} = 1$ for $1 \le i \le r$ and 0 otherwise, and the minimal value of $\|P - Q\|$ is

$$\min_{P^\dagger = P, P^2 = P}^{\mathrm{tr}P = r} \|P - Q\|^2 = \sum_{1 \le i \le r} (1 - Q'_{ii})^2 + \sum_{i > r} (Q'_{ii})^2. \tag{5.147}$$

The projector P that achieves the minimum is unique when Q' has distinct eigenvalues; if Q' is degenerate, there may also be nondiagonal P' matrices that attain the minimal $\|P - Q\|$.

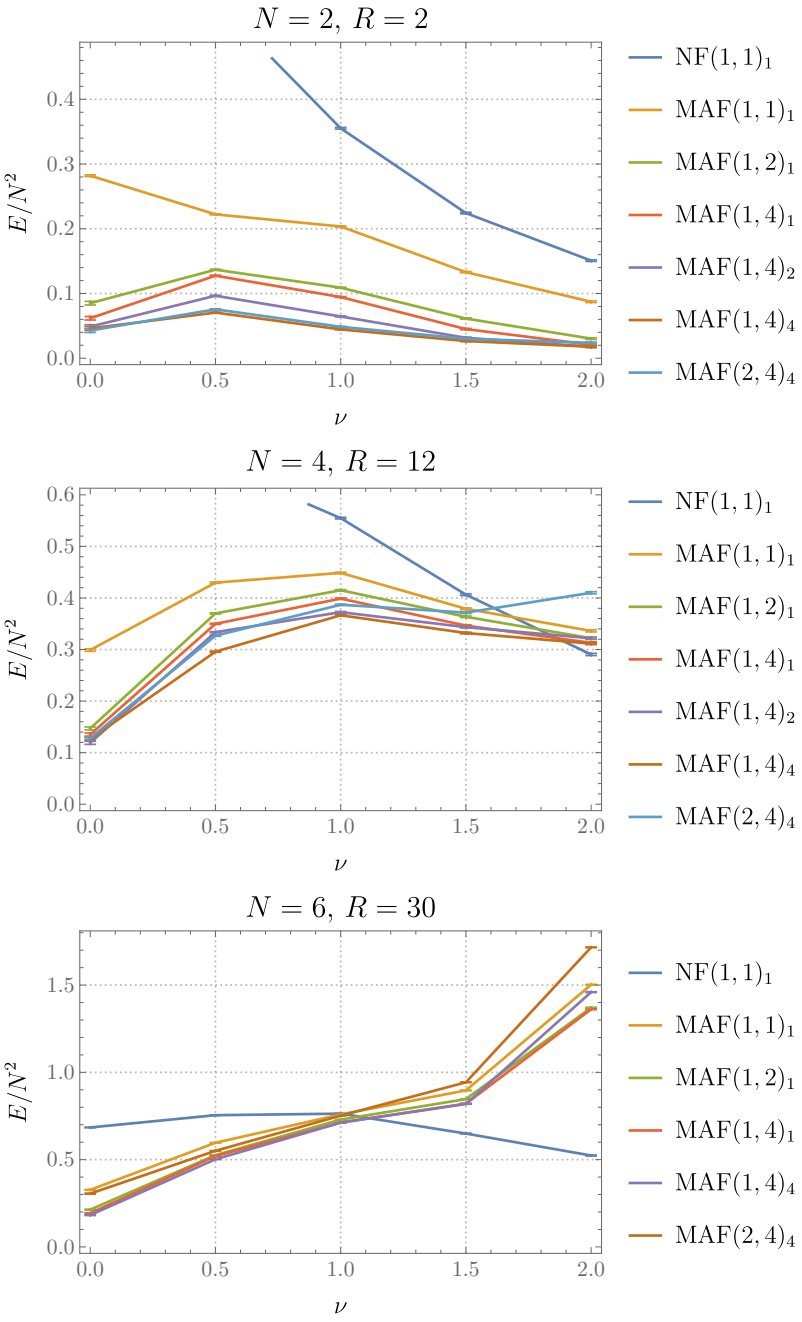


Figure 5.13: The variational energy for different N, ν and MAF architectures, in the supersymmetric sector. The wavefunctions are initialized near zero. Error bars (largely invisible) are Monte Carlo uncertainties of the final energy.

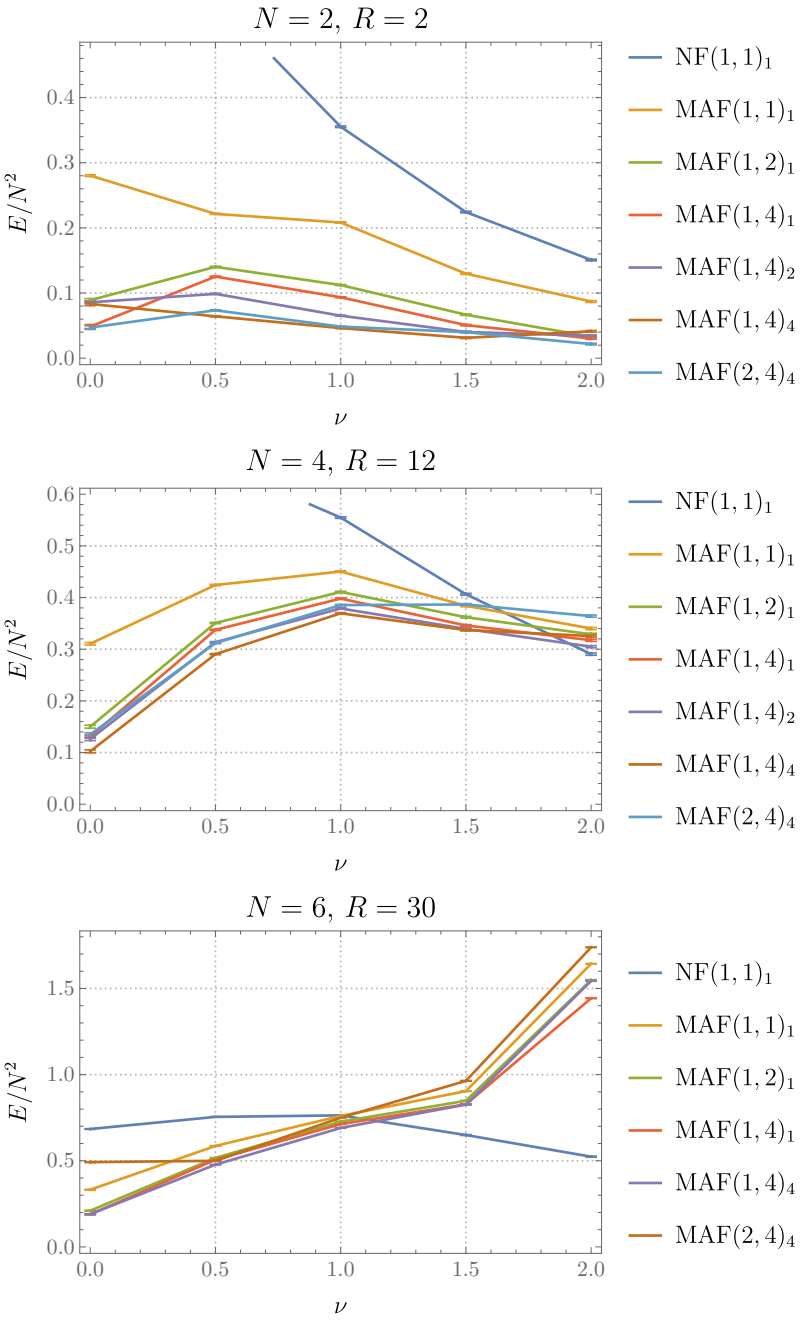


Figure 5.14: The variational energy for different N, ν and MAF architectures, in the supersymmetric sector. The wavefunctions are initialized near the fuzzy sphere. Error bars (largely invisible) are Monte Carlo uncertainties of the final energy.

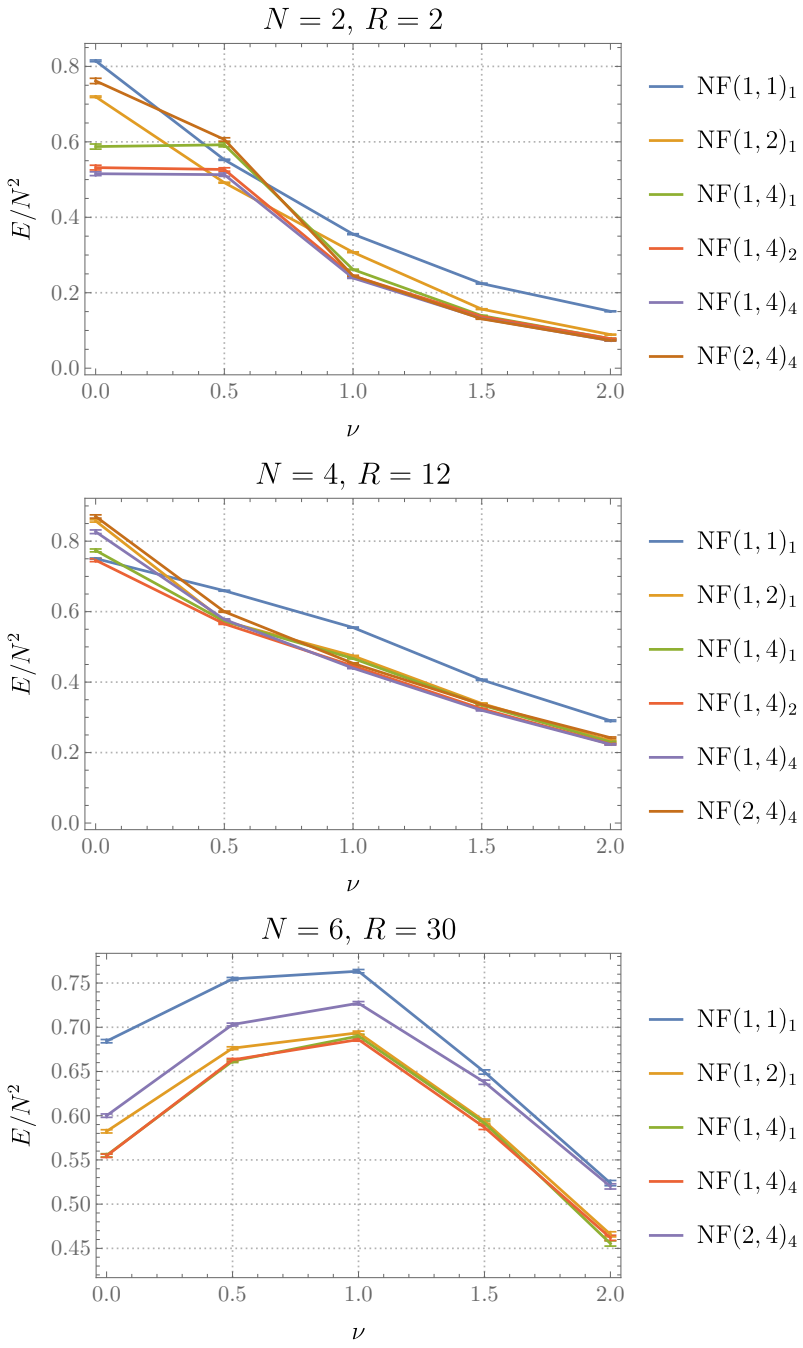


Figure 5.15: The variational energy for different N, ν and NF architectures, in the supersymmetric sector. The wavefunctions are initialized near zero. Error bars (largely invisible) are Monte Carlo uncertainties of the final energy.

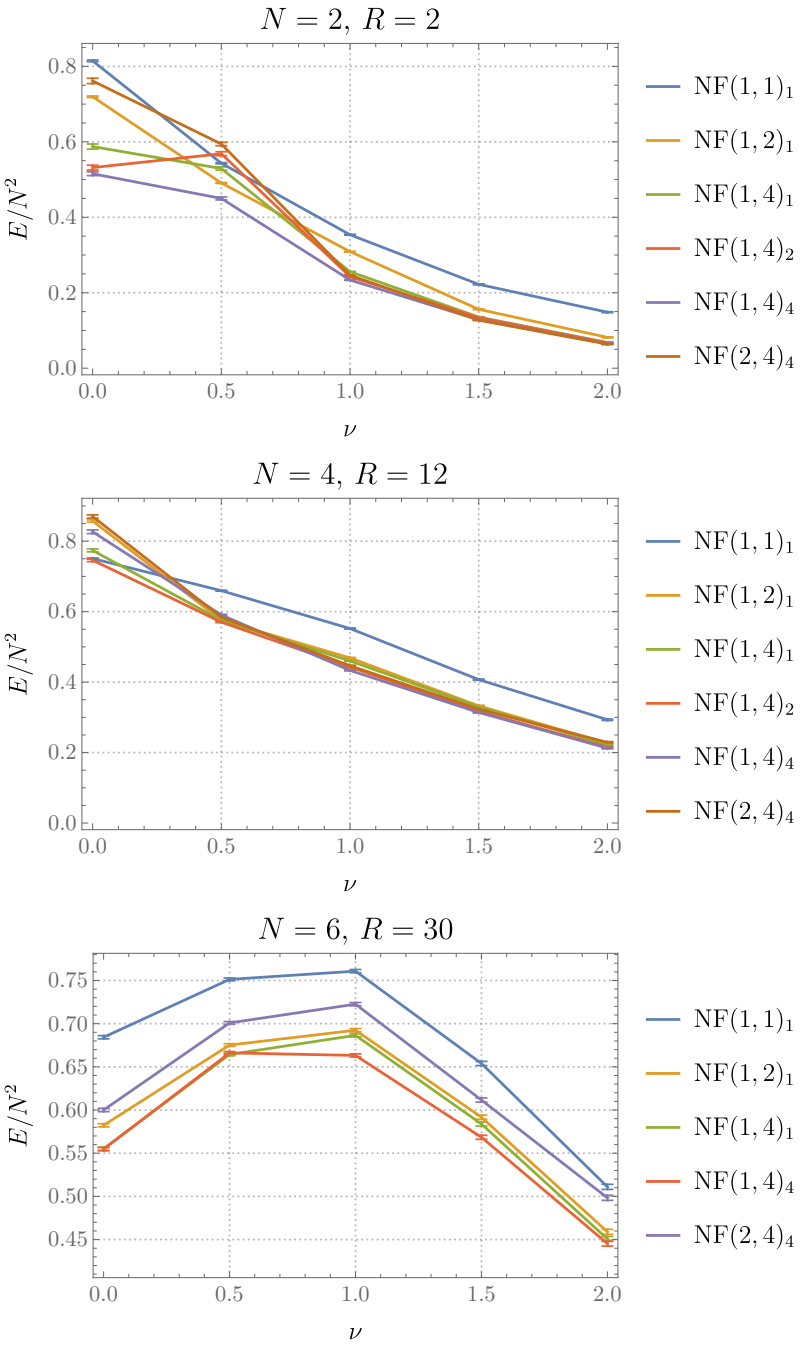


Figure 5.16: The variational energy for different N, ν and NF architectures, in the supersymmetric sector. The wavefunctions are initialized near the fuzzy sphere. Error bars are Monte Carlo uncertainties of the final energy.

The second step is to minimize (5.147) with respect to the rank r. If $Q'_{ii} \neq 1/2$, the rank should be the number of eigenvalues of Q that are above $1/2$. The minimum is then

$$\min_{P^\dagger=P, P^2=P} \|P-Q\|^2 = \sum_i \min\{(1-Q'_{ii})^2, (Q'_{ii})^2\}. \tag{5.148}$$

When one half is among the eigenvalues, there are multiple P's that minimize $\|P-Q\|$.

To summarize, let $Q = UQ'U^\dagger$ such that U is unitary and Q' is diagonal. Then the following P minimizes $\|P-Q\|_F$ among orthogonal projectors:

$$P = UP'U^\dagger, \quad P' \text{ is diagonal with } P'_{ii} = 1 \text{ if } Q'_{ii} > 1/2, \text{ and 0 otherwise.} \tag{5.149}$$

And this is the unique minimum if none of the eigenvalues of Q is $1/2$.

Evaluation of the second Rényi entropy

As discussed in the main text, in the case where the configuration space Q has a linear structure, an orthogonal decomposition $Q = Q_1 \oplus Q_2$ induces a factorization of the Hilbert space $L^2(Q) = L^2(Q_1) \otimes L^2(Q_2)$. For any pure state $|\psi\rangle \in L^2(Q)$, the entanglement entropy is computed as $S(\rho_1)$, where ρ_1 is the reduced density matrix of the subsystem $L^2(Q_1)$. For numerical simplicity, we now focus on the Rényi entropy (of order $\alpha \geq 0$):

$$S_\alpha(\rho) = \frac{1}{1-\alpha} \ln \mathrm{tr} \rho^\alpha. \tag{5.150}$$

The von Neumann entropy is recovered as the limiting case $\alpha \to 1$. And in the following consider $\alpha = 2$ for concreteness; similar methods and arguments apply to the Rényi entropies of integer orders $\alpha \geq 2$.

The decomposition $Q = Q_1 \oplus Q_2$ can be implicitly specified by an orthogonal projection operator $P : Q \to Q$, such that $Q_1 = \mathrm{im}\, P$ and $Q_2 = \ker P$. For a pure state $|\psi\rangle \in L^2(Q)$, the reduced density matrix ρ_1 is

$$\rho_1(x, x') = \int dy\, \psi(x+y)\psi^*(x'+y), \tag{5.151}$$

where $x, x' \in Q_1 = \mathrm{im}\, P$ and the integral is over the subspace $Q_2 = \ker P$. Consequently the second Rényi entropy is

$$S_2(\rho_1) = -\ln \int dx dx' dy dy'\, \psi(x+y)\psi^*(x'+y)\psi(x'+y')\psi^*(x+y'). \tag{5.152}$$

To further simplify the integral, let $z = x + y \in Q$ and $z' = x' + y' \in Q$, so that

$$x = Pz, \quad x' = Pz', \quad y = (I-P)z, \quad y' = (I-P)z'. \tag{5.153}$$

Thus the integral in (5.152) can be done over the full space Q instead:

$$S_2(\rho_1) = -\ln \int dz dz' \, \psi(z)\psi^*(Pz' + (I-P)z)\psi(z')\psi^*(Pz + (I-P)z'). \tag{5.154}$$

Numerically the integral in (5.154) can be estimated by Monte Carlo:

$$S_2(\rho_1) = -\ln \mathbb{E}_{z,z' \sim |\psi|^2} \left[\frac{\psi^*(Pz' + (I-P)z)\psi^*(Pz + (I-P)z')}{\psi^*(z)\psi^*(z')} \right], \tag{5.155}$$

where in the square bracket, the overall normalization of the wavefunction is unimportant.

The integral in (5.154) is analytically tractable for Gaussian states:

$$\psi(x) = \frac{1}{Z} \exp(-x^\dagger V x), \tag{5.156}$$

where V is some positive definite matrix and Z is the normalization factor. Up to numerical factors, for any positive definite matrix A,

$$\int dx \, \exp(-x^\dagger A x) \propto (\det A)^{-1}. \tag{5.157}$$

Substituting (5.156) into (5.154) and performing the integral using (5.157), for Gaussian pure states, one obtains

$$S_2(\rho_1) = \ln(\det R / \det S), \tag{5.158}$$

where

$$\begin{aligned} R &= \begin{pmatrix} 2V + 2PVP - PV - VP & PV + VP - 2PVP \\ VP + PV - 2PVP & 2V + 2PVP - PV - VP \end{pmatrix}, \\ S &= \begin{pmatrix} 2V & 0 \\ 0 & 2V \end{pmatrix}. \end{aligned} \tag{5.159}$$

The factor of $\det S$ comes from the normalization Z in (5.156). It is simpler to write

$$\begin{aligned} S_2(\rho_1) &= \ln \det \sqrt{S^{-1}} R \sqrt{S^{-1}} = \ln \det \begin{pmatrix} I+K & -K \\ -K & I+K \end{pmatrix} \\ &= \ln \det(I + 2K) = \operatorname{tr} \ln(I + 2K), \end{aligned} \tag{5.160}$$

where

$$K = \sqrt{V^{-1}} PVP \sqrt{V^{-1}} - \frac{1}{2} \left(\sqrt{V^{-1}} P \sqrt{V} + \sqrt{V} P \sqrt{V^{-1}} \right). \tag{5.161}$$

In the next subsection, geometric features of entanglement for free fields are understood analytically from the formulae (5.160) and (5.161).

Derivation of the geometric features of entanglement

Consider a free field on a sphere as in (5.35) with angular momentum cutoff $j \leq j_{\max}$. The ground state is a Gaussian state (5.156) with V diagonal in the basis of spherical harmonic modes with eigenvalues $\sqrt{j(j+1)+\mu^2}$ and multiplicities $2j+1$. The projector P is the one that minimizes $\|P - \chi_A\|$, with the region A being a spherical cap with polar angle θ_A. We would like to confirm the following numerical findings with analytic computations: as $j_{\max} \to \infty$, *(i)* $S_2 \propto j_{\max} \sin\theta_A \propto j_{\max}|\partial A|$ and *(ii)* $\mathrm{tr}P \propto j_{\max}^2 \int_0^{\theta_A} \sin\theta d\theta \propto j_{\max}^2 |A|$.

To start, observe that from (5.160) naively we would expect $S_2 \sim (j_{\max})^2$ because of the trace, and thus if $S_2 \sim j_{\max}$ it must be the case that the matrix K is small. Hence it is reasonable to make the approximation

$$S_2 \approx 2\mathrm{tr}K = 2\mathrm{tr}PVPV^{-1} - 2\mathrm{tr}P. \tag{5.162}$$

In terms of matrix elements of the projector, (recall that $P^\dagger = P$ and $P^2 = P$)

$$S_2 \approx \sum_{jj'm} |P_{jm,j'm}|^2 \frac{(j-j')^2}{jj'}, \tag{5.163}$$

where we have noticed that the projector preserves the J^z quantum number because of the symmetry of region A. Also the eigenvalues of V are approximated as j. Subleading terms will not modify the scaling as $j_{\max} \to \infty$, where j is typically large.

For $j, j' \ll j_{\max}$, the projector $P_{jm,j'm}$ should converge to its value at infinite $j_{\max}$, which is the matrix element of multiplication by χ_A:

$$P_{jm,j'm} \sim \frac{1}{4\pi} \int_0^{\theta_A} d\theta \sin\theta \int_0^{2\pi} d\phi\, Y^*_{jm}(\theta,\phi) Y_{j'm}(\theta,\phi), \tag{5.164}$$

where χ_A restricts the θ integral to $[0, \theta_A]$. Up to numerical factors,

$$P_{jm,j'm} \propto \sqrt{\frac{(2j+1)(2j'+1)(j-m)!(j'-m)!}{(j+m)!(j'+m)!}} \int_{\cos\theta_A}^1 dx\, P_j^m(x) P_{j'}^m(x), \tag{5.165}$$

where $P_j^m(x)$ are associated Legendre polynomials.

The asymptotic form of associated Legendre polynomials $P_j^{-m}(x)$ in the limit $j, m \to \infty$ with $\alpha = m/(j+1/2)$ fixed $(0 < \alpha < 1)$ is given by the WKB formulae eqs. (3.28) and (3.30) in [240]: for $\beta = \sqrt{1-\alpha^2}$ and $\beta < x \leq 1$,

$$P_j^{-m}(x) \sim \Lambda^{jm} (x^2 - \beta^2)^{-1/4} e^{(j+1/2)\chi_1^{jm}(x)}, \tag{5.166}$$

while for $0 \leq x < \beta$,

$$P_j^{-m}(x) \sim 2\Lambda^{jm}(\beta^2 - x^2)^{-1/4} \cos\left(\left(j + \frac{1}{2}\right)\chi_2^{jm}(x) - \frac{\pi}{4}\right), \tag{5.167}$$

where

$$\begin{aligned}
\Lambda^{jm} &= \frac{1}{\sqrt{\pi(2j+1)}}\sqrt{\frac{(j-m)!}{(j+m)!}}, \\
\chi_1^{jm}(x) &= \cosh^{-1}\left(\frac{x}{\beta}\right) - \alpha\cosh^{-1}\left(\frac{\alpha x}{\beta\sqrt{1-x^2}}\right) < 0, \\
\chi_2^{jm}(x) &= \cos^{-1}\left(\frac{x}{\beta}\right) - \alpha\cos^{-1}\left(\frac{\alpha x}{\beta\sqrt{1-x^2}}\right) > 0.
\end{aligned} \tag{5.168}$$

Let $x = \cos\theta$. At large j the oscillating region of the integral in (5.165), where (5.167) holds, is $0 < \alpha < \sin\theta$. Outside of this region, the Legendre polynomial is approximately (5.166), and hence exponentially small. We need therefore only consider the region where both Legendre polynomials are oscillating. In order to get the parametric dependence of observables right, we can furthermore restrict attention to $m \ll j, j'$. In this limit $\beta \to 1$, $\alpha \to 0$ and hence

$$\chi_2^{jm}(x) = \theta. \tag{5.169}$$

So in this limit the integrand in (5.165) can be approximated as

$$dx\, P_j^{-m}(x) P_{j'}^{-m}(x) = d\theta\, 2\Lambda^{jm}\Lambda^{j'm}\cos\left[(j-j')\theta\right] + \cdots . \tag{5.170}$$

The terms $\cdots$ necessarily oscillate strongly at large j, j' and will not contribute to leading order. In the remaining term in (5.170), in contrast, the oscillations are slower when $j \sim j'$. Performing the integral we obtain

$$P_{j(-m),j'(-m)} \propto \frac{\sin\left[(j-j')\theta_A\right]}{j-j'}\,. \tag{5.171}$$

The lower limit of integration (at $m = [\min(j,j') + 1/2]\sin\theta$) can be ignored so long as $m \ll \min(j,j')\sin\theta_A$. This is stronger than the previous assumption $m \ll j, j'$. We can now use (5.171) to evaluate observables, using the fact that $P_{j(-m),j'(-m)} = P_{jm,j'm}$.

The Rényi entropy (5.163) is now (with $j_{\rm m} = \min(j,j')$)

$$\begin{aligned}
S_2 &\propto \sum_{jj'}^{|m|\ll j_{\rm m}\sin\theta_A} \frac{\sin^2[(j-j')\theta_A]}{jj'} && (5.172)\\
&\propto \int^{j_{\max}} \frac{dj'}{j'}\int^{j'} dj\,\sin(\theta_A)\sin^2[(j-j')\theta_A] && (5.173)\\
&\propto j_{\max}\sin(\theta_A)\,. && (5.174)
\end{aligned}$$

In the second line we used $j_{\rm m} \sin\theta_A$ as a cutoff on the sum over m, to get an estimate of the scaling with $\sin\theta_A$. This is the boundary law entanglement that was observed numerically in the main text.

To get the rank of the projector one must treat the sum over m a little more carefully. In particular, we refrain from taking $\alpha \to 0$, $\beta \to 1$. Keeping $\alpha = m/(j+1/2)$,

$$\begin{aligned} {\rm tr}P &= \sum_{jm} P_{jm,jm} && (5.175)\\ &\propto \sum_{jm} \int_{\arcsin|\alpha|}^{\theta_A} \frac{\sin(\theta)d\theta}{\sqrt{\sin(\theta)^2-\alpha^2}} + \cdots . && (5.176)\end{aligned}$$

Here $\cdots$ again denote terms that oscillate strongly in the large j limit and are therefore subleading. The integrand in the second line is directly the non-oscillating part of (5.167) squared. At large $j_{\rm max}$ we therefore have, approximating the sums as integrals and letting $\alpha = \sin\gamma$,

$$\begin{aligned} {\rm tr}P &\propto j_{\rm max}^2 \int_0^{\theta_A} d\gamma \int_\gamma^{\theta_A} d\theta \frac{\sin(\theta)\cos(\gamma)}{\sqrt{\sin(\theta)^2-\sin(\gamma)^2}} && (5.177)\\ &\propto j_{\rm max}^2 \int_0^{\theta_A} d\theta \sin(\theta)\,. && (5.178)\end{aligned}$$

The integrals are most easily done by exchanging the order of integration to $\int_0^{\theta_A} d\theta \int_0^\theta d\gamma$. This result shows that the rank of the projector goes like the area of the region on the sphere, as seen numerically in the main text. The prefactor in the final result (5.178) is easily restored by noting that when $\theta_A = \pi$, corresponding to the whole sphere, ${\rm tr}P \sim j_{\rm max}^2$ at large $j_{\rm max}$.

Chapter 6

Bootstrapping Matrix Quantum Mechanics

This chapter is essentially the same as

- Han, Xizhi, Sean A. Hartnoll, and Jorrit Kruthoff. "Bootstrapping matrix quantum mechanics." *Physical Review Letters* 125.4 (2020): 041601.

Abstract

Large N matrix quantum mechanics is central to holographic duality but not solvable in the most interesting cases. We show that the spectrum and simple expectation values in these theories can be obtained numerically via a 'bootstrap' methodology. In this approach, operator expectation values are related by symmetries — such as time translation and $SU(N)$ gauge invariance — and then bounded with certain positivity constraints. We first demonstrate how this method efficiently solves the conventional quantum anharmonic oscillator. We then reproduce the known solution of large N single matrix quantum mechanics. Finally, we present new results on the ground state of large N two matrix quantum mechanics.

6.1 Introduction

Large N matrices are at the heart of the holographic emergence of semiclassical, gravitating spacetime geometry [164]. In matrix quantum mechanics geometry emerges from an underlying theory with no built in locality. The simplest such theory is the single matrix quantum mechanics description of two dimensional string theory [131], while the richest are the maximally supersymmetric multi-matrix theories of BFSS [17] and BMN [23]. There are many theories in between, with varying numbers of

matrices and degrees of supersymmetry [63]. Thus far, only the single matrix quantum mechanics has proved solvable at large N [39].

Nonzero temperature Monte Carlo studies of large N multi-matrix quantum mechanical systems have successfully captured aspects of a known dual spacetime in supersymmetric theories [7, 49, 74, 25]. Substantial Monte Carlo studies have also been performed for nonzero temperature bosonic multi-matrix theories, e.g. [15, 24]. However, recent work increasingly suggests that the quantum structure of holographic quantum states — revealed for instance in their entanglement [26, 73, 70, 94] — plays a central role in the emergence of space. It therefore behooves us to find methods suitable for studying the zero temperature quantum states of multi-matrix quantum mechanics directly. Progress was made recently in this direction by using a neural network variational wavefunction [88]. Here we describe a different approach.

Our work is directly inspired by a recent beautiful paper by Lin [149], with a similar approach also being employed in [9]. Lin's paper studied large N matrix integrals, which is an easier problem than large N quantum mechanics but shares important features. Positivity constraints and relations between correlation functions were shown to efficiently produce strong numerical bounds on correlation functions of matrix integrals. In the following we will show how this methodology can be adapted to the quantum mechanical problem.

6.2 Bootstrapping the quantum anharmonic oscillator

We first illustrate the approach with a warm-up example of a quantum anharmonic oscillator, with Hamiltonian

$$H = p^2 + x^2 + gx^4 \,. \tag{6.1}$$

Here $[p, x] = -i$. Fig. 6.1 below shows the results for this case: strong constraints on the energy E and expectation value $\langle x^2 \rangle$ of the ground state and first excited state.

The first step is to relate the expectation values of different operators. We will obtain the recursion relation in (6.6) below. In energy eigenstates, for any operator $\mathcal{O}$,

$$\langle [H, \mathcal{O}] \rangle = 0. \tag{6.2}$$

For example, let $\mathcal{O} = xp$. Eq. (6.2) is then the Virial theorem, $\langle 2p^2 \rangle = \langle 2x^2 + 4gx^4 \rangle$. The energy is therefore

$$E = 2\langle x^2 \rangle + 3g\langle x^4 \rangle. \tag{6.3}$$

More systematically, take $\mathcal{O} = x^s$ and $\mathcal{O} = x^t p$ in (6.2) for integers $s, t \geq 0$. Commuting the operators x, p with the identity $[p, x^r] = -irx^{r-1}$ and eliminating the terms with a single p operator,

we arrive at the relation

$$4t\langle x^{t-1}p^2\rangle = 8g\langle x^{t+3}\rangle + 4\langle x^{t+1}\rangle - t(t-1)(t-2)\langle x^{t-3}\rangle\,. \tag{6.4}$$

In this single particle case is there is a strengthened version of (6.2): $\langle \mathcal{O}H\rangle = E\langle\mathcal{O}\rangle$. We emphasize (6.2) instead because, as we will see later, it is more useful in the matrix case. Nonetheless, in the present anharmonic oscillator example, take $\mathcal{O} = x^{t-1}$, so that

$$\langle x^{t-1}p^2\rangle = E\langle x^{t-1}\rangle - \langle x^{t+1}\rangle - g\langle x^{t+3}\rangle. \tag{6.5}$$

Plugging (6.5) into (6.4) gives a recursive relation between expectation values of powers of x:

$$\begin{aligned} &4tE\langle x^{t-1}\rangle + t(t-1)(t-2)\langle x^{t-3}\rangle \\ &\qquad - 4(t+1)\langle x^{t+1}\rangle - 4g(t+2)\langle x^{t+3}\rangle = 0, \end{aligned} \tag{6.6}$$

where E is given by (6.3). Also we know that $\langle x^0\rangle = 1$ and $\langle x^t\rangle = 0$ if t is odd, so all expectation values of x^t can be computed from E and $\langle x^2\rangle$ with (6.6).

With the recursion relation (6.6) at hand we move onto the second step. We wish to solve for E and $\langle x^2\rangle$, the only two unknown variables, by bootstrapping. This step works as in [149]. The basic positivity constraint is that

$$\langle \mathcal{O}^\dagger\mathcal{O}\rangle \geq 0\,, \qquad \forall \mathcal{O} = \sum_{i=0}^{K} c_i x^i\,, \tag{6.7}$$

which means that the matrix $\mathcal{M}$ of size $(K+1)\times(K+1)$, $\mathcal{M}_{ij} = \langle x^{i+j}\rangle$, should be positive semidefinite. The constraint becomes stronger as we increase K, thus enlarging the space of trial operators. For a given K and test values of E and $\langle x^2\rangle$, the $\mathcal{M}_{ij}$ can be computed using the recursion relation (6.6). The bootstrap consists in scanning over these test values, computing the eigenvalues of the matrix $\mathcal{M}$, and thereby determining if positivity excludes the test values as inconsistent.

The result is shown in Fig. 6.1. Even for moderate K the values of E and $\langle x^2\rangle$ are determined quite accurately. The region of allowed values splits into a discrete set of islands. These converge to the spectrum of the Hamiltonian in the limit $K\to\infty$ [1]. Higher energy states require more constraints to be computed accurately.

[1] If $\langle I\rangle = 1$, $\langle\mathcal{O}^\dagger\rangle = \langle\mathcal{O}\rangle^*$ and $\langle\mathcal{O}^\dagger\mathcal{O}\rangle \geq 0$ for *all* operators $\mathcal{O}$, then $\langle\mathcal{O}\rangle = \mathrm{tr}(\rho\mathcal{O})$ for some quantum state ρ. If furthermore $\langle\mathcal{O}H\rangle = E\langle\mathcal{O}\rangle$, then ρ must be an eigenstate with energy E. Therefore as $K\to\infty$, wherein the constraints are indeed imposed for all operators, the allowed region of energies necessarily shrinks to the spectrum of the Hamiltonian, with $\langle\mathcal{O}\rangle$ the expectation value in energy eigenstates

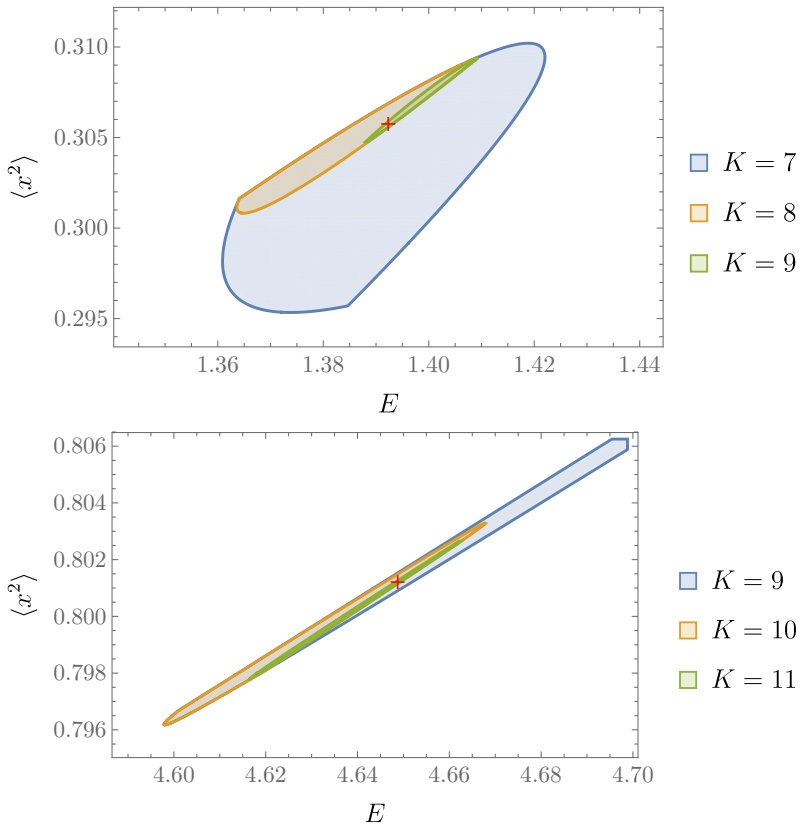


Figure 6.1: Bootstrap allowed region (shaded) for the anharmonic oscillator (6.1) with $g = 1$. Upper plot: the allowed region for $(E, \langle x^2 \rangle)$ near the ground state solution (marked by the red cross) for different sizes of the bootstrap matrix $K = 7, 8, 9$; lower plot: the allowed region near the first excited state.

6.3 One matrix quantum mechanics

Now we generalize the bootstrap method to matrix quantum mechanics at $N = \infty$. The momentum operators can no longer be eliminated explicitly in favor of the energy, and we do not use a closed form recursion relation for all expectation values. However, the energy and expectation values of short operators can still be efficiently constrained.

Consider the single-matrix quantum mechanics with

$$H = \mathrm{tr} P^2 + \mathrm{tr} X^2 + \frac{g}{N} \mathrm{tr} X^4, \tag{6.8}$$

where P and X are N-by-N Hermitian matrices with quantum commutators $[P_{ij}, X_{kl}] = -i\delta_{il}\delta_{jk}$. The theory (6.8) can be solved by mapping onto N free fermions [39]. The bootstrap reproduces this solution in Fig. 6.2.

Operator expectation values are related by symmetries. In the following, denote $\langle\mathcal{O}\rangle = \mathrm{tr}\rho\,\mathcal{O}$. If the state ρ commutes with the Hamiltonian then

$$\langle[H,\mathcal{O}]\rangle = 0, \quad \forall\mathcal{O}. \tag{6.9}$$

For example, ρ could be a pure energy eigenstate or a mixed thermal state. Choosing $\mathcal{O} = \mathrm{tr}XP$,

$$2\langle\mathrm{tr}P^2\rangle = 2\langle\mathrm{tr}X^2\rangle + \frac{4g}{N}\langle\mathrm{tr}X^4\rangle. \tag{6.10}$$

The $SU(N)$ symmetry of (6.8) has generators

$$G = i[X,P] + NI\,. \tag{6.11}$$

The final identity piece ensures that $\langle\mathrm{tr}G\rangle = 0$, with the operator ordering $[X,P] = XP - PX$ in (6.11). In gauged matrix quantum mechanics, physical states must be invariant under this symmetry. In particular,

$$\langle\mathrm{tr}G\mathcal{O}\rangle = 0, \quad \forall\mathcal{O}_{ij}. \tag{6.12}$$

For example, $\langle\mathrm{tr}G\rangle = 0$ implies $\langle\mathrm{tr}XP\rangle - \langle\mathrm{tr}PX\rangle = iN^2$. Combining this constraint with $\langle[H,\mathrm{tr}X^2]\rangle = 0$ gives

$$\langle\mathrm{tr}XP\rangle = -\langle\mathrm{tr}PX\rangle = \frac{iN^2}{2}. \tag{6.13}$$

Cyclicity of the trace gives another set of relations between operators. Commuting quantum operators may be necessary in applying the cyclic formula. For example, using large N factorization to leading order in $N \to \infty$,

$$\langle\mathrm{tr}XP^3\rangle = \langle\mathrm{tr}P^3X\rangle + 2iN\langle\mathrm{tr}P^2\rangle + i\langle\mathrm{tr}P\rangle\langle\mathrm{tr}P\rangle. \tag{6.14}$$

Equations (6.9), (6.12), cyclicity of the trace, and reality conditions $\langle\mathcal{O}^\dagger\rangle = \langle\mathcal{O}\rangle^*$ generate all relations between expectation values that we will use for the bootstrap.

As a mini-bootstrap example, consider trial operators I, X, X^2 and P. From the condition (6.7), the following bootstrap matrix should be positive semidefinite:

	I	X^2	X	P
I	$\langle\mathrm{tr}I\rangle$	$\langle\mathrm{tr}X^2\rangle$	0	0
X^2	$\langle\mathrm{tr}X^2\rangle$	$\langle\mathrm{tr}X^4\rangle$	0	0
X	0	0	$\langle\mathrm{tr}X^2\rangle$	$\langle\mathrm{tr}XP\rangle$
P	0	0	$\langle\mathrm{tr}PX\rangle$	$\langle\mathrm{tr}P^2\rangle$

(6.15)

Trial operators are built from both X and P. The expectation value for an odd number of matrices

vanishes. Positivity of (6.15) implies

$$\langle \mathrm{tr} X^2 \rangle \geq 0, \quad N\langle \mathrm{tr} X^4 \rangle \geq \langle \mathrm{tr} X^2 \rangle^2,$$
$$\langle \mathrm{tr} X^2 \rangle \left(\langle \mathrm{tr} X^2 \rangle + \frac{2g}{N} \langle \mathrm{tr} X^4 \rangle \right) \geq \frac{N^4}{4}, \tag{6.16}$$

where equations (6.10) and (6.13) are used. The inequalities (6.16) are the bootstrap constraints in this simple example. At $g = 0$, $\langle \mathrm{tr} X^2 \rangle = \frac{1}{2}N^2$ and $\langle \mathrm{tr} X^4 \rangle = \frac{1}{2}N^3$, so the last inequality in (6.16) is saturated and the other two are not.

The bootstrap constraints become stronger as we include more trial operators. Firstly, take all possible strings of X and P of length $\leq L$, and write down the matrix analogous to (6.15). This matrix must be positive semidefinite. Secondly, regard each of the $\sim 2^{2L}$ entries in the matrix as a variable (which is the expectation value of a single-trace operator with length $\leq 2L$), and write down the equalities between them following from (6.9), (6.12), cyclicity of the trace, $\langle \mathcal{O}^\dagger \rangle = \langle \mathcal{O} \rangle^*$ and that the expectation value of an odd number of matrices vanishes. The technical implementation of these constraints, as well as the minimization described in the following paragraph, is detailed in [2].

Unlike in the single-particle case, we do not necessarily require that the state be an energy eigenstate and the energy E does not appear explicitly in the bootstrap constraints. At infinite N the matrix quantum mechanics has a continuous spectrum and therefore we proceed to use gradient descent to minimize the energy in the allowed region of expectation values. In this way we obtain a lower bound on the ground state energy of the theory. The result is a lower bound because certainly the true ground state energy is allowed, and hence above the minimal allowed energy that we find. In Fig. 6.2 we observe that the lower bound is very close to the true ground state value, already for $L = 3$, and other observables, such as $\langle \mathrm{tr} X^2 \rangle$, are also solved accurately.

6.4 Two matrix quantum mechanics

One matrix quantum mechanics are tractable analytically as one can diagonalize the matrix. This is not the case for multi-matrix quantum mechanics. In the following we illustrate how bootstrap methods can successfully be used for such theories, focussing on a relatively simple two-matrix quantum mechanics with a global $O(2)$ symmetry (in addition to the large N gauge symmetry). The Hamiltonian is

$$H = \mathrm{tr}\left(P_X^2 + P_Y^2 + m^2(X^2 + Y^2) - g^2[X, Y]^2 \right), \tag{6.17}$$

with X and Y being N-by-N Hermitian matrices, with conjugate momenta P_X and P_Y, and m and g coupling constants. This theory is not exactly solvable. An early discussion of the massless ($m = 0$) limit of the theory is [106]. By rescaling the matrices we see that dimensionless physical

[2] See supplementary material below.

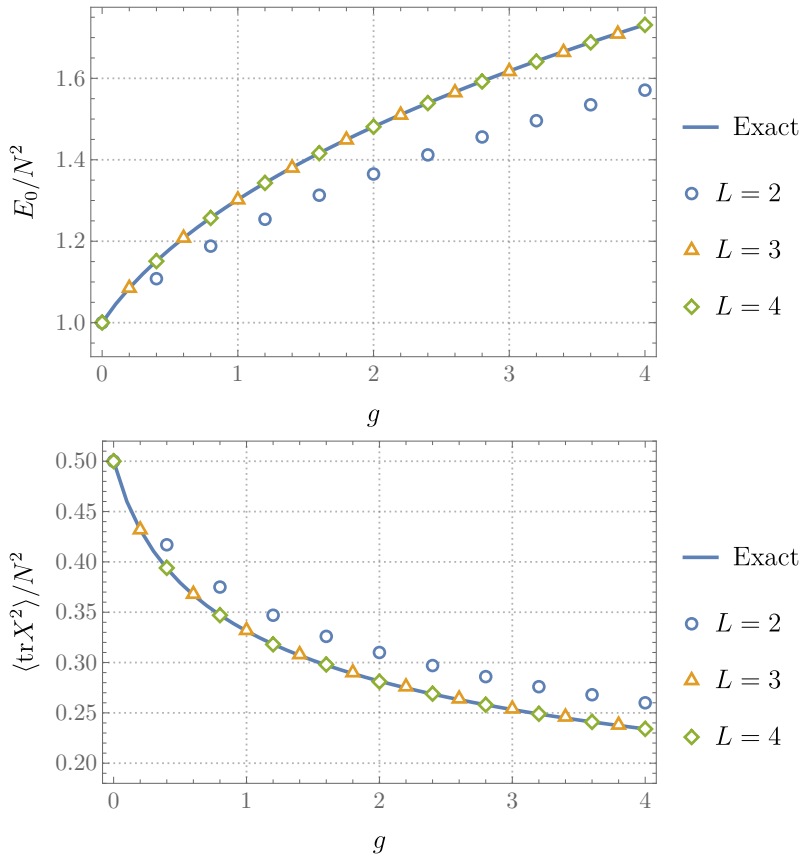


Figure 6.2: One matrix quantum mechanics bootstrap for the Hamiltonian (6.8). L is the maximal length of trial operators. Upper: The markers show the minimal energies allowed by the bootstrap constraints, in comparison with the exact ground state solution. Lower: the expectation values of $\mathrm{tr} X^2$, for the minimal energy parameters found in the upper plot.

quantities can only depend on the ratio $m^2/g^{4/3}$.

Imposing rotational invariance gives more relations between observables. We expect the ground state to be rotationally invariant. Rotations are generated by

$$S = \mathrm{tr}(XP_Y - YP_X)\,. \tag{6.18}$$

For states ρ with $[S, \rho] = 0$, including eigenstates of S,

$$\langle [S, \mathcal{O}] \rangle = 0, \quad \forall \mathcal{O}. \tag{6.19}$$

Thus in the two matrix quantum mechanics, equations (6.9), (6.12), (6.19), cyclicity of the trace, and $\langle \mathcal{O}^\dagger \rangle = \langle \mathcal{O} \rangle^*$ will be used to generate all equations between expectation values that we will use. The

bootstrap then proceeds in exactly the same way as for the case of a single matrix, now with $\sim 4^{2L}$ variables prior to imposing constraints. The results for the ground state energy, $\langle \mathrm{tr}X^2 + \mathrm{tr}Y^2 \rangle$ and $\langle \mathrm{tr}[X,Y]^2 \rangle$ are in Fig. 6.3. The Virial theorem relates these: $E_0 = 2m^2\langle \mathrm{tr}X^2 + \mathrm{tr}Y^2 \rangle - 3g^2 \langle \mathrm{tr}[X,Y]^2 \rangle$.

In order to corroborate the accuracy of the $L = 4$ results, we obtain rigorous upper and lower bounds on the true ground state energy using a Born-Oppenheimer wavefunction. We see in Fig. 6.3 that the $L = 4$ bootstrap results indeed lie within a narrow window allowed by these bounds. We briefly describe the wavefunction in the following paragraph, with details given in the appendices. As further evidence that the $L = 4$ bootstrap results are close to convergence, we compare our results to existing low temperature Monte Carlo simulations of the massless theory. At large g, $E_0/N^2 \approx 1.40\,(Ng^2)^{1/3} + 1.01\,m^2/(Ng^2)^{1/3}$ from data in Fig. 6.3. The factor of 1.40 agrees precisely with the Monte Carlo result in [174], corresponding to the value of 0.70 in the conventions of that paper. An analogous fit gives the leading order behavior $\langle \mathrm{tr}X^2 + \mathrm{tr}Y^2 \rangle / N^2 \approx 1.22/(Ng^2)^{1/3}$. The numerical factor here is close to the Monte Carlo result of 1.15 in [174].

The $SU(N)$ gauge invariance allows us to diagonalize one of the two matrices, say X. Let the eigenvalues be x_i. The Hamiltonian for the entries y_{ij} of the remaining matrix is a sum of harmonic oscillators, with frequencies $\omega_{ij}^2 = m^2 + g^2(x_i - x_j)^2$. We can therefore write down a Born-Oppenheimer wavefunction in which these oscillators are placed in their ground state:

$$\Psi(X,Y) = \psi(x_i) \prod_{i,j=1}^{N} (2\omega_{ij}/\pi)^{1/4} e^{-\frac{1}{2}\omega_{ij}|y_{ij}|^2} \,. \tag{6.20}$$

That is, the y_{ij} are treated as 'fast' compared to the eigenvalues x_i. Born-Oppenheimer wavefunctions lead to both upper and lower bounds on the ground state energy. The upper bound follows from treating the wavefunction as a variational ansatz. The lower bound is obtained by finding the ground state of the eigenvalues in an effective potential due to the zero point energy of the y_{ij} oscillators. The advantage of the form (6.20) is that computing the upper and lower bounds reduces to a solvable single-matrix large N eigenvalue problem. In Fig. 6.3 we see that the bounds following from the wavefunction (6.20) turn out to be remarkably tight.

From the results in Fig. 6.3 one can verify that the ratio $N\mathrm{tr}[X,Y]^2/(\mathrm{tr}X^2)^2$ tends to a nonzero constant at large Ng^2. This means that the matrices do not commute in this limit. This can be constrasted with the analogous two matrix integral, with no time, that does become commuting at large Ng^2 [22]. This is consistent with the fact that the two matrix integral diverges in the massless limit [135, 136], as the eigenvalues spread far apart along the classically flat directions of the potential due to commuting matrices, while the massless matrix quantum mechanics still has a discrete spectrum of normalizable states [234].

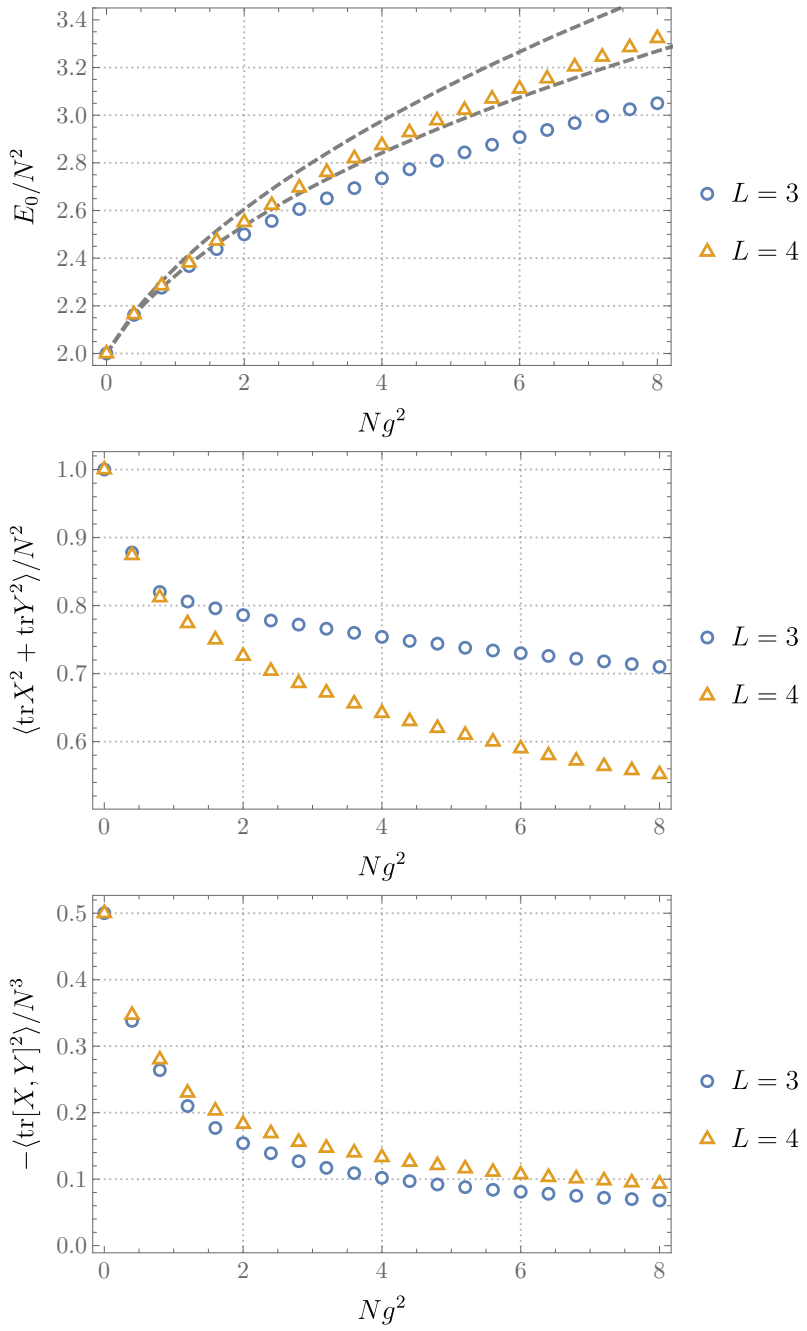


Figure 6.3: Minimal energy configuration in the bootstrap allowed region for $L = 3, 4$. The gray dashed curves are rigorous lower and upper bounds of the ground state energy from the Born-Oppenheimer approximation. In the plots we have set $m = 1$.

6.5 Final comments

In summary, we have introduced a systematic numerical method to obtain energies and expectation values of large N matrix quantum mechanics states. The method involves establishing relationships between expectation values and then imposing positivity of a certain matrix of expectation values, in the spirit of [149]. In Fig. 6.2 we see that the known analytic results for one-matrix large N quantum mechanics are readily reproduced. In Fig. 6.3 we have obtained new results for the ground state energy and expectation values of a two-matrix large N quantum mechanics.

The extension to more matrices should be possible with increased computing power or perhaps by optimizing the algorithm. Looking at supersymmetric states in supersymmetric theories may allow for stronger relationships between expectation values, using the supersymmetry generators. Both more matrices and supersymmetry will of course be necessary to tackle the full blown BFSS and BMN theories. Finally, extensions to Gibbs states (or, to high energy eigenstates) may allow nonzero temperature quantum physics to be accessed with our bootstrap methods. This could give an alternative probe of the thermal phase transitions studied via Monte Carlo in e.g. [15, 24], as well as a new window onto black hole microstates.

Acknowledgements

This work arose from discussions with Edward Mazenc and Daniel Ranard, who also collaborated on the early stages of the project. JK is supported by the Simons Foundation. SAH is partially supported by DOE award de-sc0018134 and by a Simons Investigator award.

6.6 Appendix A: Born-Oppenheimer wavefunction

This section gives details of computations involving a Born-Oppenheimer wavefunction for the two matrix quantum mechanics:

$$H = \mathrm{tr}\left(P_X^2 + P_Y^2 + m^2(X^2+Y^2) - g^2[X,Y]^2\right). \tag{6.21}$$

The role of this wavefunction is to give a lower and an upper bound on the actual ground state energy. This gives a check on the accuracy of our numerical bootstrap in this case. The results of this section are the effective Hamiltonians (6.33) and (6.34) for the eigenvalues of one of the two matrices. These will be solved in the following section 6.7, giving the upper and lower bounds respectively.

The wavefunction that we are searching for is a complex function $\Psi(X,Y)$ of Hermitian matrices X and Y. The state should be $SU(N)$ gauge invariant and hence for any unitary matrix $W \in SU(N)$,

$$\Psi(X,Y) = \Psi(WXW^{-1}, WYW^{-1}). \tag{6.22}$$

It will be convenient to parametrize such a state with the following set of variables: a diagonal real matrix x_i, a Hermitian matrix y_{ij} and a unitary matrix $U \in SU(N)$, such that

$$X = U\mathrm{diag}(x_i)U^{-1}, \quad Y = UyU^{-1}. \tag{6.23}$$

In these variables we can write down the following Born-Oppenheimer ansatz, in which the y_{ij} oscillators are put in their ground state for a fixed configuration of eigenvalues x_i:

$$\Psi(X,Y) = \psi(x_i)\phi(x_i,y_{ij}), \quad \phi(x_i,y_{ij}) = \prod_{i,j=1}^{N} (2\omega_{ij}/\pi)^{1/4} e^{-\frac{1}{2}\omega_{ij}|y_{ij}|^2}, \tag{6.24}$$

with $\omega_{ij}^2 = m^2 + g^2(x_i - x_j)^2$. Equation (6.24) defines a gauge invariant wavefunction by specifying its values on the gauge slice where X is diagonal. However, we should check that (6.24) is well-defined because (6.23) does not uniquely determine x_i and y_{ij} as a function of X and Y. Indeed, there is a residual $U(1)^{N-1}$ gauge symmetry after fixing X to be diagonal: if we choose $U = \mathrm{diag}(\exp i\theta_i)$ in (6.23), $X = \mathrm{diag}(x_i)$ but $Y_{ij} = y_{ij}\exp i(\theta_i - \theta_j)$. Because (6.24) is invariant under this residual gauge symmetry as well, $\Psi(X,Y)$ in (6.24) is well-defined.

To obtain a variational upper bound, we wish to find an effective Hamiltonian for the 'slow' x_i degrees of freedom that calculates the expectation value of the full Hamiltonian (6.21) in the variational state (6.24). The expectation value of the Hamiltonian in the state Ψ consists of a

kinetic part and a potential part:

$$\langle\Psi|H|\Psi\rangle = \int dX dY\, \Psi^*(X,Y)(H_{\rm kin}+H_{\rm pot})\Psi(X,Y)\,. \tag{6.25}$$

We discuss these in turn. The kinetic energy is

$$\langle\Psi|H_{\rm kin}|\Psi\rangle = \sum_{i,j=1}^{N}\int dX dY\, \left(\left|\frac{\partial\Psi(X,Y)}{\partial X_{ij}}\right|^2 + \left|\frac{\partial\Psi(X,Y)}{\partial Y_{ij}}\right|^2\right). \tag{6.26}$$

Here $\partial/\partial X_{ij} = \frac{1}{2}(\partial/\partial {\rm Re}X_{ij} - i\partial/\partial {\rm Im}X_{ij})$ are complex derivatives because the matrices are Hermitian. Because the kinetic energy operator is also gauge invariant, the integrand in (6.26) is constant along gauge orbits. So it suffices to evaluate it on the gauge slice where U in (6.23) is the identity. Then by the chain rule and (6.23), at $U = I$,

$$\frac{\partial\Psi}{\partial x_i} = \frac{\partial\Psi}{\partial X_{ii}}, \quad \frac{\partial\Psi}{\partial y_{ij}} = \frac{\partial\Psi}{\partial Y_{ij}}, \tag{6.27}$$

and

$$\frac{\partial\Psi}{\partial U_{ij}} = (x_j - x_i)\frac{\partial\Psi}{\partial X_{ij}} + \sum_{m,n=1}^{N}(\delta_{im}y_{jn} - \delta_{jn}y_{mi})\frac{\partial\Psi}{\partial Y_{mn}}. \tag{6.28}$$

Because Ψ is gauge invariant as in (6.22), $\partial\Psi/\partial U = 0$ so for $i \neq j$,

$$\frac{\partial\Psi}{\partial X_{ij}} = \frac{1}{x_i - x_j}\sum_{m,n=1}^{N}(\delta_{im}y_{jn} - \delta_{jn}y_{mi})\frac{\partial\Psi}{\partial y_{mn}}. \tag{6.29}$$

Plug (6.27) and (6.29) into (6.26) and evaluate the y_{ij} integrals in the state (6.24),

$$\langle\Psi|H_{\rm kin}|\Psi\rangle = \int \Delta(x_i)dx_i\, \left(\sum_{i=1}^{N}\left|\frac{\partial\psi}{\partial x_i}\right|^2 + |\psi|^2\sum_{i,j=1}^{N}\frac{\omega_{ij}}{2} + |\psi|^2\sum_{i,j,k=1}^{N}\frac{(\omega_{ik}-\omega_{jk})^2}{4\omega_{ik}\omega_{jk}(x_i-x_j)^2}\right), \tag{6.30}$$

where $\Delta = \prod_{i<j}(x_i - x_j)^2$ is the usual Vandermonde determinant, with $dXdY = \Delta dx_i dy_{ij}$.

The potential term on the gauge slice $U = I$ is

$$H_{\rm pot} = \sum_{i=1}^{N} m^2 x_i^2 + \sum_{i,j=1}^{N}\omega_{ij}^2|y_{ij}|^2, \tag{6.31}$$

and thus

$$\langle\Psi|H_{\rm pot}|\Psi\rangle = \int \Delta(x_i)dx_i\,\psi^*(x_i)\left(\sum_{i=1}^{N} m^2x_i^2 + \sum_{i,j=1}^{N}\frac{\omega_{ij}}{2}\right)\psi(x_i). \tag{6.32}$$

Overall the effective variational Hamiltonian on x_i, such that $\langle\Psi|H|\Psi\rangle = \langle\psi|H_{\rm var}|\psi\rangle$, is therefore

$$H_{\rm var} = \sum_{i=1}^{N}\left(-\frac{1}{\Delta}\frac{\partial}{\partial x_i}\left(\Delta\frac{\partial}{\partial x_i}\right) + m^2 x_i^2\right) + \sum_{i,j=1}^{N}\omega_{ij} + \sum_{i,j,k=1}^{N}\frac{(\omega_{ik}-\omega_{jk})^2}{4\omega_{ik}\omega_{jk}(x_i-x_j)^2}\,. \tag{6.33}$$

The choice of gauge and the form of the ansatz (6.24) break rotational symmetry. We have done this because it has allowed the problem to be reduced to a single-matrix eigenvalue Hamiltonian (6.33), which we will be able to solve explicitly. It is possible to restore rotational symmetry by acting on the wavefunction with the generator of rotations. This will not change the energy of the variational state.

From the variational principle we know that the ground state energy of the reduced Hamiltonian (6.33) is an upper bound on the ground state energy of the original Hamiltonian (6.21). However, it is well-known that Born-Oppenheimer wavefunctions also give a lower bound on the ground state energy. In the present context (as we prove below) this means that if we drop the final term in (6.33), the ground state energy of the Born-Oppenheimer Hamiltonian

$$H_{\rm BO} = \sum_{i=1}^{N}\left(-\frac{1}{\Delta}\frac{\partial}{\partial x_i}\left(\Delta\frac{\partial}{\partial x_i}\right) + m^2 x_i^2\right) + \sum_{i,j=1}^{N}\omega_{ij}\,, \tag{6.34}$$

is a lower bound on the ground state energy of (6.21).

A short proof of this fact is as follows: split the kinetic term into three parts $H_{\rm kin} = H^1_{\rm kin} + H^2_{\rm kin} + H^3_{\rm kin}$, where $H^1_{\rm kin}$ is the $\partial\Psi/\partial X_{ij}$ contribution in (6.26), but where the derivative does not act on the ϕ part of the wavefunction (6.24), $H^2_{\rm kin}$ is the $\partial\Psi/\partial X_{ij}$ contribution in (6.26) minus $H^1_{\rm kin}$, and $H^3_{\rm kin}$ is the remaining $\partial\Psi/\partial Y_{ij}$ term. Also split the potential term (6.31) into two pieces: $H_{\rm pot} = H^1_{\rm pot} + H^2_{\rm pot}$, where $H^1_{\rm pot}$ is the first sum in (6.31) and $H^2_{\rm pot}$ the second. Now note that $\phi(x_i, y_{ij})$ in (6.24) is the ground state of the harmonic oscillator Hamiltonian $H^3_{\rm kin} + H^2_{\rm pot}$ and that $H^2_{\rm kin}$ is positive semidefinite, so for any gauge invariant state $\Phi(x_i, y_{ij})$,

$$\begin{aligned}\langle\Phi|H|\Phi\rangle &\geq \langle\Phi|H^1_{\rm kin} + H^2_{\rm kin} + H^1_{\rm pot} + E_{\rm BO}(x_i)|\Phi\rangle \\ &\geq \langle\Phi|H^1_{\rm kin} + H^1_{\rm pot} + E_{\rm BO}(x_i)|\Phi\rangle = \langle\Phi|H_{\rm BO}|\Phi\rangle,\end{aligned} \tag{6.35}$$

where $E_{\rm BO}(x_i) = \sum_{i,j=1}^{N}\omega_{ij}$ is the ground state energy of the harmonic oscillator Hamiltonian for y_{ij}'s:

$$H^3_{\rm kin} + H^2_{\rm pot} = -\sum_{i,j=1}^{N}\frac{\partial^2}{\partial y_{ij}\partial y_{ji}} + \sum_{i,j=1}^{N}\omega_{ij}^2|y_{ij}|^2. \tag{6.36}$$

6.7 Appendix B: Large N collective field solution

In this section we solve for the ground state energies of the effective eigenvalue Hamiltonians (6.33) and (6.34), using the large N collective field method. We thereby obtain an upper and a lower bound for the ground state energy of (6.21). As is well known, at large N the collective field of eigenvalues

$$\rho(x) = \sum_{i=1}^{N} \delta(x - x_i)\,, \tag{6.37}$$

becomes classical. We can follow the established steps [58] to obtain the energy as a functional of this collective field. To obtain the Hamiltonian for $\rho(x)$ we must relate the derivative ∂_{x_i} to the conjugate collective variable $\pi(x) = -i\delta/\delta\rho(x)$. The chain rule shows that

$$\partial_{x_i} = i\pi'(x_i)\,, \qquad \partial_{x_i}^2 = i\pi''(x_i) - \pi'(x_i)^2\,. \tag{6.38}$$

Plugging these into (6.34) and defining

$$\rho_H(x) = \mathcal{P}\int dy \frac{\rho(y)}{x-y}, \tag{6.39}$$

where $\mathcal{P}$ denotes taking the principal value, one finds

$$H_{\rm BO} = \int dx \rho(x) \left[\pi'(x)^2 - 2i\rho_H(x)\pi'(x) + V(x)\right]\,, \tag{6.40}$$

with

$$V(x) = m^2x^2 + \int dy \rho(y)\sqrt{m^2 + g^2(x-y)^2}\,. \tag{6.41}$$

We also used the fact that

$$\mathcal{P}\int dxdy\rho(x)\rho(y)\frac{\pi'(x)}{x-y} = \sum_{i\neq j}\frac{\pi'(x_i)}{x_i - x_j} + \frac{1}{2}\int dx\rho(x)\pi''(x)\,. \tag{6.42}$$

The Hamiltonian in (6.40) is not manifestly Hermitian. This can be cured by performing a canonical transformation that shifts π' by $i\rho_H$, resulting in the Hamiltonian,

$$H_{\rm BO} = \int dx \rho(x) \left[\pi'(x)^2 + \rho_H(x)^2 + V(x)\right]\,. \tag{6.43}$$

With this Hamiltonian we can straightforwardly compute the ground state energy and certain observables in the ground state. At large N the eigenvalue distribution becomes classical and hence the momentum $\pi(x)$ vanishes in the ground state. Therefore it is sufficient to minimize the potential

energy functional. Using the identity

$$\int dx\rho(x)\rho_H(x)^2 = \frac{\pi^2}{3}\int dx\rho(x)^3\,, \tag{6.44}$$

(here π is the irrational number, not the conjugate momentum) this can be written as

$$E_{\rm BO}[\rho] = \int dx\rho(x)\left(\frac{\pi^2}{3}\rho(x)^2 + m^2x^2\right) + \int dxdy\rho(x)\rho(y)\omega(x,y)\,, \tag{6.45}$$

with

$$\omega(x,y) = \sqrt{m^2 + g^2(x-y)^2}\,. \tag{6.46}$$

Equation (6.45) must be minimized subject to the normalization constraint $\int dx\rho(x) = N$ and the constraint that $\rho(x)$ be pointwise non-negative. In the large N limit, this normalization combined with balancing the terms in the energy functional and taking the mass to be fixed at order one (recall that the mass can be removed by rescaling the matrices) requires the scaling

$$x \sim N^{1/2}\,, \qquad \rho \sim N^{1/2}\,, \qquad g^2 \sim \frac{1}{N}\,. \tag{6.47}$$

This is the familiar large N scaling of these quantities. In particular the 't Hooft coupling $\lambda = g^2N$ is finite in this limit.

The minimization of (6.45) is straightforward to perform numerically, by discretizing the integral. With the numerical solution at hand one can evaluate the energy $E_{\rm BO}$ of the state. These results are shown in Fig. 6.3 in the main text.

Similarly we can minimize the effective variational Hamiltonian (6.33) to obtain an upper bound on E_0. The steps are the same as above, and the functional to minimize is now

$$\begin{aligned} E_{\rm var}[\rho] = &\int dx\rho(x)\left(\frac{\pi^2}{3}\rho(x)^2 + m^2x^2\right) + \int dxdy\rho(x)\rho(y)\omega(x,y) \\ &+ \int dxdydz\rho(x)\rho(y)\rho(z)\frac{\left(\omega(x,z)-\omega(y,z)\right)^2}{4\omega(x,z)\omega(y,z)(x-y)^2}\,. \end{aligned} \tag{6.48}$$

As discussed in section 6.6, we expect that the true ground state energy E_0 is bounded above and below as

$$E_0^{\rm low} \equiv \min_\rho E_{\rm BO}[\rho] \;\le\; E_0 \;\le\; \min_\rho E_{\rm var}[\rho] \equiv E_0^{\rm high}\,. \tag{6.49}$$

We can verify explicitly that these inequalities are obeyed in perturbation theory in small g^2N. The ground state energy of the full Hamiltonian (6.21) may be evaluated using standard quantum mechanical perturbation theory directly. The functionals $E_{\rm BO}[\rho]$ and $E_{\rm var}[\rho]$ are minimized within perturbation theory by a distribution of the form $\rho(x) = \sqrt{x_\star^2 - x^2}P(x)$, with $P(x)$ a polynomial (whose degree increases order by order in perturbation theory). At large N we obtain (with $\lambda = Ng^2$

and $m = 1$)

$$\frac{E_0^{\text{low}}}{N^2} = 2 + \frac{1}{2}\lambda - \frac{7}{16}\lambda^2 + \frac{59}{64}\lambda^3 + \cdots , \tag{6.50}$$

$$\frac{E_0}{N^2} = 2 + \frac{1}{2}\lambda - \frac{11}{32}\lambda^2 + \frac{137}{256}\lambda^3 + \cdots , \tag{6.51}$$

$$\frac{E_0^{\text{high}}}{N^2} = 2 + \frac{1}{2}\lambda - \frac{1}{4}\lambda^2 + \frac{3}{64}\lambda^3 + \cdots . \tag{6.52}$$

In these expressions we see that the Born-Oppenheimer results only start to differ from the full answer at order λ^2 and that the inequalities (6.49) are obeyed. Similar perturbative expansions have previously been considered at nonzero temperature in [2]. The opposite limit of $\lambda \to \infty$ should approach the massless ($m = 0$) result. It is simple to evaluate the lower bound in this limit. With $m = 0$ and $\lambda = 1$ we find $E_0^{\text{low}}/N^2 \approx 1.308$. This is indeed lower than the Monte Carlo result of $E_0^{\text{MC}}/N^2 \approx 1.40$ for the massless theory given in [174], which we matched with the boostrap in the main text.

In Fig. 6.3 of the main text we see that for all couplings the $L = 4$ bootstrap results lie within a narrow range bounded by (6.49).

The expectation values $\langle \text{tr} X^2 \rangle$ and $\langle \text{tr}[X, Y]^2 \rangle$ in the trial wavefunction (6.24) do not provide bounds in the way that the energy does, and therefore we have not included them in Fig. 6.3. For completeness we note that these expectation values can be computed from the minimizing numerical distribution $\rho(x)$ as

$$\langle \text{tr} X^2 \rangle = \int dx \rho(x) x^2 , \tag{6.53}$$

$$\langle \text{tr}[X, Y]^2 \rangle = -\sum_{i,j=1}^{N} \langle (x_i - x_j)^2 |y_{ij}|^2 \rangle = -\int \frac{dx dx' \rho(x)\rho(x')(x - x')^2}{2\sqrt{m^2 + g^2(x - x')^2}} , \tag{6.54}$$

$$\langle \text{tr} Y^2 \rangle = \sum_{i,j=1}^{N} \langle |y_{ij}|^2 \rangle = \int \frac{dx dx' \rho(x)\rho(x')}{2\sqrt{m^2 + g^2(x - x')^2}} . \tag{6.55}$$

The wavefunction (6.24) is not rotationally symmetric and hence $\langle \text{tr} X^2 \rangle \neq \langle \text{tr} Y^2 \rangle$ in general.

6.8 Appendix C: Numerical implementation

In this section we provide more details about the bootstrap numerics. A Python implementation is available at https://github.com/hanxzh94/matrix-bootstrap. The variables under consideration are expectation values of single trace operators, with three types of constraints: linear, quadratic and semidefinite. In the following we discuss the representations of the variables and the constraints, some tricks in the implementation, and the non-convex optimization algorithm.

The variables to solve for are expectation values of single trace operators, which are represented

as strings of matrices. Denote the set of all possible matrix symbols as $\mathcal{A}$, and strings of length $\leq L$, constructed from matrices in $\mathcal{A}$, as $\mathcal{S}_L$. For example, in the single matrix case, $\mathcal{A} = \{X, P\}$, $\mathcal{S}_2 = \{\emptyset, X, P, XX, XP, PX, PP\}$, where $\emptyset$ denotes the empty string. The corresponding expectation values are $\langle \mathrm{tr} I\rangle$, $\langle \mathrm{tr} X\rangle$, $\langle \mathrm{tr} P\rangle$, ..., $\langle \mathrm{tr} PP\rangle$. Note that the matrices are non-commutative quantum operators. The expectation values v_i are then labeled by an index i, e.g., $v_0 = \langle \mathrm{tr} I\rangle = N$, $v_1 = \langle \mathrm{tr} X\rangle$, $v_2 = \langle \mathrm{tr} P\rangle$ and so on. Represented as matrices and vectors, the linear constraints can be written as $\sum_j M_{ij} v_j = 0$, the quadratic constraints $\sum_{jk} M_{ijk} v_j v_k + \sum_j N_{ij} v_j = 0$, and semidefinite constraints $\mathcal{M}_{ij} = v_{k_{ij}} \succeq 0$. In the semidefinite constraint each matrix entry $\mathcal{M}_{ij}$ is a single trace expectation value $v_{k_{ij}}$ at index k_{ij}, and k_{ij} is a function of i and j to be discussed later.

Linear equalities come from symmetry, gauge and reality constraints. Symmetry constraints take the form of $\langle [H, \mathcal{O}]\rangle = 0$, where H is the symmetry generator, and $\mathcal{O}$ is an arbitrary single trace operator in $\mathcal{S}_{2L}$. If the commutator generates operators outside $\mathcal{S}_{2L}$, the constraint is discarded. The quantum commutator of two single trace operators is also a single trace, so $\langle [H, \mathcal{O}]\rangle = 0$ is a linear equality of some single trace expectation values. Equation (6.10) in the main text is an example.

For gauge constraints $\langle \mathrm{tr} G\mathcal{O}\rangle = 0$ as in (6.12), both G and $\mathcal{O}$ are matrices instead of trace operators. In this case $\mathcal{O}$ runs over strings in $\mathcal{S}_{2L-2}$, and $\mathrm{tr} G\mathcal{O}$ is a linear combination of single trace variables. For example, in the one matrix case, G is given by (6.11). Then if we take $\mathcal{O} = XX$, the equality is

$$i\langle \mathrm{tr} XPXX\rangle - i\langle \mathrm{tr} PXXX\rangle + N\langle \mathrm{tr} XX\rangle = 0. \tag{6.56}$$

The reality constraints are $\langle \mathcal{O}^\dagger\rangle - \langle \mathcal{O}\rangle^* = 0$, for $\mathcal{O}$ a single trace operator in S_{2L}. If all matrices in $\mathcal{A}$ are Hermitian, $\mathcal{O}^\dagger$ is simply the reversed string of $\mathcal{O}$. The constraint then identifies two single trace expectation values.

Quadratic constraints result from cyclicity of the trace. Classically $\mathrm{tr} AB = \mathrm{tr} BA$, but operators in A and B may not commute quantum mechanically. For any string in $\mathcal{S}_{2L}$, we impose the equality from trying to move the first matrix in the trace to the last. Specifically, let the single trace operator be $A_{i_0 i_1} B^{(1)}_{i_1 i_2} \dots B^{(r)}_{i_r i_0}$, where $A, B^{(k)} \in \mathcal{A}$ and the repeated indices are summed over. The corresponding constraint is

$$A_{i_0 i_1} B^{(1)}_{i_1 i_2} \dots B^{(r)}_{i_r i_0} - B^{(1)}_{i_1 i_2} \dots B^{(r)}_{i_r i_0} A_{i_0 i_1} = \sum_{k=1}^{r} B^{(1)}_{i_1 i_2} \dots [A_{i_0 i_1}, B^{(k)}_{i_k i_{k+1}}] \dots B^{(r)}_{i_r i_0}, \tag{6.57}$$

where the bracket is the quantum commutator. Assume that commutators of single matrices are $[A_{ij}, B_{kl}] = c_{AB} \delta_{il} \delta_{jk}$ for some constant c_{AB}. The right hand side of (6.57) is then a sum of double trace operators

$$\sum_{k=1}^{r} c_{AB^{(k)}} \mathrm{tr} B^{(1)} \dots B^{(k-1)} \mathrm{tr} B^{(k+1)} \dots B^{(r)}. \tag{6.58}$$

An explicit example is given in equation (6.14) of the main text. At large N the expectation values of double trace operators factorize, so the left side of (6.57) is linear in expectation values v_i and the other side is quadratic. These equalities are the quadratic relations $\sum_{jk} M_{ijk} v_j v_k + \sum_j N_{ij} v_j = 0$ mentioned previously.

As discussed in the main text, positivity of certain operator expectation values requires that the matrix $\mathcal{M}_{ij} = \langle \mathrm{tr} \mathcal{O}_i^\dagger \mathcal{O}_j \rangle$ be positive semidefinite. Here $\mathcal{O}_i$ and $\mathcal{O}_j$ run over strings in $\mathcal{S}_L$, so that $\langle \mathrm{tr} \mathcal{O}_i^\dagger \mathcal{O}_j \rangle$ is an expectation value $v_{k_{ij}}$ in $\mathcal{S}_{2L}$. The index k_{ij}, as a function of i and j, is determined by the fact that the string $\mathcal{O}_{k_{ij}}$ is the string $\mathcal{O}_i^\dagger \mathcal{O}_j$. In terms of the variables v_i, the positivity constraint is then that the matrix $\mathcal{M}_{ij} = v_{k_{ij}}$ should be positive semidefinite.

Before delivering the variables and constraints to optimization, we discuss several implementation tricks used to simplify coding or improve computational efficiency. Firstly, all expectation values are scaled by proper factors of N so that N is not explicit in the numerics. The N scaling can be determined from free theories and is $N^{l/2+1}$ for a single trace operator with l matrices.

Secondly, some expectation values must vanish due to symmetries and hence are not included in the constraints. For one matrix quantum mechanics (6.8) expectation values of an odd number of matrices must vanish. For two matrix quantum mechanics (6.17) it is more efficient to work with the following matrix basis $\mathcal{A} = \{A, B, C, D\}$:

$$\begin{aligned} A &= P - iX - i(Q - iY), \quad & B &= P + iX + i(Q + iY), \\ C &= P - iX + i(Q - iY), \quad & D &= P + iX - i(Q + iY). \end{aligned} \tag{6.59}$$

The four matrices are eigenvectors of the $SO(2) \cong U(1)$ action with eigenvalues $-1, 1, 1, -1$. Hence $SO(2)$ rotation invariance is imposed if we only consider strings with $n(A) - n(B) - n(C) + n(D) = 0$, where, for example, $n(A)$ is the number of A's in the string. The number of possible strings is thus significantly reduced.

Thirdly, for bosonic matrix models the wavefunction can be chosen as real, and hence expectation values of strings with an odd number of P's (and an arbitrary number of X's) must be purely imaginary, while strings with an even number of P's must be real. This fact simplifies the reality constraints and reduces the number of real variables to optimize over.

Lastly, the linear constraints $\sum_j M_{ij} v_j = 0$ can be solved to obtain a linearly independent set of variables $\widetilde{v}_i$, where $v_i = \sum_j K_{ij} \widetilde{v}_j$ and $\sum_j M_{ij} K_{jk} = 0$. Then the quadratic and semidefinite constraints are rewritten in terms of $\widetilde{v}_i$. The optimization is more efficient on this reduced set of variables.

In the optimization, the energy $\langle H \rangle$ is minimized subject to the constraints $\sum_j M_{ij} v_j = 0$, $\sum_{jk} M_{ijk} v_j v_k + \sum_j N_{ij} v_j = 0$ and $\mathcal{M}_{ij} \succeq 0$. The constraints are generally non-convex due to the presence of quadratic equalities. We employ a trust-region sequential semidefinite programming algorithm for the non-convex optimization [184]. The algorithm iteratively searches for a local minimum of the goal function, and the basic idea is as follows. At each step, the quadratic constraint

is approximated by its local linearization. With only linear and semidefinite constraints, the problem is convex and solved with semidefinite programming. The variables v_i (or $\widetilde{v}_i$) are then updated with the solution of this local convex approximation, and the algorithm proceeds to the next step. Optimization finishes when the updates are smaller than some threshold. Expectation values of the energy and other trace operators at the local minimum are returned.

Bibliography

[1] Yakir Aharonov, Sandu Popescu, and Jeff Tollaksen. Each instant of time a new universe. In *Quantum theory: a two-time success story*, pages 21–36. Springer, 2014.

[2] Ofer Aharony, Joseph Marsano, Shiraz Minwalla, Kyriakos Papadodimas, Mark Van Raamsdonk, and Toby Wiseman. The Phase structure of low dimensional large N gauge theories on Tori. *JHEP*, 01:140, 2006.

[3] Anton Yu. Alekseev, Andreas Recknagel, and Volker Schomerus. Brane dynamics in background fluxes and noncommutative geometry. *JHEP*, 05:010, 2000.

[4] John-Mark A Allen, Jonathan Barrett, Dominic C Horsman, Ciarán M Lee, and Robert W Spekkens. Quantum common causes and quantum causal models. *Physical Review X*, 7(3):031021, 2017.

[5] Ahmed Almheiri, Xi Dong, and Daniel Harlow. Bulk locality and quantum error correction in ads/cft. *Journal of High Energy Physics*, 2015(4):163, 2015.

[6] Konstantinos N. Anagnostopoulos, Takehiro Azuma, Keiichi Nagao, and Jun Nishimura. Impact of supersymmetry on the nonperturbative dynamics of fuzzy spheres. *JHEP*, 09:046, 2005.

[7] Konstantinos N. Anagnostopoulos, Masanori Hanada, Jun Nishimura, and Shingo Takeuchi. Monte Carlo studies of supersymmetric matrix quantum mechanics with sixteen supercharges at finite temperature. *Phys. Rev. Lett.*, 100:021601, 2008.

[8] P. W. Anderson. Limits on the energy of the antiferromagnetic ground state. *Phys. Rev.*, 83:1260–1260, Sep 1951.

[9] Peter D. Anderson and Martin Kruczenski. Loop equations and bootstrap methods in the lattice. *Nuclear Physics B*, 921:702 – 726, 2017.

[10] Tarek Anous and Cameron Cogburn. Mini-BFSS in Silico. 2017.

[11] Huzihiro Araki. Gibbs states of a one dimensional quantum lattice. *Communications in Mathematical Physics*, 14(2):120–157, Jun 1969.

[12] Curtis T. Asplund, Frederik Denef, and Eric Dzienkowski. Massive quiver matrix models for massive charged particles in AdS. *JHEP*, 01:055, 2016.

[13] David D Awschalom and Michael E Flatté. Challenges for semiconductor spintronics. *Nature physics*, 3(3):153, 2007.

[14] Takehiro Azuma, Subrata Bal, Keiichi Nagao, and Jun Nishimura. Nonperturbative studies of fuzzy spheres in a matrix model with the Chern-Simons term. *JHEP*, 05:005, 2004.

[15] Takehiro Azuma, Takeshi Morita, and Shingo Takeuchi. Hagedorn Instability in Dimensionally Reduced Large-N Gauge Theories as Gregory-Laflamme and Rayleigh-Plateau Instabilities. *Phys. Rev. Lett.*, 113:091603, 2014.

[16] Waseem S. Bakr, Jonathon I. Gillen, Amy Peng, Simon Fölling, and Markus Greiner. A quantum gas microscope for detecting single atoms in a Hubbard-regime optical lattice. *Nature*, 462:74, 11 2009.

[17] Tom Banks, W. Fischler, S.H. Shenker, and Leonard Susskind. M theory as a matrix model: A Conjecture. *Phys. Rev. D*, 55:5112–5128, 1997.

[18] Tillmann Baumgratz and Martin B Plenio. Lower bounds for ground states of condensed matter systems. *New Journal of Physics*, 14(2):023027, feb 2012.

[19] Charles H Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K Wootters. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Physical review letters*, 70(13):1895, 1993.

[20] Cédric Bény, Achim Kempf, and David W Kribs. Generalization of quantum error correction via the heisenberg picture. *Physical review letters*, 98(10):100502, 2007.

[21] Cédric Bény, Achim Kempf, and David W Kribs. Quantum error correction of observables. *Physical Review A*, 76(4):042303, 2007.

[22] David E. Berenstein, Masanori Hanada, and Sean A. Hartnoll. Multi-matrix models and emergent geometry. *JHEP*, 02:010, 2009.

[23] David Eliecer Berenstein, Juan Martin Maldacena, and Horatiu Stefan Nastase. Strings in flat space and pp waves from N=4 superYang-Mills. *JHEP*, 04:013, 2002.

[24] Georg Bergner, Norbert Bodendorfer, Masanori Hanada, Enrico Rinaldi, Andreas Schäfer, and Pavlos Vranas. Thermal phase transition in Yang-Mills matrix model. *JHEP*, 01:053, 2020.

[25] Evan Berkowitz, Enrico Rinaldi, Masanori Hanada, Goro Ishiki, Shinji Shimasaki, and Pavlos Vranas. Precision lattice test of the gauge/gravity duality at large-N. *Phys. Rev.*, D94(9):094501, 2016.

[26] Eugenio Bianchi and Robert C. Myers. On the Architecture of Spacetime Geometry. *Class. Quant. Grav.*, 31:214002, 2014.

[27] Thomas Bilitewski, Subhro Bhattacharjee, and Roderich Moessner. Temperature dependence of butterfly effect in a classical many-body system. August 2018.

[28] Mike Blake. Universal Charge Diffusion and the Butterfly Effect in Holographic Theories. *Phys. Rev. Lett.*, 117(9):091601, 2016.

[29] Immanuel Bloch, Jean Dalibard, and Sylvain Nascimbène. Quantum simulations with ultracold quantum gases. *Nature Physics*, 8:267, 04 2012.

[30] Lapo Bogani and Wolfgang Wernsdorfer. Molecular spintronics using single-molecule magnets. *Nature materials*, 7(3):179, 2008.

[31] A Bohrdt, C B Mendl, M Endres, and M Knap. Scrambling and thermalization in a diffusive quantum many-body system. *New Journal of Physics*, 19(6):063001, 2017.

[32] A. Bohrdt, C. B. Mendl, M. Endres, and M. Knap. Scrambling and thermalization in a diffusive quantum many-body system. *New J. Phys.*, 19(6):063001, 2017.

[33] Luca Bombelli, Rabinder K. Koul, Joohan Lee, and Rafael D. Sorkin. A Quantum Source of Entropy for Black Holes. *Phys. Rev.*, D34:373–383, 1986.

[34] Gabriel Bouch. Complex-time singularity and locality estimates for quantum lattice systems. *Journal of Mathematical Physics*, 56(12):123303, 2015.

[35] Raphael Bousso and Douglas Stanford. Measurements without probabilities in the final state proposal. *Physical Review D*, 89(4):044038, 2014.

[36] Peter Braun-Munzinger and Johanna Stachel. The quest for the quark–gluon plasma. *Nature*, 448:302, 07 2007.

[37] S. Bravyi, M. B. Hastings, and F. Verstraete. Lieb-Robinson Bounds and the Generation of Correlations and Topological Quantum Order. *Phys. Rev. Lett.*, 97:050401, 2006.

[38] Heinz-Peter Breuer and Francesco Petruccione. *The Theory of Open Quantum Systems*. Oxford University Press, 2007.

[39] E. Brezin, C. Itzykson, G. Parisi, and J.B. Zuber. Planar Diagrams. *Commun. Math. Phys.*, 59:35, 1978.

[40] Winton Brown and Omar Fawzi. Scrambling speed of random quantum circuits. *arXiv preprint arXiv:1210.6644*, 2012.

[41] J. A. N. Bruin, H. Sakai, R. S. Perry, and A. P. Mackenzie. Similarity of Scattering Rates in Metals Showing T-Linear Resistivity. *Science*, 339(6121):804–807, 2013.

[42] Časlav Brukner. Quantum causality. *Nature Physics*, 10(4):259, 2014.

[43] Časlav Brukner. Bounding quantum correlations with indefinite causal order. *New Journal of Physics*, 17(8):083034, 2015.

[44] Giuseppe Carleo and Matthias Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602–606, 2017.

[45] Juan Carrasquilla, Giacomo Torlai, Roger G. Melko, and Leandro Aolita. Reconstructing quantum states with generative models. *Nature Machine Intelligence*, 1(3):155–161, 2019.

[46] Horacio Casini, Marina Huerta, and Jose Alejandro Rosabal. Remarks on entanglement entropy for gauge fields. *Phys. Rev.*, D89(8):085012, 2014.

[47] Esteban Castro-Ruiz, Flaminia Giacomini, and Časlav Brukner. Dynamics of quantum causal structures. *Physical Review X*, 8(1):011047, 2018.

[48] P. Castro-Villarreal, R. Delgadillo-Blando, and Badis Ydri. A Gauge-invariant UV-IR mixing and the corresponding phase transition for U(1) fields on the fuzzy sphere. *Nucl. Phys.*, B704:111–153, 2005.

[49] Simon Catterall and Toby Wiseman. Black hole thermodynamics from simulations of lattice Yang-Mills theory. *Phys. Rev.*, D78:041502, 2008.

[50] P. M. Chaikin and T. C. Lubensky. *Principles of Condensed Matter Physics*. CUP, Cambridge, 1995.

[51] Hong Zhe Chen and Joanna L. Karczmarek. Entanglement entropy on a fuzzy sphere with a UV cutoff. *JHEP*, 08:154, 2018.

[52] Chih-Chun Chien, Sebastiano Peotta, and Massimiliano Di Ventra. Quantum transport in ultracold atoms. *Nature Physics*, 11:998, Dec 2015.

[53] Richard Cleve, Daniel Gottesman, and Hoi-Kwong Lo. How to share a quantum secret. *Physical Review Letters*, 83(3):648, 1999.

[54] Fabio Costa and Sally Shrapnel. Quantum causal modelling. *New Journal of Physics*, 18(6):063032, 2016.

[55] Jordan Cotler, Nicholas Hunter-Jones, Junyu Liu, and Beni Yoshida. Chaos, complexity, and random matrices. *Journal of High Energy Physics*, 2017(11):48, 2017.

[56] Jordan Cotler, Chao-Ming Jian, Xiao-Liang Qi, and Frank Wilczek. Superdensity operators for spacetime quantum mechanics. *arXiv:1711.03119*, 2017.

[57] Sumit R. Das. Degrees of freedom in two-dimensional string theory. *Nucl. Phys. Proc. Suppl.*, 45BC:224–233, 1996.

[58] Sumit R. Das and Antal Jevicki. String Field Theory and Physical Interpretation of $D = 1$ Strings. *Mod. Phys. Lett. A*, 5:1639–1650, 1990.

[59] Sankar Das Sarma, Dong-Ling Deng, and Lu-Ming Duan. Machine learning meets quantum physics. *Physics Today*, 72(3):48–54, 2019.

[60] Keshav Dasgupta, Mohammad M. Sheikh-Jabbari, and Mark Van Raamsdonk. Matrix perturbation theory for M theory on a PP wave. *JHEP*, 05:056, 2002.

[61] Richard A. Davison, Wenbo Fu, Antoine Georges, Yingfei Gu, Kristan Jensen, and Subir Sachdev. Thermoelectric transport in disordered metals without quasiparticles: The Sachdev-Ye-Kitaev models and holography. *Phys. Rev. B*, 95:155131, Apr 2017.

[62] B. de Wit, J. Hoppe, and H. Nicolai. On the quantum mechanics of supermembranes. *Nuclear Physics B*, 305(4):545 – 581, 1988.

[63] Bernard de Wit. Supersymmetric quantum mechanics, supermembranes and Dirichlet particles. *Nucl. Phys. Proc. Suppl.*, 56B:76–87, 1997.

[64] Rodrigo Delgadillo-Blando, Denjoe O'Connor, and Badis Ydri. Geometry in Transition: A Model of Emergent Geometry. *Phys. Rev. Lett.*, 100:201601, 2008.

[65] Dong-Ling Deng, Xiaopeng Li, and S. Das Sarma. Quantum entanglement in neural network states. *Phys. Rev. X*, 7:021021, May 2017.

[66] Laurent Dinh, David Krueger, and Yoshua Bengio. NICE: Non-linear Independent Components Estimation. page arXiv:1410.8516, Oct 2014.

[67] Laurent Dinh, Jascha Sohl-Dickstein, and Samy Bengio. Density estimation using real NVP. *CoRR*, 2016.

[68] William Donnelly. Decomposition of entanglement entropy in lattice gauge theory. *Phys. Rev.*, D85:085004, 2012.

[69] William Donnelly. Entanglement entropy and nonabelian gauge symmetry. *Class. Quant. Grav.*, 31(21):214003, 2014.

[70] William Donnelly and Laurent Freidel. Local subsystems in gauge theory and gravity. *JHEP*, 09:102, 2016.

[71] Djamel Dou and Badis Ydri. Entanglement entropy on fuzzy spaces. *Phys. Rev.*, D74:044014, 2006.

[72] David Fattal, Toby S Cubitt, Yoshihisa Yamamoto, Sergey Bravyi, and Isaac L Chuang. Entanglement in the stabilizer formalism. *arXiv preprint quant-ph/0406168*, 2004.

[73] Thomas Faulkner, Aitor Lewkowycz, and Juan Maldacena. Quantum corrections to holographic entanglement entropy. *JHEP*, 11:074, 2013.

[74] Veselin G. Filev and Denjoe O'Connor. The BFSS model on the lattice. *JHEP*, 05:167, 2016.

[75] Thomas M. Fiola, John Preskill, Andrew Strominger, and Sandip P. Trivedi. Black hole thermodynamics and information loss in two-dimensions. *Phys. Rev.*, D50:3987–4014, 1994.

[76] Joseph F Fitzsimons, Jonathan A Jones, and Vlatko Vedral. Quantum correlations which imply causation. *Scientific reports*, 5:18281, 2015.

[77] D. Forster. *Hydrodynamic Fluctuations, Broken Symmetry and Correlation Functions*. Perseus Books, 1975.

[78] Xun Gao and Lu-Ming Duan. Efficient representation of quantum many-body states with deep neural networks. *Nature Communications*, 8(1):662, 2017.

[79] Mathieu Germain, Karol Gregor, Iain Murray, and Hugo Larochelle. {MADE:} Masked Autoencoder for Distribution Estimation. *CoRR*, 2015.

[80] Ivan Glasser, Nicola Pancotti, Moritz August, Ivan D. Rodriguez, and J. Ignacio Cirac. Neural-network quantum states, string-bond states, and chiral topological states. *Phys. Rev. X*, 8:011006, Jan 2018.

[81] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

[82] Sarang Gopalakrishnan, David A. Huse, Vedika Khemani, and Romain Vasseur. Hydrodynamics of operator spreading and quasiparticle diffusion in interacting integrable systems. 2018.

[83] Daniel Gottesman. Stabilizer codes and quantum error correction. *arXiv preprint quant-ph/9705052*, 1997.

[84] Daniel Gottesman and John Preskill. Comment on"the black hole final state". *Journal of High Energy Physics*, 2004(03):026, 2004.

[85] Yingfei Gu, Xiao-Liang Qi, and Douglas Stanford. Local criticality, diffusion and chaos in generalized Sachdev-Ye-Kitaev models. *JHEP*, 05:125, 2017.

[86] O. Gunnarsson, M. Calandra, and J. E. Han. Colloquium: Saturation of electrical resistivity. *Rev. Mod. Phys.*, 75:1085–1099, Oct 2003.

[87] Wei Han, Roland K Kawakami, Martin Gmitra, and Jaroslav Fabian. Graphene spintronics. *Nature nanotechnology*, 9(10):794, 2014.

[88] Xizhi Han and Sean A. Hartnoll. Deep Quantum Geometry of Matrices. *Phys. Rev. X*, 10(1):011069, 2020.

[89] Xizhi Han, Sean A. Hartnoll, and Jorrit Kruthoff. Bootstrapping matrix quantum mechanics. *Phys. Rev. Lett.*, 125(4):041601, 2020.

[90] Zhao-Yu Han, Jun Wang, Heng Fan, Lei Wang, and Pan Zhang. Unsupervised generative modeling using matrix product states. *Phys. Rev. X*, 8:031012, Jul 2018.

[91] Lucien Hardy. The operator tensor formulation of quantum theory. *Phil. Trans. R. Soc. A*, 370(1971):3385–3417, 2012.

[92] Lucien Hardy. Operational general relativity: Possibilistic, probabilistic, and quantum. *arXiv preprint arXiv:1608.06940*, 2016.

[93] Daniel Harlow. Jerusalem lectures on black holes and quantum information. *Reviews of Modern Physics*, 88(1):015002, 2016.

[94] Daniel Harlow. The Ryu-Takayanagi Formula from Quantum Error Correction. *Commun. Math. Phys.*, 354(3):865–912, 2017.

[95] Aram W Harrow and Richard A Low. Random quantum circuits are approximate 2-designs. *Communications in Mathematical Physics*, 291(1):257–302, 2009.

[96] Thomas Hartman, Sean A. Hartnoll, and Raghu Mahajan. Upper Bound on Diffusivity. *Phys. Rev. Lett.*, 119(14):141601, 2017.

[97] S. A. Hartnoll, A. Lucas, and S. Sachdev. *Holographic Quantum Matter*. MIT Press, 2018.

[98] Sean A Hartnoll. Theory of universal incoherent metallic transport. *Nature Physics*, 11(1):54, 2015.

[99] Sean A. Hartnoll and Edward Mazenc. Entanglement entropy in two dimensional string theory. *Phys. Rev. Lett.*, 115(12):121602, 2015.

[100] Koji Hashimoto, Sotaro Sugishita, Akinori Tanaka, and Akio Tomiya. Deep learning and the AdS/CFT correspondence. *Phys. Rev. D*, 98:046019, Aug 2018.

[101] Matthew B. Hastings and Tohru Koma. Spectral gap and exponential decay of correlations. *Communications in Mathematical Physics*, 265(3):781–804, Aug 2006.

[102] Patrick Hayden, Sepehr Nezami, Xiao-Liang Qi, Nathaniel Thomas, Michael Walter, and Zhao Yang. Holographic duality from random tensor networks. *Journal of High Energy Physics*, 2016(11):9, 2016.

[103] Patrick Hayden and John Preskill. Black holes as mirrors: Quantum information in random subsystems. *JHEP*, 09:120, 2007.

[104] Sebastian Hild, Takeshi Fukuhara, Peter Schauß, Johannes Zeiher, Michael Knap, Eugene Demler, Immanuel Bloch, and Christian Gross. Far-from-Equilibrium Spin Transport in Heisenberg Quantum Magnets. *Phys. Rev. Lett.*, 113:147205, Oct 2014.

[105] G. E. Hinton and R. R. Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507, 2006.

[106] Jens Hoppe. *Quantum theory of a massless relativistic surface and a two-dimensional bound state problem.* PhD thesis, Massachusetts Institute of Technology, 1982.

[107] Jens Hoppe. Diffeomorphism Groups, Quantization and SU(infinity). *Int. J. Mod. Phys.*, A4:5235, 1989.

[108] Gary T Horowitz and Juan Maldacena. The black hole final state. *Journal of High Energy Physics*, 2004(02):008, 2004.

[109] Pavan Hosur, Xiao-Liang Qi, Daniel A. Roberts, and Beni Yoshida. Chaos in quantum channels. *JHEP*, 02:004, 2016.

[110] Hong-Ye Hu, Shuo-Hui Li, Lei Wang, and Yi-Zhuang You. Machine Learning Holographic Mapping by Neural Network Renormalization Group. 2019.

[111] Edwin W. Huang, Christian B. Mendl, Shenxiu Liu, Steve Johnston, Hong-Chen Jiang, Brian Moritz, and Thomas P. Devereaux. Numerical evidence of fluctuating stripes in the normal state of high-Tc cuprate superconductors. *Science*, 358(6367):1161–1164, 2017.

[112] J. Hubbard and Brian Hilton Flowers. Electron correlations in narrow energy bands. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 276(1365):238–257, 1963.

[113] N. E. Hussey, K. Takenaka, and H. Takagi. Universality of the Mott–Ioffe–Regel limit in metals. *Philosophical Magazine*, 84(27):2847–2864, 2004.

[114] Marko Žnidarič. Dephasing-induced diffusive transport in the anisotropic Heisenberg model. *New Journal of Physics*, 12(4):043001, 2010.

[115] Marko Žnidarič. Spin Transport in a One-Dimensional Anisotropic Heisenberg Model. *Phys. Rev. Lett.*, 106:220601, May 2011.

[116] Marko Žnidarič. Relaxation times of dissipative many-body quantum systems. *Phys. Rev. E*, 92:042143, Oct 2015.

[117] Igor Žutić, Jaroslav Fabian, and S. Das Sarma. Spintronics: Fundamentals and applications. *Rev. Mod. Phys.*, 76:323–410, Apr 2004.

[118] Satoshi Iso, Yusuke Kimura, Kanji Tanaka, and Kazunori Wakatsuki. Noncommutative gauge theory on fuzzy sphere from matrix model. *Nucl. Phys.*, B604:121–147, 2001.

[119] Nissan Itzhaki, Juan Martin Maldacena, Jacob Sonnenschein, and Shimon Yankielowicz. Supergravity and the large N limit of theories with sixteen supercharges. *Phys. Rev.*, D58:046004, 1998.

[120] D. Jaksch and P. Zoller. The cold atom Hubbard toolbox. *Annals of Physics*, 315(1):52 – 79, 2005. Special Issue.

[121] Dileep P. Jatkar, Gautam Mandal, Spenta R. Wadia, and K. P. Yogendran. Matrix dynamics of fuzzy spheres. *JHEP*, 01:039, 2002.

[122] Ding Jia et al. Generalizing entanglement. *Physical Review A*, 96(6):062132, 2017.

[123] Ding Jia et al. Quantum theories from principles without assuming a definite causal structure. *Physical Review A*, 98(3):032112, 2018.

[124] Ding Jia, Nitica Sakharwade, et al. Tensor products of process matrices with indefinite causal structure. *Physical Review A*, 97(3):032110, 2018.

[125] Danilo Jimenez Rezende and Shakir Mohamed. Variational Inference with Normalizing Flows. page arXiv:1505.05770, May 2015.

[126] Joanna L. Karczmarek and Philippe Sabella-Garnier. Entanglement entropy on the fuzzy sphere. *JHEP*, 03:129, 2014.

[127] C. Karrasch, J. H. Bardarson, and J. E. Moore. Finite-Temperature Dynamical Density Matrix Renormalization Group and the Drude Weight of Spin-1/2 Chains. *Phys. Rev. Lett.*, 108:227206, May 2012.

[128] Vedika Khemani, David A. Huse, and Adam Nahum. Velocity-dependent Lyapunov exponents in many-body quantum, semiclassical, and classical chaos. *Phys. Rev. B*, 98:144304, Oct 2018.

[129] Vedika Khemani, Ashvin Vishwanath, and David A. Huse. Operator spreading and the emergence of dissipative hydrodynamics under unitary evolution with conservation laws. *Phys. Rev. X*, 8:031057, Sep 2018.

[130] Diederik P Kingma, Tim Salimans, and Max Welling. Improving Variational Inference with Inverse Autoregressive Flow. *CoRR*, 2016.

[131] Igor R. Klebanov. String theory in two-dimensions. In *Spring School on String Theory and Quantum Gravity (to be followed by Workshop) Trieste, Italy, April 15-23, 1991*, pages 30–101, 1991.

[132] M. Kliesch, C. Gogolin, M. J. Kastoryano, A. Riera, and J. Eisert. Locality of temperature. *Phys. Rev. X*, 4:031019, Jul 2014.

[133] Gerald Knizia and Garnet Kin-Lic Chan. Density matrix embedding: A simple alternative to dynamical mean-field theory. *Phys. Rev. Lett.*, 109:186404, Nov 2012.

[134] Marco Koschorreck, Daniel Pertot, Enrico Vogt, and Michael Köhl. Universal spin dynamics in two-dimensional Fermi gases. *Nature Physics*, 9:405, May 2013.

[135] Werner Krauth, Hermann Nicolai, and Matthias Staudacher. Monte Carlo approach to M theory. *Phys. Lett.*, B431:31–41, 1998.

[136] Werner Krauth and Matthias Staudacher. Finite Yang-Mills integrals. *Phys. Lett.*, B435:350–355, 1998.

[137] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'12, pages 1097–1105, USA, 2012. Curran Associates Inc.

[138] Nima Lashkari, Douglas Stanford, Matthew Hastings, Tobias Osborne, and Patrick Hayden. Towards the fast scrambling conjecture. *Journal of High Energy Physics*, 2013(4):22, 2013.

[139] J. P. F. LeBlanc, Andrey E. Antipov, Federico Becca, Ireneusz W. Bulik, Garnet Kin-Lic Chan, Chia-Min Chung, Youjin Deng, Michel Ferrero, Thomas M. Henderson, Carlos A. Jiménez-Hoyos, E. Kozik, Xuan-Wen Liu, Andrew J. Millis, N. V. Prokof'ev, Mingpu Qin, Gustavo E. Scuseria, Hao Shi, B. V. Svistunov, Luca F. Tocchio, I. S. Tupitsyn, Steven R. White, Shiwei Zhang, Bo-Xiao Zheng, Zhenyue Zhu, and Emanuel Gull. Solutions of the two-dimensional hubbard model: Benchmarks and results from a wide range of numerical algorithms. *Phys. Rev. X*, 5:041041, Dec 2015.

[140] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521:436, 05 2015.

[141] Patrick A. Lee, Naoto Nagaosa, and Xiao-Gang Wen. Doping a Mott insulator: Physics of high-temperature superconductivity. *Rev. Mod. Phys.*, 78:17–85, Jan 2006.

[142] Eyal Leviatan, Frank Pollmann, Jens H. Bardarson, and Ehud Altman. Quantum thermalization dynamics with Matrix-Product States. 2017.

[143] Eyal Leviatan, Frank Pollmann, Jens H. Bardarson, David A. Huse, and Ehud Altman. Quantum thermalization dynamics with Matrix-Product States. 2017.

[144] Michael Levin and Cody P Nave. Tensor renormalization group approach to two-dimensional classical lattice models. *Physical review letters*, 99(12):120601, 2007.

[145] Elliott H. Lieb and Derek W. Robinson. The finite group velocity of quantum spin systems. *Comm. Math. Phys.*, 28(3):251–257, 1972.

[146] Elliott H. Lieb and Derek W. Robinson. The finite group velocity of quantum spin systems. *Communications in Mathematical Physics*, 28(3):251–257, Sep 1972.

[147] Elliott H. Lieb and F. Y. Wu. Absence of mott transition in an exact solution of the short-range, one-band model in one dimension. *Phys. Rev. Lett.*, 20:1445–1448, Jun 1968.

[148] Cheng-Ju Lin and Olexei I. Motrunich. Out-of-time-ordered correlators in a quantum ising chain. *Phys. Rev. B*, 97:144304, Apr 2018.

[149] Henry W. Lin. Bootstraps to strings: solving random matrix models with positivite. *JHEP*, 06:090, 2020.

[150] Jennifer Lin and ĐRadičević. Comments on Defining Entanglement Entropy. 2018.

[151] G. Lindblad. On the generators of quantum dynamical semigroups. *Comm. Math. Phys.*, 48(2):119–130, 1976.

[152] Yi Ling, Peng Liu, and Jian-Pin Wu. Holographic Butterfly Effect at Quantum Critical Points. *JHEP*, 10:025, 2017.

[153] F. Lizzi, R. J. Szabo, and A. Zampini. Geometry of the gauge algebra in noncommutative Yang-Mills theory. *JHEP*, 08:032, 2001.

[154] Marko Ljubotina, Marko Žnidarič, and Tomaž Prosen. Spin diffusion from an inhomogeneous quench in an integrable system. *Nature communications*, 8:16117, 2017.

[155] Seth Lloyd. Almost certain escape from black holes in final state projection models. *Physical review letters*, 96(6):061302, 2006.

[156] Zhou Lu, Hongming Pu, Feicheng Wang, Zhiqiang Hu, and Liwei Wang. The expressive power of neural networks: A view from the width. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 6231–6239. Curran Associates, Inc., 2017.

[157] Andrew Lucas. Constraints on hydrodynamics from many-body quantum chaos. 2017.

[158] C. Luciuk, S. Smale, F. Böttcher, H. Sharum, B. A. Olsen, S. Trotzky, T. Enss, and J. H. Thywissen. Observation of Quantum-Limited Spin Transport in Strongly Interacting Two-Dimensional Fermi Gases. *Phys. Rev. Lett.*, 118:130405, Mar 2017.

[159] Jean-Philippe W MacLean, Katja Ried, Robert W Spekkens, and Kevin J Resch. Quantum-coherent mixtures of causal relations. *Nature communications*, 8:15149, 2017.

[160] J. Madore. The Fuzzy sphere. *Class. Quant. Grav.*, 9:69–88, 1992.

[161] Juan Maldacena. The large-n limit of superconformal field theories and supergravity. *International journal of theoretical physics*, 38(4):1113–1133, 1999.

[162] Juan Maldacena, Stephen H. Shenker, and Douglas Stanford. A bound on chaos. *JHEP*, 08:106, 2016.

[163] Juan Maldacena, Douglas Stanford, and Zhenbin Yang. Diving into traversable wormholes. *Fortsch. Phys.*, 65(5):1700034, 2017.

[164] Juan Martin Maldacena. The Large N limit of superconformal field theories and supergravity. *Int. J. Theor. Phys.*, 38:1113–1133, 1999. [Adv. Theor. Math. Phys.2,231(1998)].

[165] P. Mazur. Non-ergodicity of phase functions in certain systems. *Physica*, 43(4):533 – 545, 1969.

[166] David A. Mazziotti. Two-electron reduced density matrix as the basic variable in many-electron quantum chemistry and physics. *Chemical Reviews*, 112(1):244–262, 01 2012.

[167] David A. Mazziotti. Enhanced constraints for accurate lower bounds on many-electron quantum energies from variational two-electron reduced density matrix theory. *Phys. Rev. Lett.*, 117:153001, Oct 2016.

[168] John McGreevy, Leonard Susskind, and Nicolaos Toumbas. Invasion of the giant gravitons from Anti-de Sitter space. *JHEP*, 06:008, 2000.

[169] Marko Medenjak, Christoph Karrasch, and Tomaz Prosen. Lower Bounding Diffusion Constant by the Curvature of Drude Weight. *Phys. Rev. Lett.*, 119:080602, Aug 2017.

[170] Mariya V. Medvedyeva, Tomas Prosen, and Marko Žnidarič. Influence of dephasing on many-body localization. *Phys. Rev. B*, 93:094205, Mar 2016.

[171] J J Mendoza-Arenas, S Al-Assam, S R Clark, and D Jaksch. Heat transport in the XXZ spin chain: from ballistic to diffusive regimes and dephasing enhancement. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(07):P07007, 2013.

[172] Márk Mezei and Douglas Stanford. On entanglement spreading in chaotic systems. *Journal of High Energy Physics*, 2017(5):65, 2017.

[173] Márk Mezei and Douglas Stanford. On entanglement spreading in chaotic systems. *JHEP*, 05:065, 2017.

[174] Takeshi Morita and Hiroki Yoshida. Critical Dimension and Negative Specific Heat in One-dimensional Large-N Reduced Models. *Phys. Rev. D*, 101(10):106010, 2020.

[175] Robert C. Myers. Dielectric branes. *JHEP*, 12:022, 1999.

[176] Robert C. Myers. NonAbelian phenomena on D branes. *Class. Quant. Grav.*, 20:S347–S372, 2003.

[177] B. Nachtergaele and R. Sims. Lieb-Robinson Bounds in Quantum Many-Body Physics. *ArXiv e-prints*, April 2010.

[178] Bruno Nachtergaele, Volkher B. Scholz, and Reinhard F. Werner. Local approximation of observables and commutator bounds. In Jan Janas, Pavel Kurasov, Ari Laptev, and Sergei Naboko, editors, *Operator Methods in Mathematical Physics*, pages 143–149, Basel, 2013. Springer Basel.

[179] Bruno Nachtergaele and Robert Sims. Lieb-robinson bounds and the exponential clustering theorem. *Communications in Mathematical Physics*, 265(1):119–130, Jul 2006.

[180] Bruno Nachtergaele and Robert Sims. Locality estimates for quantum spin systems. In Vladas Sidoravičius, editor, *New Trends in Mathematical Physics*, pages 591–614, Dordrecht, 2009. Springer Netherlands.

[181] Bruno Nachtergaele, Anna Vershynina, and Valentin A Zagrebnov. Lieb-Robinson bounds and existence of the thermodynamic limit for a class of irreversible quantum dynamics. *AMS Contemporary Mathematics*, 552:161–175, 2011.

[182] Adam Nahum, Sagar Vijay, and Jeongwan Haah. Operator spreading in random unitary circuits. *Phys. Rev. X*, 8:021014, Apr 2018.

[183] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. CUP, New York, NY, USA, 10th edition, 2011.

[184] Jorge Nocedal and Stephen Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

[185] B. O'Donoghue, E. Chu, N. Parikh, and S. Boyd. Conic optimization via operator splitting and homogeneous self-dual embedding. *Journal of Optimization Theory and Applications*, 169(3):1042–1068, June 2016.

[186] B. O'Donoghue, E. Chu, N. Parikh, and S. Boyd. SCS: Splitting conic solver, version 2.1.2. https://github.com/cvxgrp/scs, November 2019.

[187] Shizuka Okuno, Mariko Suzuki, and Asato Tsuchiya. Entanglement entropy in scalar field theory on the fuzzy sphere. *PTEP*, 2016(2):023B03, 2016.

[188] Ognyan Oreshkov and Nicolas J Cerf. Operational formulation of time reversal in quantum theory. *Nature Physics*, 11(10):853, 2015.

[189] Ognyan Oreshkov and Nicolas J Cerf. Operational quantum theory without predefined time. *New Journal of Physics*, 18(7):073037, 2016.

[190] Ognyan Oreshkov, Fabio Costa, and Časlav Brukner. Quantum correlations with no causal order. *Nature communications*, 3:1092, 2012.

[191] Ognyan Oreshkov and Christina Giarmatzi. Causal and causally separable processes. *New Journal of Physics*, 18(9):093020, 2016.

[192] Román Orús. A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Annals of Physics*, 349:117 – 158, 2014.

[193] L. D. Paniak and R. J. Szabo. Instanton expansion of noncommutative gauge theory in two dimensions. *Commun. Math. Phys.*, 243:343–387, 2003.

[194] George Papamakarios, Iain Murray, and Theo Pavlakou. Masked autoregressive flow for density estimation. In *Advances in Neural Information Processing Systems*, pages 2338–2347, 2017.

[195] Fernando Pastawski, Beni Yoshida, Daniel Harlow, and John Preskill. Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence. *Journal of High Energy Physics*, 2015(6):149, 2015.

[196] Aavishkar A. Patel, Debanjan Chowdhury, Subir Sachdev, and Brian Swingle. Quantum butterfly effect in weakly interacting diffusive metals. *Phys. Rev. X*, 7:031047, Sep 2017.

[197] Vern Paulsen. *Completely Bounded Maps and Operator Algebras*. Cambridge Studies in Advanced Mathematics. CUP, 2003.

[198] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac. Matrix Product State Representations. *Quantum Info. Comput.*, 7(5):401–430, July 2007.

[199] Jacques Pienaar and Časlav Brukner. A graph-separation theorem for quantum causal models. *New Journal of Physics*, 17(7):073020, 2015.

[200] M. B. Plenio and P. L. Knight. The quantum-jump approach to dissipative dynamics in quantum optics. *Rev. Mod. Phys.*, 70:101–144, Jan 1998.

[201] David Poland, Slava Rychkov, and Alessandro Vichi. The conformal bootstrap: Theory, numerical techniques, and applications. *Rev. Mod. Phys.*, 91:015002, Jan 2019.

[202] Joseph Polchinski. Introduction to Gauge/Gravity Duality. In *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2010). String Theory and Its Applications: From meV to the Planck Scale: Boulder, Colorado, USA, June 1-25, 2010*, pages 3–46, 2010.

[203] David Poulin. Lieb-Robinson Bound and Locality for General Markovian Quantum Dynamics. *Phys. Rev. Lett.*, 104:190401, May 2010.

[204] Tomas Prosen. Open XXZ Spin Chain: Nonequilibrium Steady State and a Strict Bound on Ballistic Transport. *Phys. Rev. Lett.*, 106:217206, May 2011.

[205] Tomas Prosen. Lower bounds on high-temperature diffusion constants from quadratically extensive almost-conserved operators. *Phys. Rev. E*, 89:012142, Jan 2014.

[206] Xiao-Liang Qi. Exact holographic mapping and emergent space-time geometry. *arXiv preprint arXiv:1309.6282*, 2013.

[207] Xiao-Liang Qi and Alexandre Streicher. Quantum Epidemiology: Operator Growth, Thermal Effects, and SYK. 2018.

[208] Xiao-Liang Qi and Zhao Yang. Butterfly velocity and bulk causal structure. *arXiv:1705.01728*, 2017.

[209] Xiao-Liang Qi and Zhao Yang. Space-time random tensor networks and holographic duality. *arXiv preprint arXiv:1801.05289*, 2018.

[210] Xiao-Liang Qi, Zhao Yang, and Yi-Zhuang You. Holographic coherent states from random tensor networks. *Journal of High Energy Physics*, 2017(8):60, 2017.

[211] Tibor Rakovszky, Frank Pollmann, and C. W. von Keyserlingk. Diffusive Hydrodynamics of Out-of-Time-Ordered Correlators with Charge Conservation. *Phys. Rev. X*, 8:031058, Sep 2018.

[212] Stuart A Rice. *Reduced-density-matrix mechanics: with application to many-electron atoms and molecules*, volume 165. John Wiley & Sons, 2007.

[213] Katja Ried, Megan Agnew, Lydia Vermeyden, Dominik Janzing, Robert W Spekkens, and Kevin J Resch. A quantum advantage for inferring causal structure. *Nature Physics*, 11(5):414, 2015.

[214] Martin Ringbauer, Christina Giarmatzi, Rafael Chaves, Fabio Costa, Andrew G White, and Alessandro Fedrizzi. Experimental test of nonlocal causality. *Science advances*, 2(8):e1600162, 2016.

[215] Daniel A. Roberts, Douglas Stanford, and Alexandre Streicher. Operator growth in the SYK model. *JHEP*, 06:122, 2018.

[216] Daniel A Roberts, Douglas Stanford, and Leonard Susskind. Localized shocks. *Journal of High Energy Physics*, 2015(3):51, 2015.

[217] Daniel A. Roberts, Douglas Stanford, and Leonard Susskind. Localized shocks. *JHEP*, 03:051, 2015.

[218] Daniel A. Roberts and Brian Swingle. Lieb-robinson bound and the butterfly effect in quantum field theories. *Phys. Rev. Lett.*, 117:091602, Aug 2016.

[219] Daniel A. Roberts and Brian Swingle. Lieb-Robinson Bound and the Butterfly Effect in Quantum Field Theories. *Phys. Rev. Lett.*, 117:091602, Aug 2016.

[220] Daniel A. Roberts and Beni Yoshida. Chaos and complexity by design. *JHEP*, 04:121, 2017.

[221] A. Rubio-García, J. Dukelsky, D. R. Alcoba, P. Capuzzi, O. B. Oña, E. Ríos, A. Torre, and L. Lain. Variational reduced density matrix method in the doubly-occupied configuration interaction space using four-particle N-representability conditions: Application to the XXZ model of quantum magnetism. *The Journal of Chemical Physics*, 151(15):154104, 2019.

[222] Yoav Sagi, Miri Brook, Ido Almog, and Nir Davidson. Observation of Anomalous Diffusion and Fractional Self-Similarity in One Dimension. *Phys. Rev. Lett.*, 108:093002, Mar 2012.

[223] Subhayan Sahu, Shenglong Xu, and Brian Swingle. Scrambling dynamics across a thermalization-localization quantum phase transition. 2018.

[224] Anders W. Sandvik. Finite-size scaling of the ground-state parameters of the two-dimensional Heisenberg model. *Phys. Rev. B*, 56:11678–11690, Nov 1997.

[225] Ulrich Schneider, Lucia Hackermüller, Jens Philipp Ronzheimer, Sebastian Will, Simon Braun, Thorsten Best, Immanuel Bloch, Eugene Demler, Stephan Mandt, David Rasch, and Achim

Rosch. Fermionic transport and out-of-equilibrium dynamics in a homogeneous Hubbard model with ultracold atoms. *Nature Physics*, 8:213, Jan 2012. Article.

[226] Ulrich Schollwöck. The density-matrix renormalization group in the age of matrix product states. *Annals of Physics*, 326(1):96 – 192, 2011. January 2011 Special Issue.

[227] Norbert Schuch and Frank Verstraete. Computational complexity of interacting electrons and fundamental limitations of density functional theory. *Nature Physics*, 5(10):732–735, 2009.

[228] Yasuhiro Sekino and Leonard Susskind. Fast Scramblers. *JHEP*, 10:065, 2008.

[229] Stephen H. Shenker and Douglas Stanford. Black holes and the butterfly effect. *JHEP*, 03:067, 2014.

[230] Stephen H. Shenker and Douglas Stanford. Black holes and the butterfly effect. *JHEP*, 03:067, 2014.

[231] Stephen H. Shenker and Douglas Stanford. Multiple Shocks. *JHEP*, 12:046, 2014.

[232] Edward Shuryak. Physics of Strongly coupled Quark-Gluon Plasma. *Prog. Part. Nucl. Phys.*, 62:48–101, 2009.

[233] David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. Mastering the game of go with deep neural networks and tree search. *Nature*, 529:484, 01 2016.

[234] Barry Simon. Some quantum operators with discrete spectrum but classically continuous spectrum. *Annals of Physics*, 146(1):209 – 220, 1983.

[235] Mark Srednicki. Entropy and area. *Phys. Rev. Lett.*, 71:666–669, 1993.

[236] G. R. Stewart. Non-fermi-liquid behavior in d- and f-electron metals. *Rev. Mod. Phys.*, 73:797–855, Oct 2001.

[237] Leonard Susskind and John Uglum. Black hole entropy in canonical quantum gravity and superstring theory. *Phys. Rev.*, D50:2700–2711, 1994.

[238] M. Suzuki. Ergodicity, constants of motion, and bounds for susceptibilities. *Physica*, 51(2):277 – 291, 1971.

[239] Brian Swingle. Entanglement renormalization and holography. *Physical Review D*, 86(6):065007, 2012.

[240] R. C. Thorne and Harold Jeffreys. The asymptotic expansion of legendre function of large degree and order. *Phil. Trans. Roy. Soc. Lon. Series A, Mathematical and Physical Sciences*, 249(971):597–620, 1957.

[241] Roser Valentí, Claudius Gros, P. J. Hirschfeld, and Walter Stephan. Rigorous bounds for ground-state properties of correlated fermi systems. *Phys. Rev. B*, 44:13203–13212, Dec 1991.

[242] Roser Valentí, Joachim Stolze, and P. J. Hirschfeld. Lower bounds for the ground-state energies of the two-dimensional hubbard and t-J models. *Phys. Rev. B*, 43:13743–13746, Jun 1991.

[243] Mark Van Raamsdonk. Building up spacetime with quantum entanglement. *General Relativity and Gravitation*, 42(10):2323–2329, 2010.

[244] Erik Verlinde and Herman Verlinde. Black hole entanglement and quantum error correction. *Journal of High Energy Physics*, 2013(10):107, 2013.

[245] Brecht Verstichel, Ward Poelmans, Stijn De Baerdemacker, Sebastian Wouters, and Dimitri Van Neck. Variational optimization of the 2DM: Approaching three-index accuracy using extended cluster constraints. *The European Physical Journal B*, 87(3):59, 2014.

[246] Frank Verstraete and J Ignacio Cirac. Renormalization algorithms for quantum-many body systems in two and higher dimensions. *arXiv preprint cond-mat/0407066*, 2004.

[247] Guifré Vidal. Efficient classical simulation of slightly entangled quantum computations. *Physical review letters*, 91(14):147902, 2003.

[248] Guifré Vidal. Class of quantum many-body states that can be efficiently simulated. *Physical review letters*, 101(11):110501, 2008.

[249] C. W. von Keyserlingk, Tibor Rakovszky, Frank Pollmann, and S. L. Sondhi. Operator Hydrodynamics, OTOCs, and Entanglement Growth in Systems without Conservation Laws. *Phys. Rev. X*, 8:021013, Apr 2018.

[250] Christopher David White, Michael Zaletel, Roger S. K. Mong, and Gil Refael. Quantum dynamics of thermalizing systems. *Phys. Rev. B*, 97:035127, Jan 2018.

[251] Steven R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69:2863–2866, Nov 1992.

[252] Edward Witten. Anti de sitter space and holography. *arXiv preprint hep-th/9802150*, 1998.

[253] Michael M Wolf, Frank Verstraete, Matthew B Hastings, and J Ignacio Cirac. Area laws in quantum systems: mutual information and correlations. *Physical Review Letters*, 100(7):070502, 2008.

[254] S. A. Wolf, D. D. Awschalom, R. A. Buhrman, J. M. Daughton, S. von Molnár, M. L. Roukes, A. Y. Chtchelkanova, and D. M. Treger. Spintronics: A Spin-Based Electronics Vision for the Future. *Science*, 294(5546):1488–1495, 2001.

[255] Dian Wu, Lei Wang, and Pan Zhang. Solving statistical mechanics using variational autoregressive networks. *Phys. Rev. Lett.*, 122:080602, Feb 2019.

[256] Shenglong Xu and Brian Swingle. Accessing scrambling using matrix product operators. 2018.

[257] Badis Ydri. Impact of Supersymmetry on Emergent Geometry in Yang-Mills Matrix Models II. *Int. J. Mod. Phys.*, A27:1250088, 2012.

[258] Yi-Zhuang You, Zhao Yang, and Xiao-Liang Qi. Machine learning spatial geometry from entanglement features. *Phys. Rev. B*, 97:045153, Jan 2018.

[259] Shiwei Zhang. 15 auxiliary-field quantum monte carlo for correlated electron systems. *Emergent Phenomena in Correlated Matter*, 2013.

[260] Bo-Xiao Zheng and Garnet Kin-Lic Chan. Ground-state phase diagram of the square lattice Hubbard model from density matrix embedding theory. *Phys. Rev. B*, 93:035126, Jan 2016.

[261] Bo-Xiao Zheng, Chia-Min Chung, Philippe Corboz, Georg Ehlers, Ming-Pu Qin, Reinhard M. Noack, Hao Shi, Steven R. White, Shiwei Zhang, and Garnet Kin-Lic Chan. Stripe order in the underdoped region of the two-dimensional hubbard model. *Science*, 358(6367):1155–1160, 2017.

[262] J.M. Ziman. *Electrons and Phonons: The Theory of Transport Phenomena in Solids*. International series of monographs on physics. OUP Oxford, 1960.