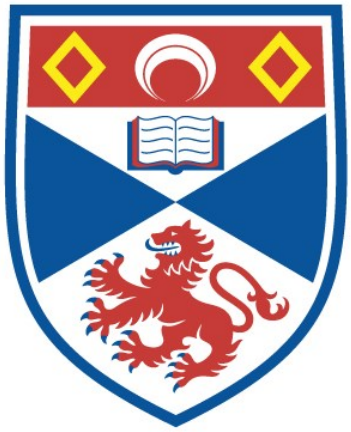
Quantum-enhanced protocols: secure quantum state sharing and noisy field estimation

Cailean James Wilkinson

A thesis submitted for the degree of PhD
at the
University of St Andrews



2025

Candidate's declaration

I, Cailean James Wilkinson, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 47,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree. I confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in September 2020.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date 31/10/2025 Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

Date Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Cailean James Wilkinson, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Electronic copy

No embargo on electronic copy.

Date 31/10/2025 Signature of candidate

Date Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Cailean James Wilkinson, understand that by declaring that I have original research data or digital outputs, I should make every effort in meeting the University's and research funders' requirements on the deposit and sharing of research data or research digital outputs.

Date 31/10/2025 Signature of candidate

Permission for publication of underpinning research data or digital outputs

We understand that for any original research data or digital outputs which are deposited, we are giving permission for them to be made available for use in accordance with the requirements of the University and research funders, for the time being in force.

We also understand that the title and the description will be published, and that the underpinning research data or digital outputs will be electronically accessible for use in accordance with the license specified at the point of deposit, unless exempt by award of an embargo as requested below.

The following is an agreed request by candidate and supervisor regarding the publication of underpinning research data or digital outputs:

No embargo on underpinning research data or digital outputs.

Date 31/10/2025 Signature of candidate

Date Signature of supervisor

COLLABORATION STATEMENT

This thesis covers work I completed between 2020 and 2025 at the University of St Andrews, Scotland, and Macquarie University, Australia. It has been supported by a number of people throughout this time.

Chapters 3 to 5 discusses work completed under the supervision of Prof. Natalia Korolkova and Dr Matthew Thornton at the University of St Andrews. This work, except section 5.4, has been previously published in C. Wilkinson, M. Thornton and N. Korolkova, Phys. Rev. A 107, 062401 (2023).

Chapters 6 and 7 were completed by me under the supervision of Prof. Natalia Korolkova.

All of part II covers work I performed at Macquarie University under the supervision of A/Prof. Alexei Gilchrist and Dr Zixin Huang.

ACKNOWLEDGEMENT OF FUNDING

This work was supported by funding from the University of St Andrews (St Leonard's College and School of Physics and Astronomy), and from the International Macquarie University Research Excellence Scholarship scheme.

RESEARCH DATA ACCESS STATEMENT

Research data underpinning this thesis is available at https://doi.org/10.17630/c34a6ddd-59a8-4b92-90b3-414673a346c4.

ABSTRACT

The first part of this thesis presents a new quantum state sharing scheme that distributes the quantum information contained within a quantum state across three shares such that it cannot be accessed from any individual share. By combining any two shares, however, the original state can be reconstructed. This requirement for collaboration ensures security against single dishonest actors. We demonstrate that the protocol is provably secure for the class of pure Gaussian states and is effective for the sharing of mixed states, although security for those cannot be guaranteed.

We then go on to discuss the use of quantum state sharing as a hybrid protocol for the distribution of discrete-variable states, including Fock states and particle-number qubits, using Gaussian entanglement. We demonstrate that, with access to suitable entanglement resources, this can be achieved with arbitrarily-high fidelity and that the security of the protocol can be guaranteed for qubit-like states.

In the second part of this thesis, we consider the potential for quantum entanglement to improve the measurement of gradients in the magnetic field. We find that in the absence of noise, it is optimal to measure orthogonal gradients individually, devoting the full measurement network to the measurement of a single gradient at a time. In the presence of high levels of environment noise, however, it becomes preferable to measure them simultaneously. We find the optimal network configurations and entanglement structures to make these measurements in the presence of three common noise sources.

ACKNOWLEDGMENTS

This thesis could not have been completed without the support of a huge number of people.

First and foremost, I would like to thank my supervisor, Professor Natalia Korolkova, for her mentorship and encouragement over the past four years; and not-least for patiently unpicking the sometimes near-impenetrable writing in drafts preceding this work. Thank you especially for pushing me to explore new areas and widen my research interests. I have enjoyed every moment of this PhD, and have learned more than I would have imagined under your supervision.

I would also like to thank Dr Alexei Gilchrist and Dr Zixin Huang at Macquarie University in Sydney, who supervised the second part of this thesis, for introducing me to the world of quantum metrology. Your passions for quantum mechanics are infectious, and I learned a huge amount from my time working with you both.

Mention should also go to Dr Matt Thornton; it is because I so enjoyed my time working with you and Natalia during my Master's project that I went on to do this PhD. I am further grateful to Dr Kirill Fedorov, Karolina Weber, and Wun Kwan Yam at the Walther-Meißner Institute in Munich for many useful conversations relating this research back to the real world.

There are too many other members of the university community who have helped me throughout my time at St Andrews to list here, but particular mention should go to the admin team at the School of Physics and Astronomy and to the St Andrews Global Office, who went above and beyond to accommodate changes to my PhD programme. My thanks also to my examining committee, Professor Brendon Lovett at St Andrews and Dr Petros Wallden from the University of Edinburgh, for their useful feedback during my viva, pointing me to many interesting applications I hadn't considered.

Thank you also to all the undergraduate students I had the pleasure to teach during my time in St Andrews for so many enjoyable conversations reminding me that physics is indeed fun. Particular mention must go to Dan Travers and Rosie Gittings, whose projects I had the privilege to supervise. You both produced some remarkable work and I found myself endlessly impressed by your insight and grasp of the field; I look forward to seeing what you achieve next.

I am fortunate to be surrounded by good friends, without whose constant support and encouragement I would not have been able to complete this PhD. To Grace, Jess, Rebecca, Kieran, and Ruth: thank you for your friendship and for many nights of good food and great fun. You are all incredible, funny, thoughtful people and I am so grateful to know you.

Finally, thank you to my parents and to my sister, Poppy, for your love and support and for giving me the encouragement to pursue this PhD. I appreciate you more than I can possibly say.

CONTENTS

1

INTRODUCTION TO THESIS

As we approach the second quantum revolution [1, 2] and the future development of a quantum internet [3–5], interest in the applications of quantum mechanics has surged within academia, industry, and governments. While some of the benefits promised — particularly the use of scalable quantum computers and consumer quantum-communication devices — may still appear only on the horizon, a huge number of explicitly quantum and quantum-informed technologies have already been commercialised and progress towards these more 'science-fiction'-style technologies is being made more rapidly every year. Existing applications of quantum technologies include the use of quantum random number generators and quantum key distribution (QKD) to support specialist[1] cryptography systems [1] and the use of quantum sensors for the detection of leaks in underground water networks [8]. Potential future uses, meanwhile, cover applications as broad as GPS-less navigation systems [9, 10] and the archeological mapping of now-buried ancient structures [11]. As Dr Cathy Foley, Australia's Chief Scientist, put it: the 'impact of the quantum revolution will be comparable to the digital revolution that brought us transistors and lasers' [12].

Countries around the globe are recognising the transformative effect quantum technologies can have, both economically[2] and socially.[3] This confidence in quantum technologies is underscored by the rush in recent years from governments to support quantum industry, with the United Kingdom [14], Australia [12], and the European Union [1, 15], among others, developing detailed (inter)national quantum strategies. Indeed, many industries are preparing for the quantum future while the capabilities they

[1] Although so-called 'plug-and-play' QKD implementations have been developed [6], the UK's National Cyber Security Centre continues to advise against their use due to the specialised hardware needed for their implementation and the requirement that the intermediate infrastructure be trustworthy [7]. Solving for this infrastructure-trust problem and improving public quantum literacy such that security conditions can be readily understood are key planks of many national quantum strategies.

[2] The Australian government estimates that the quantum technology sector will grow by as much as 30% a year globally over the next 5 years and will directly contribute between A$6.1 billion and A$9 billion to Australia's GDP by 2045 [12].

[3] The benefits of improved scans from quantum sensing and drug development from quantum computers are likely to revolutionise the delivery of medical care [13], improving the quality of diagnoses and the effectiveness of treatment. The applications to materials science, meanwhile — particularly to the development of solar panels and battery technology [14] — may accelerate our ability to tackle the climate crisis.

need do not yet exist. Transport for New South Wales, for example, are already contracting quantum computing companies to develop revolutionary traffic-routing algorithms to be ready to reap the benefits once the quantum computing sector is able to support them [12].

Broadly, quantum technologies fall into three main categories: quantum communications, in which the power of quantum mechanics is leveraged to enable more secure communication networks; quantum sensing, in which networks of entangled measurement probes are used to reach beyond the limit of what can be seen with classical approaches; and quantum computing, which is predicted to be able to deliver previously unimaginable speedups, particularly in logistical and modelling problems [12]. In this thesis we consider applications within two of these areas: quantum communication and quantum sensing. The thesis is therefore split into two parts, each designed to be read independently alongside the background chapter as follows.

PART I: QUANTUM STATE SHARING In this first part of the thesis we will consider a quantum communication protocol for the secure sharing and transmission of quantum states in the presence of potentially-untrustworthy parties. Such protocols, termed *quantum state sharing* protocols, split the information describing a quantum state between a number of players in such a way that they can only access it by working together. Crucially, though, the full group is not neccessary to reconstruct the original state so no individual player is handed a veto over the information. We propose and analyse a scheme for this to be implemented using continuous-variable entanglement — a source of entanglement readily available in any quantum optics lab — and prove its security for the whole class of pure Gaussian input states. We further show that this protocol is useful and secure beyond the realm of Gaussian states for which it was designed, and could serve as a novel way to share discrete-variable states without the requirement for complex single-photon entanglement sources.

PART II: QUANTUM-ASSISTED FIELD MEASUREMENT In the second part, we will discuss the use of entanglement to improve the measurement of gradients in the magnetic field. We will begin by showing that in the absence of noise there is no better way to estimate multi-dimensional gradients than to simply measure each individually. We will go on to show that this ceases to be the case in the presence of sufficiently strong noise fields, and that an approach that measures multiple gradients simultaneously becomes preferable. Although we find there is often an upper-limit

to the noise levels in which such a quantum approaches remains useful, we demonstrate that this region of quantum advantage can be extended significantly by making changes to the measurement setup that compensate for the noise.

2

INTRODUCTION TO QUANTUM INFORMATION THEORY

This thesis belongs to the field of quantum information theory, the study of quantum states not as descriptions of specific physical systems but as abstract carriers of so-called 'quantum information'. We will not, therefore, strictly define the type of quantum system under discussion beyond the simple geometry of the state. Nonetheless, the area of greatest interest within quantum information, and quantum technology more broadly, is found in the photonic field. For that reason, to aid readability we will often use 'photon' as a short-hand for any bosonic particle type; this should not be mistaken as an indication that such results apply only to photonic systems.

We will work in natural units throughout this thesis, such that $\hbar = 1$; we will therefore neglect $\hbar$ in many formula it might usually be found.

2.1 QUANTUM INFORMATION

2.1.1 *A brief review of quantum mechanics*

Let us begin by recapping some fundamental results from the field of quantum mechanics that we will use throughout this thesis.

PURE QUANTUM STATES AND MEASUREMENT BASES A quantum system is associated with some d-dimensional Hilbert space, $\mathcal{H}$: a vector space over the complex field equipped with an inner product $\langle\cdot|\cdot\rangle \in \mathbb{C}$ that defines the distance between two elements.[1] Every possible quantum state of the system belongs to this Hilbert space, and any normalised vector within this Hilbert space is a possible state of the system. We denote such pure states, representative of all the quantum information contained within the system, in ket notation as $|\psi\rangle \in \mathcal{H}$.

As the quantum state is a vector element in a vector space, it does not possess a unique mathematical representation; instead we define it with respect to some basis. Ordinarily, this basis will correspond to some selected observable property, $\hat{M}$, of the system. When a state has a definite value

[1] A Hilbert space additionally imposes some completeness conditions which ensure calculus works as intended but are otherwise of no direct relevance to this work.

of this property, it is known as an eigenstate of the observable, denoted $|M_i\rangle$, with eigenvalue, m_i, corresponding to the observable value. Such eigenstates do not change when acted upon by $\hat{M}$,

$$\hat{M}|M_i\rangle = m_i|M_i\rangle, \tag{1}$$

instead accumulating a scalar coefficient corresponding to its eigenvalue. This set of eigenvalues defines the possible outcomes from a measurement on the system, while the eigenvectors represent the corresponding set of potential states this measurement could leave the system in.

If two quantum systems are associated with two Hilbert spaces, $\mathcal{H}_A$ and $\mathcal{H}_B$, they can also be modelled as a single system described by the tensor product of the Hilbert spaces, $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. The state of the collective system is given by the corresponding tensor product of the states of the subsystems as

$$|\psi\rangle = |\psi_1\rangle_A \otimes |\psi_2\rangle_B \in \mathcal{H}_A \otimes \mathcal{H}_B. \tag{2}$$

When n identical subsystems are in the same state, they will be denoted by the shorthand

$$|\psi\rangle = |\psi_1\rangle^{\otimes n} = |\psi_1\rangle \otimes |\psi_1\rangle \otimes |\psi_1\rangle \otimes \dots . \tag{3}$$

SUPERPOSITION STATES A phenomenon unique to quantum mechanics is that the properties of quantum states need not have defined values; the outcome of measuring $\hat{M}$ may be probabilistic. Such states are termed *superposition states* and are written as a sum of measurement eigenstates,

$$|\psi\rangle = \alpha_1|M_1\rangle + \alpha_2|M_2\rangle + \alpha_3|M_3\rangle + \dots , \tag{4}$$

where $\alpha_i = \langle M_i|\psi\rangle$ is a probability amplitude encoding both the probability of finding the state in eigenstate $|M_i\rangle$ after measurement and a relative complex phase between the eigenstates.

As such superpositions do not exist in the classical world, at the point that $|\psi\rangle$ is measured its superposition will collapse into one of these eigenstates, $|M_i\rangle$, with probability given by the inner product $|\langle M_i|\psi\rangle|^2$. To be a complete description of the quantum state, $|\psi\rangle$ must be normalised such that these probabilities sum to 1, as $\sum_i \alpha_i^2 = 1$.

If this conversion to the classical world is inherently destructive — losing these useful superposition features — how then can quantum information be transmitted between, say, quantum computers? It is precisely this prob-

lem that underpins much of the field of quantum communication, and which we will consider in part I.

MIXED STATES AND THE DENSITY MATRIX Although these states are a complete mathematical representation of the *quantum* information contained within a system, they represent an idealised version of the real world. In addition to the quantum superposition — representative of state information 'unknown to the universe', so to speak — the state may evolve in classically probabilistic ways unknown only to us. For example, perhaps we know that a system is *either* in state $|\psi\rangle$ or in state $|\phi\rangle$ with probability 1/2, different to the quantum picture in which the state is in *both states at once*. To enable us to capture both forms of uncertainty let us first promote the ket vector describing a quantum state $|\psi\rangle$ to a projector,

$$\hat{\rho}_{\text{pure}} = |\psi\rangle\langle\psi|, \tag{5}$$

an operator from the Hilbert space of states back to itself. As these projectors are the outer product of two vectors, this representation can be written as a matrix and so is termed the *density matrix* of the state.

Classical uncertainty as to which quantum state the system is in can be represented by summing these pure-state projectors,

$$\hat{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \tag{6}$$

weighted by the probabilities p_i that the system will be in state $|\psi_i\rangle$ and normalised such that $\sum_1 p_i = 1$. States that can be written as a single projector, $|\psi\rangle\langle\psi|$ — and so consist of quantum uncertainty only — are termed *pure states*, while those that do not — and so contain classical uncertainty also — are termed *mixed states*. We can measure the extent to which a state is mixed through its purity, given by the trace of the density matrix,

$$\mathcal{P}(\hat{\rho}) = \text{Tr}(\hat{\rho}^2). \tag{7}$$

In contrast to this classical uncertainty, quantum uncertainty — or superposition — presents itself in the density matrix through the emergence of off-diagonal elements. Consider the density matrix for a state described by $\alpha_1|\psi_1\rangle + \alpha_2|\psi_2\rangle$, for example, which is given by

$$\begin{aligned}\hat{\rho} &= \left(\alpha_1^\dagger|\psi_1\rangle + \alpha_2^\dagger|\psi_2\rangle\right)\left(\alpha_1\langle\psi_1| + \alpha_2\langle\psi_2|\right) \\ &= |\alpha_1|^2|\psi_1\rangle\langle\psi_1| + |\alpha_2|^2|\psi_2\rangle\langle\psi_2| + \alpha_1^\dagger\alpha_2|\psi_1\rangle\langle\psi_2| + \alpha_2^\dagger\alpha_1|\psi_2\rangle\langle\psi_1|. \end{aligned} \tag{8}$$

Were this superposition instead described by classical uncertainty, and the density matrix given by equation (6), these *coherence terms* of the form $|\psi_1\rangle\langle\psi_2|$ would not exist.

THE UNCERTAINTY PRINCIPLE The defining feature of quantum systems — and the one that enables much of the security we rely on in quantum communication — is given by the uncertainty principle. Recall that making a measurement of a quantum system is inherently an information-destructive act. Often, making a measurement of one observable will destroy information about the state of the system with respect to a different observable.

The degree to which two observables $\hat{A}$ and $\hat{B}$ are compatible is given by their commutator,

$$[\hat{A}\,;\hat{B}] \coloneqq \hat{A}\hat{B} - \hat{B}\hat{A}. \tag{9}$$

If two observables do not commute — if it matters which is measured first — the measurement of one must have an impact on the measurement of the other. The two observables cannot therefore both simultaneously be perfectly measured and so no state can be an eigenstate of both. Making a measurement of one will induce some level of superposition in the other.

It is this principle more than any other that underpins the security of quantum protocols. As we will discuss further in section 4.3, these fundamental limits on information restrict our ability to clone quantum states,[2] and so forbid potential adversaries from eavesdropping on the communication without detection.

2.1.2 *Continuous-variable systems and the Wigner function*

Up to this point, we have considered quantum systems with discrete eigenspectra — so-called *discrete-variable* (DV) systems. Throughout the first part of this thesis, however, we will be interested in *continuous-variable* (CV) systems represented by observables with a continuous spectrum of possible measurement outcomes.

The canonical example of such a system is given by the electromagnetic field, the object of main study within quantum optics. Although we do not restrict our results to those observed by the quantum optical field, photonic systems are the predominant medium in which CV systems are implemented and so we will often use the language of the field here.

[2] Were we able to clone states, one could trivially measure incompatible observables by first duplicating the state then performing one measurement on each.

HEISENBERG PICTURE AND QUADRATURE OPERATORS Let us first consider the modelling of such systems in the Heisenberg picture, in which the measurement operators themselves represent the state of the system.

A continuous-variable state consists of a conjugate pair of observables, typically denoted $\hat{x}$ and $\hat{p}$ and referred to as the x and p *quadrature operators* describing the system. For a single-particle system, these might be the position and momentum observables. More typically, in a quantum optical systems these observables might represent the in- and out-of-phase components of the electromagnetic field.[3] Although we will continue to refer to them as the x and p quadratures, in formulae we will ordinarily denote them respectively as $\hat{X}^+$ and $\hat{X}^-$ for notational convenience. These quadrature operators do not commute, with

$$[\hat{X}^+;\hat{X}^-] = \mathrm{i}\,\hbar, \tag{10}$$

and so cannot be measured simultaneously; every continuous-variable system must exhibit a minimum level of uncertainty,

$$\Delta\hat{X}^+\Delta\hat{X}^- \geq 1, \tag{11}$$

where we have assumed, as we will throughout this thesis, that $\hbar = 1$. This base level of CV uncertainty is termed the standard quantum limit.

When both quadratures evolve in the same way, we will sometimes use the *mode operators*, $\hat{a}$ and $\hat{a}^\dagger$ to describe the evolution of the mode as a whole. These combine the two quadratures as

$$\hat{a} = \frac{1}{\sqrt{2}}(\hat{X}^+ + \mathrm{i}\,\hat{X}^-), \quad \text{and} \quad \hat{a}^\dagger = \frac{1}{\sqrt{2}}(\hat{X}^+ - \mathrm{i}\,\hat{X}^-). \tag{12}$$

SCHRÖDINGER PICTURE Often, it will be convenient to return to the Schrödinger picture and consider the probabilistic description of the state directly. A continuous-variable analogue to the state vector, $|\psi\rangle$, can be found by projecting the state onto the quadrature operators, as

$$\psi(x) = \langle\hat{x}|\psi\rangle \in L^2, \quad \text{and} \quad \tilde{\psi}(p) = \langle\hat{p}|\psi\rangle \in L^2. \tag{13}$$

The resultant probability density distribution, termed the state's *wavefunction*, is a vector in the infinite-dimensional Hilbert space of square-

[3] As we are not considering the quantum information as a component of a specified system, we will not discuss here the quantisation of the electromagnetic field. Interested readers are instead directed to the book on quantum optics by Ulf Leonhardt [16] for details.

integrable functions, L^2. As representations of a probability distribution, these wavefunctions must again be normalised such that

$$\int_{\mathbb{R}} dx|\psi(x)|^2 = 1. \tag{14}$$

For a pure state — in CV systems synonymous with a minimum-uncertainty state — either of these wavefunctions represent a complete description of the state and the two are related through the Fourier transform [17]

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dx\phi(x)\, \mathrm{e}^{\mathrm{i}\, px}\,. \tag{15}$$

For this class of states, then, the choice of quadrature representation is simply a choice of measurement eigenbasis.

WIGNER FUNCTIONS Mixed states — or CV states that do not saturate the uncertainty relation of equation (11) — cannot be properly represented as a wavefunction in L^2 any more than they can be represented as a ket vector, $|\psi\rangle$, in $\mathcal{H}$. To represent such states, we must consider the full (x, p) phase space collectively. We cannot properly represent these states through a two-dimensional probability distribution, however, as this would imply a non-zero probability of finding the state in any arbitrarily-small area in phase space, violating the uncertainty principle.

To sidestep this problem of uncertainty-respecting representation, let us introduce a pseudo-probability distribution termed the *Wigner function*, $W(x, p)$ [16, 18]. This function associates with every point in phase space a real number related to the quadrature probability distributions, but *which does not directly correspond to a probability.* Instead, it is defined such that its marginal distributions describe the probability of finding the quadrature measurements as

$$\mathrm{pr}(x) = \int_{\mathbb{R}} dpW(x, p), \quad \text{and} \quad \mathrm{pr}(p) = \int_{\mathbb{R}} dxW(x, p). \tag{16}$$

Indeed, the fact that the Wigner function is not a proper probability density is underscored by another intriguing feature of the distribution: it allows for negative values. Classically, a Wigner function must be strictly positive so this curious property can be used as a blunt metric for the 'quantumness' of a state.

The Wigner function for a pure state with wavefunction $\psi(x)$ or $\tilde{\psi}(p)$ can be found by integrating over the wavefunction as [16]

$$W(x,p) = \frac{1}{2\pi}\int_{\mathbb{R}} dx' \psi^\dagger\Big(x+\frac{x'}{2}\Big)\psi\Big(x-\frac{x'}{2}\Big)\mathrm{e}^{\mathrm{i}px'} \tag{17}$$

$$= \frac{1}{2\pi}\int_{\mathbb{R}} dp' \tilde{\psi}^\dagger\Big(p+\frac{p'}{2}\Big)\tilde{\psi}\Big(p-\frac{p'}{2}\Big)\mathrm{e}^{\mathrm{i}xp'}\,. \tag{18}$$

Precisely as a mixed-state density matrix can be constructed by summing over pure-state projectors, mixed-state Wigner functions are given simply by the sum of pure state Wigner functions as

$$W(x,p) = p_1 W_1(x,p) + p_2 W_2(x,p) + \dots, \tag{19}$$

weighted by the set of classical probabilities $\sum_i p_i = 1$.

Superposition states, by contrast, are again characterised by the presence of coherence terms taking the place of the off-diagonal elements of the density matrix. Applying equation (17) to the superposition state $\psi(x) = \alpha_1\psi_1(x) + \alpha_2\psi_2(x)$, for example, gives Wigner function

$$W(x,p) = |\alpha_1|^2 W_1(x,p) + |\alpha_2|^2 W_2(x,p) + \alpha_1^\dagger\alpha_2 I_{1,2}(x,p) + \alpha_2^\dagger\alpha_1 I_{2,1}(x,p), \tag{20}$$

for coherence terms

$$I_{i,j}(x,p) \coloneqq \frac{1}{2\pi}\int_{\mathbb{R}} dx' \psi_i^\dagger(x+\frac{x'}{2})\psi_j(x-\frac{x'}{2})\,\mathrm{e}^{\mathrm{i}px'}\,. \tag{21}$$

Tensor-product states are represented by the product of the Wigner functions,

$$W_{\psi\otimes\phi}(x_1,p_1,x_2,p_2) = W_\psi(x_1,p_1)W_\phi(x_2,p_2), \tag{22}$$

with each mode represented by its own conjugate pair of quadrature variables. Extracting a single mode from the Wigner function representing a wider system, meanwhile, can be achieved by integrating over the quadrature variables representing any unwanted modes as

$$W_{\text{subssys}}(\boldsymbol{q}_1) = \int_{\mathbb{R}^{N'}} d\boldsymbol{q}_{2,\dots,N} W_{\text{sys}}(\boldsymbol{q}), \tag{23}$$

When considering Wigner functions, we will often for simplicity denote both quadrature variables together as $\boldsymbol{q} = (x,p)^T$. When denoting multiple modes together, we will ordinarily use 'mode-ordering' with $\boldsymbol{q} = (x_1,p_1,x_2,p_2,\dots,x_n,p_n)^T$. The exception to this will be in chapter 6

and in appendix B in which we will use 'xp-ordering' in which the quadratures are grouped by type and $\boldsymbol{q} = (x_1, x_2, \dots, x_n, p_1, p_2, \dots, p_n)^T$.

GAUSSIAN STATES Of particular interest to this thesis is the subset of continuous-variable states whose Wigner function takes the form of a Gaussian,

$$W(\boldsymbol{q}) = \frac{1}{\pi^n \sqrt{\det V}} \exp\left[-(\boldsymbol{q} - \bar{\boldsymbol{r}})^T V^{-1} (\boldsymbol{q} - \bar{\boldsymbol{r}})\right], \tag{24}$$

perhaps unsurprisingly, termed the *Gaussian states*. These states are entirely characterised by the vector $\bar{\boldsymbol{r}} = (\langle \hat{x}_1 \rangle, \langle \hat{p}_1 \rangle, \langle \hat{x}_2 \rangle, \langle \hat{p}_2 \rangle, \dots)^T \in \mathbb{R}^n$ describing their mean position in phase space and a covariance matrix $V \in \mathbb{R}^{2n}$ describing the shape of the distribution.

The diagonal elements of this covariance matrix are given by the variances of each quadrature,

$$V_{i,i} = \frac{1}{2} \langle \{\Delta q_i \, ; \, \Delta q_i\} \rangle , \tag{25}$$

representing how noisy they are individually. The off-diagonal elements — termed the covariances — represent any correlation between the quadratures and are equivalently defined as

$$V_{i,j} = \frac{1}{2} \langle \{\Delta q_i \, ; \, \Delta q_j\} \rangle , \tag{26}$$

for $\boldsymbol{q} = (\hat{x}, \hat{p})^T$. This direct relationship between the quadrature operators and the form of the Wigner function makes it trivial to convert Gaussian states between the two pictures. When a mode $\hat{X}^{\pm}$ can be specified by relation to a set of (independent and uncorrelated) known modes

$$\hat{X}^{\pm} = \alpha_1 \hat{X}_1^{\pm} + \alpha_2 \hat{X}_2^{\pm}, \tag{27}$$

the mean vector and covariance matrices are given simply by the equivalent relations,

$$\bar{\boldsymbol{r}} = \alpha_1 \bar{\boldsymbol{r}}_1 + \alpha_2 \bar{\boldsymbol{r}}_2, \tag{28}$$

$$V = \alpha_1^2 V_1 + \alpha_2^2 V_2. \tag{29}$$

Recalling that the Wigner function describing a tensor-product state is simply given by the product of the Wigner functions, for two Gaussian states the combined system is again a Gaussian state characterised by the

direct product of the subsystem covariance matrices and concatenation of the mean vectors as[4]

$$V_{\psi\otimes\phi} = V_\psi \oplus V_\phi \tag{30}$$

$$\bar{\boldsymbol{r}}_{\psi\otimes\phi} = \bar{\boldsymbol{r}}_\psi \oplus \bar{\boldsymbol{r}}_\phi. \tag{31}$$

The reverse operation, tracing out unwanted modes, is as simple as removing the relevant rows and columns from the covariance matrix; the first subsystem of a state described by

$$V = \begin{pmatrix} V_A & V_{AB} \\ V_{AB}^T & V_B \end{pmatrix} \tag{32}$$

is then simply described by the V_A sub-matrix, while the second subsystem is similarly described by V_B.

2.1.3 *Quantum entanglement and EPR steering*

Quantum entanglement is a form of correlation between two modes of a quantum state that ensures some degree of coordination between their post-measurement states but which cannot be explained classically. This property is so crucial to quantum communication protocols that throughout this thesis we will refer to it as the *resource* underpinning the protocol, and the state that provides the entanglement as the *resource state*.

The wider concept of 'entanglement' is an umbrella term for a class of related properties of quantum states, which are not necessarily equivalent.[5] The simplest of these is the concept of *non-separability*, simply the idea that a multi-mode state cannot be represented as two separate, classically-correlated states. A second, more useful, phenomenon is the ability to affect, or 'steer', the state of one mode of a quantum system solely by acting on the other mode. This is *Einstein–Podolsky–Rosen (EPR) steering* [19], also simply termed *quantum steering*, and is not automatically implied by the presence of non-separability. We will primarily discuss quantum steering in part I, as this is the resource the protocol outlined there utilises, and non-separability in part II, as we are instead interested in the difference between classically-allowed states and fully-quantum states. References to

[4] In this thesis, for notational consistency we denote the concatenation of two vectors as $\boldsymbol{a} \oplus \boldsymbol{b} = (\boldsymbol{a}, \boldsymbol{b})^T$.

[5] For pure states, such as the two-mode squeezed vacuum state we will see later, there is no distinction between these forms of entanglement and any non-separable state will also exhibit some form of quantum steering. In this thesis, we are interested in the full set of in-general mixed resource states and so we will consider non-separability and steering as separate properties.

'entanglement' in those sections should therefore be taken to refer to EPR steering and non-separability respectively, except where otherwise noted.

NON-SEPARABILITY When a pure quantum system can be written as the tensor product of two smaller subsystems, as

$$|\Psi\rangle = |\psi\rangle_a \otimes |\phi\rangle_b, \tag{33}$$

they can be separated and the state of either subsystem can be fully specified without reference to the other. These are termed *separable states*.

When the wider system is itself a superposition of states, for example as

$$|\Phi\rangle = \sum_i \alpha_i |\psi_i\rangle_{ab}, \tag{34}$$

and cannot be written in the form of equation (33) a fully-quantum description is only possibly by considering the system collectively. Any description of the subsystems individually will be a mixed state, with the addition of some classical uncertainty representing our lack of knowledge of the state of the remaining component. These states are termed *non-separable* states, or simply entangled states. The existence of this non-separability in itself will be the object of interest in the second part of this thesis, where it is used to define the set of states impossible to construct classically.

Not all such entangled states are equally non-separable, however. In the latter part of this thesis, we will quantify the degree to which a quantum state is non-separable through a metric termed the *logarithmic negativity* [20].[6] Fully separable states have a logarithmic negativity of 0, with increasing values indicating greater entanglement. The logarithmic negativity corresponding to maximal entanglement is dependent on the size of the subsystems. For two subsystems consisting of a single qubits each, maximal entanglement is indicated by a logarithmic negativity of 1; for subsystems each containing two qubits the maximal entanglement corresponds to a logarithmic negativity of 2. The logarithmic negativity between two subsystems of a three subsystem state is defined as the logarithmic negativity of the state after the third subsystem has been traced out.

EPR STEERING To enable the quantum state sharing protocol outlined in this thesis to operate, a stronger form of entanglement termed *EPR steering* is required. This is the ability for a measurement on one subsystem

[6] Loosely, the logarithmic negativity is measure of how much a state ceases to be valid when one subsystem is transposed and the other is not [20]. For separable states this is a valid operation as the subsystems can be acted upon individually.

of the resource state to non-locally affect the state of the other subsystem, and is not automatically implied by the existence of non-separability.

The degree to which a state exhibits quantum steering is quantified by the *steering parameter*, $E_{1|2}(g)$, given by [21]

$$E_{1|2}(g) = \Delta^2(\hat{X}_1^+ - g\hat{X}_2^+) = \Delta^2(\hat{X}_1^- + g\hat{X}_2^-). \tag{35}$$

with $E_{1|2}(g) < 1$ required for steering to be certified, and $E_{1|2}(g) \to 0$ approaching perfect entanglement. Here, Δ^2 is again the variance operator,

$$\Delta^2\hat{O} = \langle\psi|\hat{O}^2|\psi\rangle - \langle\psi|\hat{O}|\psi\rangle^2. \tag{36}$$

For continuous-variable states, EPR steering is then equally a measure of how well the modes destroy each other when mixed with ratio g as

$$\hat{X}_1^+ - g\hat{X}_2^+, \quad \hat{X}_1^- + g\hat{X}_2^-. \tag{37}$$

Every two-mode state will have a distribution of steering parameters across the range $g \in (0, \sqrt{2})$ that describes its entanglement properties.[7] Such a state exhibits quantum steering for *any g value for which* $E_{1|2}(g) < 1$; except for very lightly-entangled states there will then be a range of g values for which the state is steerable. Notably, quantum steering is a directional property — it may well be the case that a state exhibit steering from one mode to the other, but not vice versa.

GAUSSIAN STATE ENTANGLEMENT Within Gaussian states, entanglement properties are wholly characterised by the covariance matrix describing the state, V. It turns out, though, that all two-mode Gaussian states can be brought into so-called 'standard form',

$$V = \begin{pmatrix} n & 0 & c & 0 \\ 0 & n & 0 & -c \\ c & 0 & m & 0 \\ 0 & -c & 0 & m \end{pmatrix}, \tag{38}$$

only by the action of a series of local operations acting independently on each mode [22], for $n, m > 1$ and $|c| \leq \sqrt{nm-1}$ termed the state's symplectic invariants. Consequently, in any discussion exclusively focused on their entanglement properties, we can assume all two-mode Gaussian states to take this form. In this form, the n and m parameters can be clearly

[7] Outside of this $g \in (0, \sqrt{2})$ EPR steering cannot exist, as it would otherwise be possible to use it to violate the uncertainty limit.

identified as representative of the single-mode variances of each mode while c represents the strength of the entanglement between them. Such a state is pure only when this covariance matrix has unit determinant, when $c = \sqrt{nm - 1}$. Such states exhibit the minimum single-mode variance possible for the amount of entanglement present. The steering parameter for a Gaussian state is given in terms of its symplectic invariants by

$$E_{1|2}(g) = n + g^2 m - 2gc. \tag{39}$$

2.1.4 *Quantum dynamics*

Changes in the state of the system — for example, because of the action of a quantum protocol — are represented through the application of quantum channels, transforming the states as

$$\hat{\rho}_{\text{out}} = \hat{\Lambda}(\hat{\rho}_{\text{in}}). \tag{40}$$

QUANTUM CHANNELS IN THE DIRAC FORMALISM When such a quantum channel always maps pure states to pure states it is termed a *quantum unitary*,

$$|\psi\rangle \mapsto \hat{U}|\psi\rangle, \tag{41}$$

and can be represented as a matrix transform on the ket vectors, as

$$|\psi_{\text{out}}\rangle = U|\psi_{\text{in}}\rangle, \tag{42}$$

for U a matrix representing the action of the unitary. Such a channel represents only the movement of quantum information between the modes present in the state, with no interference from outside sources. These unitaries can equally act on mixed states described by density matrices through the matrix transform

$$\hat{\rho}_{\text{out}} = U\hat{\rho}_{\text{in}}U^{\dagger}. \tag{43}$$

In general, a quantum channel does not have to preserve the purity of its input, though, and will map pure-state ket vectors to mixed-state density matrices. These more-general quantum channels cannot be represented

through unitary matrices and are instead defined by a set of *Kraus operators*, $\{K_i\}$, describing the evolution of a quantum state as [23]

$$\hat{\rho}_{\text{out}} = \sum_i K_i \hat{\rho}_{\text{in}} K_i^\dagger. \tag{44}$$

To represent a physical process, these operators must be collectively normalised such that $\sum_i K_i^\dagger K_i = I_2$.

The mixedness these channels induce stems from our lack of knowledge of the state of the environment the state interacts with. By extending the quantum system to include modes explicitly representing the environment, though, any n-mode quantum channel can be represented as a quantum unitary over some larger $n+m$-mode system. The n-mode output can then be obtained by tracing out the environment modes from the output state, as

$$\hat{\rho}_{\text{out}} = \text{Tr}_m\big[\hat{U}(\hat{\rho}_{\text{in}} \otimes \hat{\rho}_{\text{env}})\big], \tag{45}$$

where $\hat{\rho}_{\text{env}}$ represents the initial state of the environment system.

GAUSSIAN CHANNELS In the Wigner formalism, the action of a quantum unitary is equivalent to a coordinate transform, acting such that

$$W_{\text{out}}(\boldsymbol{q}) = W_{\text{in}}(\Lambda \cdot \boldsymbol{q}), \tag{46}$$

for $\Lambda \in \text{Sp}(2n;\mathbb{R})$ a symplectic matrix (defined below) representing the unitary. As Gaussian states are wholly characterised by their mean vectors, $\bar{\boldsymbol{r}}$, and covariance matrices, V, the action of a Gaussian channel can be modelled by considered the change in each of these, as

$$\bar{\boldsymbol{r}} \mapsto \Lambda^{-1}\bar{\boldsymbol{r}}, \tag{47}$$

$$V \mapsto \Lambda^{-1} V (\Lambda^{-1})^T. \tag{48}$$

In the Heisenberg picture, quantum unitaries can similarly be modelled through their impact on the quadratures as

$$\hat{\boldsymbol{X}}^{\pm} = T\hat{\boldsymbol{X}}^{\pm}, \tag{49}$$

for transformation matrix $T \in \text{Sp}(2n;\mathbb{R})$ related to the coordinate transformation by $T = \Lambda^{-1}$. To be valid quantum states, both the input and

output states of this transformation must respect the canonical commutation relations that

$$[\hat{x}_i\,;\hat{p}_j] = \mathrm{i}\,\delta_{i,j}, \tag{50}$$

which can be written compactly as

$$[\hat{q}_i\,;\hat{q}_j] = \mathrm{i}\,\Omega_{i,j} \tag{51}$$

for Ω the symplectic form given by

$$\Omega = \bigoplus \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{52}$$

Only those matrices which preserve this symplectic form — those for which $T\Omega T^T = \Omega$ — produce valid output states and so represent physical quantum channels. These matrices are termed *symplectic matrices*, and belong to the *symplectic group*, $\mathrm{Sp}(2n;\mathbb{R})$. Consequently, continuous-variable quantum channels that are also quantum unitaries are often termed *symplectic channels*.

2.1.5 *State discrimination and fidelity*

Finally, let us briefly touch on the subject of state discrimination, the measurement of how different two quantum states are. Throughout the first part of this thesis we will use the similarity between the input and output states of the quantum protocol as a measure of its effectiveness. This can be quantified through the *fidelity* between them, a generalisation of the concept of state overlap to mixed states, defined as [24]

$$\mathcal{F}(\hat{\rho}_1,\hat{\rho}_2) = \left[\mathrm{Tr}(\sqrt{\sqrt{\hat{\rho}_1}\hat{\rho}_2\sqrt{\hat{\rho}_1}})\right]^2. \tag{53}$$

This expression reduces to simply the state overlap

$$\mathcal{F}(|\psi_1\rangle,\hat{\rho}_2) = \langle\psi_1|\hat{\rho}_2|\psi_1\rangle \tag{54}$$

when one of the states is pure. A fidelity of 0 indicates the states are wholly orthogonal, while a fidelity of 1 indicates they are identical.

When both states are Gaussian, the fidelity can be found through the covariance matrix and mean vectors characterising the states as [23]

$$\mathcal{F} = \frac{2}{\sqrt{\Delta + \delta} - \sqrt{\delta}} \exp\Big[-(\bar{\boldsymbol{r}}_1 - \bar{\boldsymbol{r}}_2)^T (V_1 + V_2)^{-1} (\bar{\boldsymbol{r}}_1 - \bar{\boldsymbol{r}}_2)\Big], \tag{55}$$

for

$$\Delta = \det(V_1 + V_2), \tag{56}$$

$$\delta = (\det V_1 - 1)(\det V_2 - 1). \tag{57}$$

In the special case in which one of the states is pure, the δ contribution vanishes and the fidelity reduces to

$$\mathcal{F} = \langle\psi|\hat{\rho}|\psi\rangle = \frac{2^n}{\sqrt{\det(V_1 + V_2)}} \exp[-(\bar{\boldsymbol{r}}_1 - \bar{\boldsymbol{r}}_2)^T (V_1 + V_2)^{-1} (\bar{\boldsymbol{r}}_1 - \bar{\boldsymbol{r}}_2)]. \tag{58}$$

When both states share the same mean vector, $\bar{\boldsymbol{r}}_1 = \bar{\boldsymbol{r}}_2$, the exponential vanishes and the fidelity reduces further,

$$\mathcal{F} = \frac{2^n}{\sqrt{\det(V_1 + V_2)}}, \tag{59}$$

and is proportional simply to the inverse covariance matrices.

2.2 SOME EXAMPLES OF QUANTUM STATES

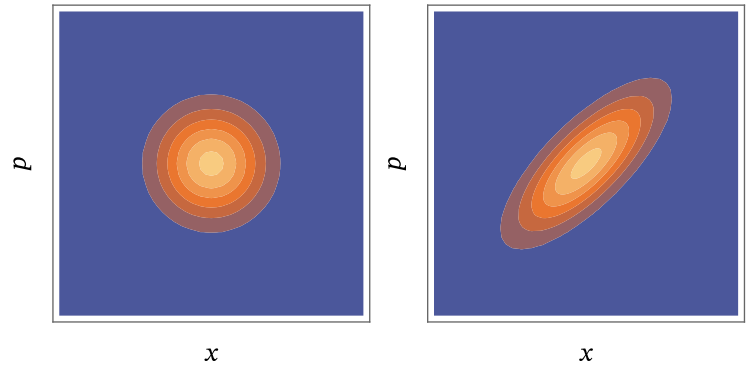
Let us now consider some examples of quantum states that we will encounter in this thesis.

2.2.1 *Gaussian states*

In the first part of the thesis, we will be considering a Gaussian quantum communication protocol. Let us give a brief overview here of some notable Gaussian states.

COHERENT STATES & SQUEEZED STATES The standard Gaussian state is the *coherent state*, characterised by its symmetric saturation of the uncertainty limit such that $V_{\text{coh}} = I_2$ and the variance in each quadrature is equal. The trivial example of this class of state is the vacuum state $|0\rangle$ with Wigner function

$$W(x, p) = \frac{1}{\pi} \exp\big[-x^2 - p^2\big]. \tag{60}$$



(a) Coherent state (b) Squeezed state

Figure 1: Wigner-function representations of (a) a displaced coherent state and (b) a squeezed state with squeezed quadrature variance ≈ 4.5 dB below the vacuum limit at a 45° squeezing angle. Both states saturate the uncertainty limit. In the coherent state, though, the uncertainty is symmetric in both quadratures, while in the squeezed state it has been reduced below the standard quantum limit in one quadrature at the expense of increased uncertainty in the conjugate quadrature.

The remaining class of coherent states carry information through their displacement away from this vacuum state, characterised by the position of their mean in phase space, $\bar{\boldsymbol{r}} \in \mathbb{R}^2$.

A related class of state is the *squeezed coherent state*, often simply termed the squeezed state. These states continue to saturate the uncertainty limit, such that $\Delta^2\hat{X}^+\Delta^2\hat{X}^- = 1$, and so remain pure; however, they do so in an asymmetric way that decreases uncertainty in one quadrature at the expense of increasing it in the other. The covariance matrix for such states is then given by

$$V_{\text{sqz}} = \begin{pmatrix} \mathrm{e}^{-2\zeta}\cos\theta + e^{2\zeta}\sin\theta & 2\sinh(2\zeta)\cos\theta\sin\theta \\ 2\sinh(2\zeta)\cos\theta\sin\theta & \mathrm{e}^{2\zeta}\cos\theta + e^{-2\zeta}\sin\theta \end{pmatrix}, \tag{61}$$

for some squeezing parameter $\zeta > 0$ at an angle θ in phase-space, quantifying the degree to which one quadrature is squeezed below the standard quantum limit. Notably, this covariance matrix always has unit determinant, so continues to saturate the uncertainty limit.

The level of squeezing is more commonly referenced in decibels describing the reduction in noise it causes below the standard quantum limit, given by

$$s = -10\log_{10}\exp(-2\zeta). \tag{62}$$

THERMAL STATES A coherent state that has interacted with some noisy environment and so accumulated additional variance above the

standard quantum limit is termed a *thermal state*.[8] These states have covariance matrix $V = (2\bar{n} + 1)I_2$ for $\bar{n} \geq 0$ representative of the number of photons likely to be found in the state above those that would be found in an equivalent coherent state. Thermal states can also be squeezed, in which case the identity in the covariance matrix is replaced by the squeezed-state covariance matrix of equation (61).

TWO-MODE SQUEEZED VACUUM The standard example of a Gaussian entangled state is the two-mode squeezed vacuum (TMSV) state. This state is constructed by first squeezing two undisplaced vacuum states along complementary axes, such that each is squeezed to the same magnitude, $\pm\zeta \in \mathbb{R}$, and the two have covariance matrix

$$V_{s1} = \begin{pmatrix} e^{2\zeta} & 0 \\ 0 & e^{-2\zeta} \end{pmatrix} \qquad V_{s2} = \begin{pmatrix} e^{-2\zeta} & 0 \\ 0 & e^{2\zeta} \end{pmatrix}. \tag{63}$$

The two squeezed states are then passed through a 50:50 beamsplitter to construct the 2-mode state

$$\hat{X}^{\pm}_{\text{TMSV1}} = \frac{1}{\sqrt{2}}(\hat{X}^{\pm}_{s1} + \hat{X}^{\pm}_{s2}) \tag{64}$$

$$\hat{X}^{\pm}_{\text{TMSV2}} = \frac{1}{\sqrt{2}}(\hat{X}^{\pm}_{s1} - \hat{X}^{\pm}_{s2}), \tag{65}$$

with covariance matrix

$$V_{\text{TMSV}} = \begin{pmatrix} \cosh 2\zeta & 0 & \sinh 2\zeta & 0 \\ 0 & \cosh 2\zeta & 0 & -\sinh 2\zeta \\ \sinh 2\zeta & 0 & \cosh 2\zeta & 0 \\ 0 & -\sinh 2\zeta & 0 & \cosh 2\zeta \end{pmatrix}. \tag{66}$$

Each of these modes are individually indistinguishable from a thermal state, having single-mode covariance matrix $V = \cosh(2\zeta)I_2$, but collectively the system exhibits entanglement between the modes such that

$$\hat{X}^{+}_{\text{TMSV1}} \sim \hat{X}^{+}_{\text{TMSV2}}, \qquad \hat{X}^{-}_{\text{TMSV1}} \sim -\hat{X}^{-}_{\text{TMSV2}}. \tag{67}$$

The quality of this entanglement, quantified by the strength of these correlations, is dependent on the level of squeezing in the two input states. Applying equation (39) to the covariance matrix given in equation (66), we

[8] In this thesis, the thermal states we refer to are those which take Gaussian form: the thermal states of second order Hamiltonians [23]. These are a subset of the more general set of Gibbs thermal states [16].

can find the steering parameter for two-mode squeezed vacuum states to be

$$E_{1|2}(g) = (1 + g^2)\cosh(2\zeta) - 2g\sinh(2\zeta), \tag{68}$$

where ζ is the level of squeezing used to construct the states and g is the mixing ratio described in section 2.1.3.

As TMSV states are the standard entanglement tool in quantum optics, we will often use TMSV squeezing in place of the EPR steering parameter to enable an easier connection to experimental implementations. In such cases, the squeezing will be quantified through decibels using equation (62).

Despite being continuous-variable in nature, these states can be equally represented as a particle-number superposition in an infinite Hilbert space. In that regime, the two-mode squeezed vacuum has representation

$$|\mathrm{TMSV}(\zeta)\rangle = \frac{1}{\cosh\zeta}\sum_{n=0}^{\infty}\tanh^n\zeta|n,n\rangle, \tag{69}$$

where n represents the number of photons in the state and ζ the amount of squeezing present in the state's construction.

2.2.2 *Discrete-variable states*

As the number of particles present and the dimension of each's state space increases, the varieties of possible entanglement also increase hugely. Here, let us focus on a couple of the common ones we will use in our discussion of entangled probe networks in part II. In that part of the thesis, we will be interested in spin-$\frac{1}{2}$ particles, defined by being either in the spin-up or spin-down state, and so let us here use as our example the two-level quantum systems composed of the $\{|\uparrow\rangle, |\downarrow\rangle\}$ spin eigenstates.

BELL STATES When only two particles are present in the state, all possible entanglement is characterised by the set of four Bell states:

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}\big(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle\big), \quad \text{and} \quad |\Phi^-\rangle = \frac{1}{\sqrt{2}}\big(|\uparrow\uparrow\rangle - |\downarrow\downarrow\rangle\big), \tag{70}$$

defining entanglement in which the particles are always found in the same state and

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}\big(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\big), \quad \text{and} \quad |\Psi^-\rangle = \frac{1}{\sqrt{2}}\big(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\big), \tag{71}$$

defining entanglement in which they are always in opposing states. In this thesis, we will only make use of the first form of entanglement, $|\Phi^+\rangle$, which we will denote as simply $|\Phi\rangle$.

GHZ STATES The first two forms of Bell state, $|\Phi^\pm\rangle$ generalise naturally to systems composed of a larger number of particles. The state in which n particles are known to be in the same state, but where that state is not known, is termed the *Greenberger–Horne–Zeilinger*, or GHZ, state, and has form

$$|\mathrm{GHZ}\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle^{\otimes n} + |\downarrow\rangle^{\otimes n}\right). \tag{72}$$

Although a form of maximally-entangled state, the GHZ state is extremely fragile. Making a measurement of the spin of one mode (along the spin axis the GHZ state is defined against) will reveal the state of the entire system, and so leave the remaining system in a fully known eigenstate with no entanglement [25]. Consequently, tracing out any single mode — if the measurement outcome is unknown, for example, or due to the presence of certain types of noise — the remaining system will be left in the maximally mixed state, with the quantum superposition replaced by classical uncertainty.

DICKE STATES Finally, let us consider the class of *Dicke states*: a generalisation of the latter two Bell states $|\Psi^\pm\rangle$. Consider a three-particle state whose total energy is known: we might deduce therefore that a total of 2 of the 3 particles must be in the excited state. If the configuration of these particles is not known, however, the overall state is written

$$|\mathcal{D}(3,2)\rangle = \frac{1}{\sqrt{3}}\left(|\uparrow\uparrow\downarrow\rangle + |\uparrow\downarrow\uparrow\rangle + |\downarrow\uparrow\uparrow\rangle\right). \tag{73}$$

The state of each probe is then dependent on the state of the wider system, and no single probe can be fully specified alone. This is the $(3,2)$ Dicke state. These states remain fragile but are slightly more resilient than GHZ states, with the measurement of a single probe potentially destroying the entanglement (if the measurement result is $|\downarrow\rangle$) or leaving the remaining system in a Bell state (if the measurement is $|\uparrow\rangle$).

The general Dicke state of k excitations distributed across n particles is given by

$$\mathcal{D}(n,k) = \frac{1}{\mathcal{N}} \sum_{P(n,k)} |P\rangle \tag{74}$$

for $P(n,k)$ the set of all possible such permutations and $\mathcal{N}$ a normalisation constant.

2.3 FURTHER READING

For more information about continuous-variable systems and the Wigner function, the reader is directed to the canonical book by Alessio Serafini in Ref. [23] or the review article by Christian Weedbrook *et al.* in Ref. [24]. For an overview of quantum optical systems (although broadly applicable more widely to any bosonic system) the textbooks by Ulf Leonhardt in [16] and Pieter Kok and Brendon Lovett [26] are an excellent resource. For further information on discrete-variable systems the reader is directed to the reference book by Nielsen and Chuang [27].

Part I

QUANTUM STATE SHARING

3

INTRODUCTION

Imagine an aquarium, facing the ever-present risk of theft from rivals, feels the need to up security for its prize hammerhead. Naturally, no single employee can be allowed individual access lest they steal the fish away. On the other hand, access must be possible at any time: what if the shark were to fall ill or otherwise need care? Specified groups of trusted keyholders cannot be relied upon to be present every time the tank is opened. What is needed is a system whereby the tank can be accessed by *any* suitably-large group of employees while any smaller group remains locked out. The solution to this problem lies in encryption techniques that give each person only one part of the key, but allow for *any* combination of a set number of these parts to together form the full key. Schemes of this form are known as secret sharing protocols and are in use in security systems around the globe.

Classically, this is a perfectly solved problem. In 1979, Shamir [28] and Blakley [29] independently presented protocols that are provably unbreakable so long as the individual players' shares remain themselves secure. Shamir's protocol, for example, hides the secret information as the intersection of a $(k-1)$-dimensional polynomial with the y-axis. Each of the n players is then distributed a randomly selected point $(x_i, f(x_i))$ along this curve. As there exists only one $(k-1)$-dimensional polynomial passing through k points, access to any k of these points allows the original curve can be identified exactly, and so the secret information recovered. By contrast, there are an infinite number of such curves passing through $k-1$ points — and one can be found passing through every point along the y-axis — so no information can be recovered from any smaller set of shares. An illustration of this scheme is shown for $k = 2$ in figure 2. Blakley's protocol [29], meanwhile, hides the secret information as the intersection of $(k-1)$-dimensional planes, with one dealt to each player. Only with the knowledge of k such planes is their intersection unique, and thus the secret retrievable.

Although this classical solution is well-established, there remains a place for quantum mechanics in such a setup ensuring the secure distribution of the shares to each player and forbidding their duplication. Such quantum-

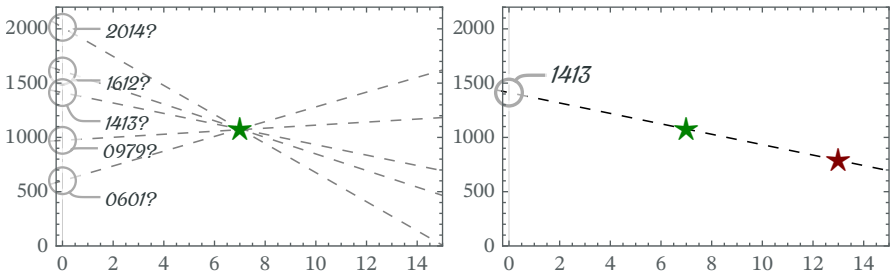


Figure 2: An example of a $k = 2$ classical secret sharing scheme in which a 4-digit secret is encoded as the intersection with the y-axis. (a) With the knowledge of only one point (green star) it is impossible to identify the original line, and so the secret is secure. (b) With access to a second point (red star), though, the original line can be established and the secret code recovered.

augmented protocols are termed quantum secret sharing and are discussed extensively in Refs [30–33] so we will not consider them further here.

In considering this protocol's relationship to quantum mechanics, though, another question arises: *what if the secret itself was a quantum state?* Such a state cannot be converted into classical information without destroying some part of it, so classical information security protocols are of little use to us here. It is this class of protocol, termed quantum state sharing (QSS), that we will consider in this first part of the thesis.

These quantum protocols differ from their classical cousins in a number of important ways as the quantum information must obey the quantum limitations. First, as we will discuss in section 4.1.2, some of the complex access structures possible with classical secret sharing cannot be implemented with QSS. In particular, it is impossible to permit reconstruction with fewer than half of the shares. Further, as the quantum information encoded within the states is vulnerable to decoherence, more care is required in their storage and transmission which may increase infrastructure costs. On the other hand, though, the cloning-prevention built in to quantum mechanics prevents careless parties from sharing their copies with others. These restrictions may make quantum approaches unsuitable for some use-cases, in which case a classical secret sharing scheme be necessary.

In practical terms, these protocols involve mixing a single secret (and unknown to the players) quantum state into a larger entangled system to distribute the quantum information across the modes. Collectively, this wider system perfectly represents the original state but no single mode, traced out, can be used to recover it. By exploiting the entanglement between the modes (a task for which one must necessarily have access to multiple shares), however, the original state can be disentangled from this wider

system and reproduced in a single mode. The quality of this 'reconstructed' copy of the input state will depend on how well-entangled the original 'resource' system was prior to the introduction of the secret state.

Protocols of this form have the potential to find uses across the quantum communication field. There are obvious applications to the task of secure state distribution, where QSS solves a subtly different problem to that of quantum teleportation: how to communicate securely when one has access to a multiple communication routes but cannot guarantee that any is individually secure. By splitting the state using a QSS scheme prior to transmission, the loss of individual transmissions is secured against. A potential adversary would then have to intercept a larger number of transmissions to decode the original information, increasing the cost and access requirements of a successful attack. An adversary failing this stronger requirement would further be unable to prevent successful reconstruction of the secret by the desired recipients. Crucially, as we will see in this thesis, quantum state sharing provides this security at a lower resource cost than quantum teleportation, offering cheaper security against less-sophisticated threat models.

QSS protocols may well find applications outside of simple state distribution, however. In the field of secure distributed computing, there has been promising early work in the discrete-variable regime looking at the use of quantum state sharing schemes for so-called 'blind computation' protocols [34, 35] and secure multipartite quantum computation [36]. In such a scenario, a service-user might utilise a QSS scheme to obscure confidential data before sending each share to be processed by quantum servers operated by different (untrusted) parties. No individual quantum computer then has access to either the original underlying data or the computation output. It is only by bringing the shares back together (at a point controlled by the service-user) that either this information becomes accessible again. Such privacy guarantees could enable sensitive information to be analysed by cloud-based quantum computers without the requirement that the operators be trusted [34]. Although such a solution necessarily involves the use of additional resources — with access to multiple quantum computers required — it guarantees the underlying data remains secure.

Finally, quantum state sharing schemes may find uses as a rudimentary form of quantum error correction, a continuous-variable analogue to Reed-Solomon codes [35, 37]. The ability to split information into a number of shares and reconstruct it, in the ideal case perfectly, with access to only a subset of those shares allows for the possibility to recover from the complete loss of a number of transmissions; so long as more than

half of transmissions are successfully received the original information is preserved. Such quantum error correction protocols are termed erasure correction protocols and serve an important role in making quantum communication robust [38, 39]. In fact, in the discrete-variable picture, it has been shown that *any* $(2k-1)$-dimensional erasure correction protocol that corrects $k-1$ errors can be made into a $\{k, 2k-1\}$-threshold quantum state sharing protocol, and vice versa. It is likely that this also applies to continuous-variable protocols.

For quantum state sharing to be useful in any of these fields, however, its security must be provably guaranteed. It is well-known that for quantum information protocols better security can be achieved by increasing the entanglement present in the resource state, but this comes with significant tradeoffs by increasing the energy and equipment requirements. It is essential, then, that we know precisely how much entanglement is necessary to achieve secure QSS to enable its use in the widest possible variety of setups. As well as considering its general effectiveness, it is primarily this question of security that we will return to throughout our discussion of QSS here.

In chapter 4 we will introduce the continuous-variable quantum state sharing protocol utilising Gaussian entanglement that we are presenting in this thesis. Chapter 5 will consist of the main results of this part, an analysis of the use of this protocol for the whole class of single-mode Gaussian secret states including coherent states, squeezed states, and thermal states, as well as considering the potential for the protocol to be extended to allow for the sharing of multi-mode Gaussian states. In this chapter we will be particularly interested in the conditions under which this protocol is provably secure against bad actors, and the degree of resource entanglement required for this security to be guaranteed. Finally, in chapter 7 we will consider our protocol as a 'hybrid' protocol allowing for Gaussian entanglement to be used to securely share and reconstruct discrete-variable states, using the framework developed in chapter 6. We will consider the protocols effectiveness for the set of Fock eigenstates and a number of superpositions of those.

4

CONTINUOUS-VARIABLE QUANTUM STATE SHARING PROTOCOLS

In this chapter, we will outline the form of the quantum state sharing protocol under consideration. We have focused on the simplest non-trivial quantum state sharing protocol here: (2, 3)-threshold state sharing in which the quantum information is split between three players, and any two of them can collaborate to reconstruct it. This is only one element of the much larger (k, n)-threshold class of quantum state sharing protocols; we discuss in chapter 8 the potential for this work to be extended to cover the larger protocols.

Quantum state sharing consist of two distinct stages, which we will consider in turn: the *dealer protocol*, in which the single-mode secret is split into multiple shares; and the *reconstruction protocol*, in which a subset of these shares are recombined into the original secret state. After the dealer protocol, the quantum information describing the original state is distributed across the three shares, each given to a different player, in such a way that no player acting alone can access it — two players must collaborate to reconstruct the original secret state. Crucially, the participation of the third player is not required for the secret to be accessed. In addition to being locked out from accessing the secret information, an individual bad actor is therefore also unable to sabotage the others' access. A top-level overview of this process is shown in figure 3.

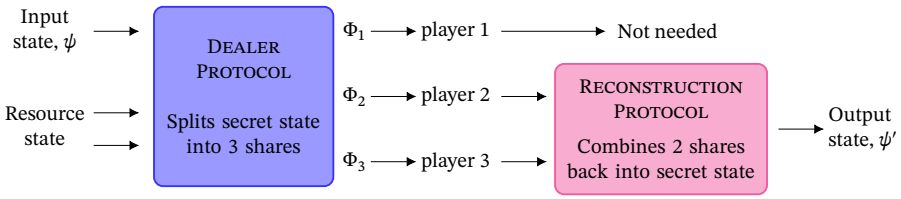


Figure 3: A high-level overview of a {2,3} quantum state sharing protocol. The *dealer* is provided with a secret state unknown to them and splits it into three shares. Any two of these shares can then be combined into the original secret state through a *reconstruction protocol*.

The quantum state sharing protocol we present here belongs to the class of continuous-variable protocols, those that exploit Gaussian entanglement to share potentially infinite-dimensional systems. Discrete-variable quantum state sharing has been the subject of extensive previous study

[31, 32, 40]. In 1999, Cleve *et al.* [40] demonstrated that an n-level GHZ state[1] provided sufficient entanglement to share an n-level superposition state with guaranteed security. Indeed, they showed that in this case (in the absence of transmission loss) not only can the secret be reconstructed perfectly, but any adversaries cannot retrieve any information whatsoever about the original state. Similar protocols have since been presented utilising Dicke-like entanglement [41] and cluster state entanglement [42].

In the continuous-variable regime, such absolute statements about performance and security are not possible, however. As perfect entanglement is not achievable for continuous-variable states, the reconstruction will always be approximate and some information will always leak to the adversaries. We will discuss this problem in more detail in section 4.3, where we introduce the primary security concern in CV quantum communication: ensuring authorised players gain the most information about the original state.

Our protocol is ultimately a generalisation of that proposed by Tyc and Sanders in 2002 [43] and demonstrated by Lance *et al.* in 2003 [44]. We have extended that protocol to allow for the original state to be reconstructed with a non-unity amplification, similar to a recent extension of CV quantum teleportation by He *et al.* in 2016 [45]. This allows for the full exploitation of generally-asymmetric entanglement resources. Before we go on to discuss our protocol, though, let us briefly consider those prior similar protocols.

4.1 A SURVEY OF CONTINUOUS-VARIABLE QUANTUM STATE SHARING

4.1.1 *Tyc & Sanders quantum state sharing*

Quantum state sharing in the continuous-variable regime was first formulated by Tyc & Sanders in 2002 [43] and later expanded by them and Rowe in 2003 and 2004 [46, 47]. They showed, using simple linear algebra, that so long as one could find a dealer protocol whose output satisfies certain linear-independence conditions then later state reconstruction is always possible. Let us briefly sketch their argument here.

[1] In fact, this remains true in the continuous-variable regime. Recall from equation (69) in section 2.2.1 that the maximally-entangled TMSV state has (unnormalised) form

$$|\mathrm{EPR}\rangle = \sum_{n=0}^{\infty} |n, n\rangle, \tag{75}$$

for photon number n. This is an example of an infinite-dimensional GHZ state! Unfortunately, such perfect entanglement cannot exist in the CV regime, and so the resource states we will use here are at-best approximations of this GHZ-like entanglement.

The dealer has access to the secret quantum state $|\psi\rangle$, which is the object of the protocol, and $k-1$ two-mode idealised EPR states, $|\Theta\rangle$. The initial state of the system is then given by

$$|\Phi\rangle = \int_{\mathbb{R}^k} d^k\boldsymbol{x}\ \psi(x_1)|x_1\rangle_1\,|x_2\rangle_2|x_2\rangle_{k+1}\ldots|x_k\rangle_k|x_k\rangle_{2k-1}, \tag{76}$$

where

$$|\psi\rangle = \int_{\mathbb{R}} dx_1\ \psi(x_1)|x_1\rangle_1 \tag{77}$$

represents the secret state and

$$|\Theta\rangle = \int_{\mathbb{R}} dx_i\ |x_i\rangle_i|x_i\rangle_{k-1+i} \tag{78}$$

is a single two-mode EPR state.[2]

Now let us consider the k-dimensional vector $\boldsymbol{x} = (x_1, \ldots x_k)^T$ representing the initial canonical positions of the secret state and each EPR state. The dealer can select a linear transform

$$L : \mathbb{R}^k \to \mathbb{R}^{2k-1} : \boldsymbol{x} \mapsto \big[L_1(\boldsymbol{x}), L_2(\boldsymbol{x}), \ldots, L_{2k-1}(\boldsymbol{x})\big]^T, \tag{79}$$

which they use to encode the secret state into the wider system as

$$|\Phi\rangle = \int_{\mathbb{R}^k} d^k\,\boldsymbol{x}\,\psi(x_1)|L_1(\boldsymbol{x})\rangle_1|L_2(\boldsymbol{x})\rangle_2\ldots|L_{2k-1}(\boldsymbol{x})\rangle_{2k-1}, \tag{80}$$

producing a single $(2k-1)$-mode system collectively representing the original secret state.

Let us assume that the dealer has chosen this linear mapping such that every k-element subset of

$$\big\{x_1, L_1(\boldsymbol{x}), L_2(\boldsymbol{x}), \ldots, L_{2k-1}(\boldsymbol{x})\big\}, \tag{81}$$

is linearly independent. That is to say, any k of the individual transforms L_i are linearly independent, and any $k-1$ of the transforms are linearly-independent with the original x_1 vector.

Now, let us assume that a set of k players wishes to reconstruct the secret state. As we have not introduced any particular ordering to the set of linear mappings, we can assume without losing generality that the players have access to the first k shares, represented by the mappings

[2] As we are considering idealised EPR states here, the state of both modes can be perfectly represented by the state of only one mode and only $k-1$ variables are required.

$\{L_1(\boldsymbol{x}) \ldots L_k(\boldsymbol{x})\}$. This set and the disjoint set $\{x_1, L_{k+1}(\boldsymbol{x}) \ldots L_{2k-1}(\boldsymbol{x})\}$ of the remaining transforms are both linearly independent, as k-element subsets of the set in equation (81). Consequently, there will exist a transformation matrix T between them such that

$$T\begin{pmatrix} L_1(\boldsymbol{x}) \\ L_2(\boldsymbol{x}) \\ \vdots \\ L_k(\boldsymbol{x}) \end{pmatrix} = \begin{pmatrix} x_1 \\ L_{k+1}(\boldsymbol{x}) \\ \vdots \\ L_{2k-1}(\boldsymbol{x}) \end{pmatrix}. \tag{82}$$

This can be used to construct an equivalent unitary transform, $\hat{T}$ on the first set of states such that[3]

$$\hat{T}|L_1(\boldsymbol{x})\rangle_1 \ldots |L_k(\boldsymbol{x})\rangle_k = |x_1\rangle_1 |L_{k+1}(\boldsymbol{x})\rangle_2 \ldots |L_{2k-1}(\boldsymbol{x})\rangle_k. \tag{83}$$

Re-parametersing the system using $\boldsymbol{x}' = (x_1, L_{k+1} \ldots L_{2k-1})^T$ such that

$$\hat{T}|\Phi\rangle = \int_{\mathbb{R}^k} d^k \boldsymbol{x}'\, \psi(x_1)|x_1\rangle|L_{k+1}\rangle_2|L_{k+1}\rangle_{k+1} \ldots |L_{2k-1}\rangle_k|L_{2k-1}\rangle_{2k-1}, \tag{84}$$

the system is returned to its original form

$$\hat{T}|\Phi\rangle = |\psi\rangle_1|\Theta\rangle_{2,k+1} \ldots |\Theta\rangle_{k,2k-1}, \tag{85}$$

for $|\Theta\rangle$ representing the idealised EPR states.

So long as a linear transform L can be found satisfying the linear-independence constraints, therefore, it can be used as a QSS protocol. This turns out to always be possible [40], so a $(k, 2k-1)$-threshold QSS protocol exists for any k. Further, as a (k, n)-threshold scheme can be construction from a $(k, 2k-1)$ scheme for any $n \leq 2k-1$ simply by discarding unwanted shares, any threshold quantum state sharing protocol can be constructed using continuous-variables.

A further proof extending this idea to real-world two-mode squeezed vacuum entanglement was presented in Ref. [46].

[3] Strictly, there should be an additional $\sqrt{|\det T|}$ term here to account for the possibility that the transformation matrix is not itself unit-determinant. We neglect this, and the Jacobian for the coordinate transformation in equation (84), as the states are not normalised to begin with.

In the language of their paper, Tyc & Sanders proposed the form of the dealer protocol that forms the basis of the QSS protocol presented here, described by the set of mappings

$$\left\{L_1 = \frac{x_1 + x_2}{\sqrt{2}},\, L_2 = \frac{x_1 - x_2}{\sqrt{2}},\, L_3 = x_2\right\}. \tag{86}$$

The first two mappings correspond to the output of a balanced beamsplitter mixing the secret state with one mode of the EPR resource state and the third represents the unchanged second mode of the EPR state. As given in that paper, though, this is only one possible QSS dealer protocol; we will prove the optimality of this setup in section 4.2.1.

4.1.2 *Access schemes*

This class of quantum state protocol is termed (k, n)-threshold quantum state sharing; the quantum state is split into n shares with *any* set of shares exceeding the k threshold able to reconstruct it. The subclass of protocols proposed by Tyc & Sanders [43] is specific to $\{k, 2k-1\}$-threshold QSS; those with the minimum threshold required by the uncertainty principle. Any protocol which allowed for a reconstruction with $n/2$ or fewer shares would permit multiple independent reconstructions of the state and therefore violate the no-cloning theorem.

While this may initially present as a limited form of QSS, more complex access schemes can be built up from this protocol through the distribution of shares. In particular, a (k, n)-threshold scheme can be constructed for any $k > n/2$ by starting with a $(k, 2k-1)$-threshold scheme, for example as proposed by Ref. [43], before discarding unnecessary additional shares to reach $n < 2k-1$. Alternatively, a hierarchical access structure could be constructed by distributing shares unevenly to each player. For example, a protocol which always requires the CEO of a company plus any of three executives to reconstruct a state can be constructed from a $(4, 7)$-threshold scheme by distributing 3 shares to the CEO and 1 share to each executive, discarding the remaining share. Meanwhile, a protocol that additionally allows the three executives together to reconstruct the state without the CEO could be constructed from a $(5, 3)$-threshold scheme by giving the CEO two shares and the executives a share each. The flexibility in such schemes remains limited, however; it is not possible to implement an access structure that gives access to disjoint sets of players — any such scheme would again violate the no-cloning theorem.

4.1.3 *Lance et al. quantum state sharing*

The quantum state sharing protocol proposed by Ref. [43] was then expanded upon by Lance *et al.* [44, 48, 49], who turned this mathematical description into a physical implementation for $(2, 3)$-threshold QSS. While the dealer protocol was prescribed by Ref. [43], they proposed a number of potential setups for the subsequent reconstruction of the state, depending on the quantum resources one had available.[4]

The same authors later went on [48] to experimentally demonstrate the quantum nature of the protocol, using two-mode squeezed vacuum (TMSV) entanglement to share and reconstruct coherent states. Their implementations were tuned to the specifics of TMSV resource states, however, and did not translate to the full class of Gasussian entanglement.

4.1.4 *Prior work on this protocol*

A previous iteration of the protocol presented here was presented in my dissertation submitted for the degree of Master of Physics in 2020 [50]. That work presented an adapted version of one specific experimental implementation from the Lance *et al.* protocol [49] that allowed for entanglement asymmetries in the resource state. A looser version of some of the results presented here for coherent states were found in that study, which demonstrated that *for that particular implementation* any form of two-way steering was sufficient for the secure sharing of coherent states. We will show here that *for any such continuous-variable* $\{2, 3\}$*-threshold QSS protocol* in fact only one-way steering is sufficient (and necessary).

There is little overlap with this thesis; indeed, as will be seen in section 5.2.4 we have found here that in fact those results were unnecessarily restrictive and that only *one-way* steering is required. Further, all results in this thesis have been re-derived from solid information-theoretic principles that do not tie the results to a specific setup. For clarity, a full accounting of any similarity between these results and those previously presented can be found in appendix A

[4] The two main options required the use of either two coordinated single-mode squeezing operations or a measure-and-displace 'feed-forward loop'. Both have different experimental challenges and the preferred implementation may depend on the specific domain.

4.2 OUR QUANTUM STATE SHARING PROTOCOL

Let us now finally turn our attention to consider the quantum state sharing protocol presented in this thesis. Like all quantum state sharing schemes, our protocol consists of two broad sub-protocols, as shown in figure 3.

1. The *dealer protocol* takes as input the original secret state and a two-mode entangled resource state. The properties of this resource state are pre-defined and considered public knowledge. The secret state is mixed with one of the resource modes to produce a wider three-mode system collectively describing the secret state. Although all the quantum information describing the original secret state is represented within this system, it is inaccessible from any single mode alone. Each of these modes is then distributed to a single player. This subprotocol is outlined in section 4.2.1.

2. Any two players can work together through the relevant *reconstruction protocol* to combine their modes and reconstruct a (generally imperfect) copy of the original secret state by exploiting the entanglement present in the resource state. The specific subprotocol which reconstructs the secret state will depend on which two players collaborate, which we denote $\{i, j\}$ *reconstruction* when shares i and j are used. $\{1, 2\}$ reconstruction is discussed in section 4.2.2 while $\{1, 3\}$ and $\{2, 3\}$ reconstruction subprotocols are outlined in sections 4.2.3 and 4.2.4.

An illustrative example of the Wigner functions representing the individual modes at each stage of the process for a coherent-state secret is shown in figure 4. In the intermediate share in figure 4b the information describing the secret state has been drowned-out by the relatively

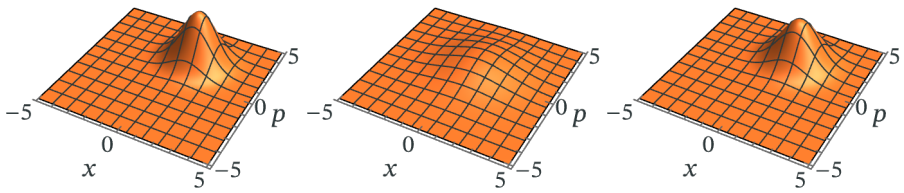


Figure 4: Wigner functions at each stage of the QSS process: (a) input state, (b) intermediate distributed share (share 1), and (c) reconstructed output state. This figure shows the sharing of a coherent-state secret with mean $\bar{\boldsymbol{r}} = (2, 2)^{\mathrm{T}}$ utilising a two-mode squeezed vacuum resource state with 10 dB squeezing.

huge single-mode variance of the resource state. By applying a reconstruction protocol to two of these shares, however, the original state can be reconstructed with increasingly good accuracy as the entanglement in the resource state increases, seen in figure 4c.

4.2.1 *Dealer protocol*

Recall from section 4.1.1 that Tyc & Sanders showed that *any* dealer protocol satisfying certain linear-independence conditions would allow for a later (generally-imperfect) reconstruction of the original state [43]. There are a number of possible dealer protocols satisfying this constraint, though, so a question remains: which to choose? In selecting our dealer protocol let us put aside their linear-independence condition and define our own constraints on what makes a suitable QSS protocol.

Definition 1. *DEALER CHANNEL CONSTRAINTS*

1. *The dealer channel and every reconstruction channel must, like all quantum channels, preserve the canonical commutation relations (CCRs).*
2. *The original secret state must be reconstructable from any two outputs of the dealer protocol.*
3. *The protocol must be able to use suitably-strong resource states exhibiting any form of quantum steering — even highly-asymmetric resource states. The collaborating players must be able to set up their reconstruction protocol to mix the resource mode contributions with ratio g for any $g \in (0, \sqrt{2})$.*
4. *Any noise introduced by the protocol should be equally distributed between the quadratures. There should not be more noise introduced to one quadrature of the reconstruction than the other.*

Having in mind a desire to minimise the need for expensive and noise-inducing quantum resources such as squeezers or nonlinear optics, let us add another (technically optional) condition that further narrows the options.

Definition 1. *DEALER CHANNEL CONSTRAINTS* (continued)

5. *The dealer channel should be composed only of passive optics.*

In the next subsection, we will go on to derive the optimal protocol under these constraints. An overview of the selected dealer protocol is presented in section 4.2.1.2 on page 42.

4.2.1.1 *Optimal dealer protocol*

We will in this subsection explore the class of quantum unitaries that satisfy definition 1 and so could serve as a tripartite QSS dealer protocol. Recalling that we have restricted the dealer to passive optics, the channel is fully characterised by the real matrix,

$$D = \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{pmatrix}, \tag{87}$$

for $\alpha_i, \beta_i, \gamma_i \in \mathbb{R}$. The three dealt shares are therefore described by

$$\hat{a}_1 = \alpha_1 \hat{a}_\psi + \beta_1 \hat{a}_{\mathrm{r}1} + \gamma_1 \hat{a}_{\mathrm{r}2}, \tag{88}$$

$$\hat{a}_2 = \alpha_2 \hat{a}_\psi + \beta_2 \hat{a}_{\mathrm{r}1} + \gamma_2 \hat{a}_{\mathrm{r}2}, \tag{89}$$

$$\hat{a}_3 = \alpha_3 \hat{a}_\psi + \beta_3 \hat{a}_{\mathrm{r}1} + \gamma_3 \hat{a}_{\mathrm{r}2}, \tag{90}$$

for $\hat{a}_\psi$, $\hat{a}_{ri}$ representing the secret and resource modes respectively.

CONSTRAINTS IMPOSED BY EACH RECONSTRUCTION PERMUTATION Let us now take two arbitrary shares, $\hat{a}_i$ and $\hat{a}_j$, and consider the conditions under which they can be recombined into the original secret state. The most general reconstruction channel can be written

$$\begin{pmatrix} \hat{X}_1^\pm \\ \hat{X}_2^\pm \end{pmatrix} \to T \begin{pmatrix} \hat{X}_1^\pm \\ \hat{X}_2^\pm \end{pmatrix} = \begin{pmatrix} \hat{X}_{\psi'}^\pm \\ \hat{X}_{2'}^\pm \end{pmatrix}, \tag{91}$$

for the real matrix

$$T = \begin{pmatrix} A^+ & 0 & B^+ & 0 \\ 0 & A^- & 0 & B^- \\ C^+ & 0 & D^+ & \\ 0 & C^- & 0 & D^- \end{pmatrix} \tag{92}$$

representing the reconstruction protocol, with $A^\pm, B^\pm, C^\pm, D^\pm \in \mathbb{R}$.

We now ask the question: what T produces an output ψ' most closely matching the input state? As we are interested here only in the form of the output state, we can disregard the second mode and consider only the first output from this channel given by

$$\hat{X}_{out}^\pm = (A^\pm \alpha_i + B^\pm \alpha_j)\hat{X}_\psi^\pm + (A^\pm \beta_i + B^\pm \beta_j)\hat{X}_{r1}^\pm + (A^\pm \gamma_i + B^\pm \gamma_j)\hat{X}_{r2}^\pm. \tag{93}$$

Recall now our 3rd condition: that the reconstruction be able to achieve maximal destructive interference for *any* resource state regardless of its optimal reconstruction parameter g. Reparameterising this output state such that the coefficients for $\hat{X}^{\pm}_{r1}$ and $\hat{X}^{\pm}_{r2}$ differ by a factor of $\mp g$ then gives an expression for $B^{\pm}$ in terms of $A^{\pm}$ as

$$B^{\pm} = -\frac{\gamma_i \pm g\beta_i}{\gamma_j \pm g\beta_j} A^{\pm}, \tag{94}$$

which we know must be achievable for all $g \in (0, \sqrt{2})$.

Imposing now condition 1 (that the reconstruction preserve the CCRs)[5] then specifies the form of A^{-} (as well as C^{-}, $D^{\pm}$) as

$$A^{-} = \frac{(g\beta_j - \gamma_j)(g\beta_j + \gamma_j)}{g^2(\beta_i^2 g^2 + \beta_j^2) - \gamma_i^2 - \gamma_j^2} \frac{1}{A^{+}}, \tag{95}$$

$$C^{-} = \frac{g^2\beta_i^2 - \gamma_i^2}{g^2(\beta_i^2 + \beta_j^2) - \gamma_i^2 - \gamma_j^2} \frac{1}{C^{+}}, \tag{96}$$

$$D^{+} = \frac{g\beta_j - \gamma_j}{g\beta_i - \gamma_i} C^{+}, \tag{97}$$

$$D^{-} = \frac{(g\beta_i - \gamma_i)(g\beta_j + \gamma_j)}{g^2(\beta_i^2 + \beta_j^2) - \gamma_i^2 - \gamma_j^2} \frac{1}{C^{+}}, \tag{98}$$

leaving the channel fully parameterised by A^{+} and C^{+}. We are not interested here in the state of the second output mode, so we will not consider the $C^{\pm}, D^{\pm}$ terms further.

This allows us, finally, to write the output state simply as a function of A^{+} and of the dealer parameters as

$$\hat{X}^{\pm}_{out} = \eta^{\pm}\left(\hat{X}^{\pm}_{\psi} + \lambda^{\pm}(\hat{X}^{\pm}_{r1} \mp g\hat{X}^{\pm}_{r2})\right), \tag{99}$$

for

$$\eta^{+} = \frac{\left(\alpha_i\gamma_j - \alpha_j\gamma_i + g\left(\alpha_i\beta_j - \alpha_j\beta_i\right)\right)}{\gamma_j + g\beta_j} A^{+}, \tag{100}$$

$$\eta^{-} = -\frac{(\gamma_j + g\beta_j)\left(\alpha_j\gamma_i - \alpha_i\gamma_j + g\left(\alpha_i\beta_j - \alpha_j\beta_i\right)\right)}{\gamma_i^2 + \gamma_j^2 - g^2\left(\beta_i^2 + \beta_j^2\right)} \frac{1}{A^{+}}, \tag{101}$$

$$\lambda^{\pm} = \frac{\beta_j\gamma_i - \beta_i\gamma_j}{\pm g(\alpha_j\beta_i - \alpha_i\beta_j) + \alpha_j\gamma_i - \alpha_i\gamma_j}, \tag{102}$$

and where choice of A^{+} is left free.

[5] This CCR condition is equivalent to the matrix transform preserving the symplectic form, $T\Omega T^{T} = \Omega$.

We are now in a position to consider what constraints must be imposed on the dealer to ensure this suitable reconstruction is possible. We shall primarily lean on condition 4 to help us here: that any additional noise present in the output state be equally distributed between the two quadratures. For the output state shown in equation (99), this is equivalent to the condition that $\lambda^+ = \lambda^-$ as the $\eta^\pm$ coefficients can already be equalised through judicious choice of A^+.

It is clear from the form of $\lambda^\pm$ in equation (102) that this constraint can be satisfied in two ways. It is trivially satisfied when

$$\beta_i\gamma_j - \beta_j\gamma_i = 0, \tag{103}$$

in which case $\lambda^\pm = 0$ and no additional noise is added to either quadrature. This will be the case, for example, for the {1,2}-reconstruction protocol we describe in section 4.2.2 in which a second beam splitter exactly reverses the dealer protocol.

When this added noise is unavoidable, however, the condition that it be equally distributed between quadratures is equivalent to imposing that

$$\alpha_j\beta_i - \alpha_i\beta_j = 0, \tag{104}$$

in which case the $\pm$ component of equation (102) vanishes and

$$\lambda^\pm = \frac{\beta_j\gamma_i - \beta_i\gamma_j}{\alpha_j\gamma_i - \alpha_i\gamma_j} \tag{105}$$

no longer differs between λ^+ and λ^-.

One of these two conditions must be satisfied for the original state to be reconstructable from the given two shares.

Finally, to ensure the added noise remains finite let us also impose that

$$g(\alpha_j\beta_i - \alpha_i\beta_j) + \alpha_j\gamma_i - \alpha_i\gamma_j \neq 0. \tag{106}$$

DEALER PROTOCOLS SATISFYING THESE CONSTRAINTS To convert these individual constraints into a set of permitted dealer protocols, let us now recall condition 2: that reconstruction be possible for *any* permutation of output modes. These conditions must therefore be satisfied for *every*

combination of $i, j \in \{1, 2, 3\}$ and the dealer channel must simultaneously satisfy

$$\beta_1\gamma_2 = \beta_2\gamma_1 \quad \text{or} \quad \alpha_1\beta_2 = \alpha_2\beta_1, \tag{107}$$

$$\beta_1\gamma_3 = \beta_3\gamma_1 \quad \text{or} \quad \alpha_1\beta_3 = \alpha_3\beta_1, \tag{108}$$

$$\beta_2\gamma_3 = \beta_3\gamma_2 \quad \text{or} \quad \alpha_2\beta_3 = \alpha_3\beta_2, \tag{109}$$

as well as the condition that

$$\alpha_j\beta_i - \alpha_i\beta_j \quad \text{and} \quad \alpha_j\gamma_i - \alpha_i\gamma_j \tag{110}$$

not both be simultaneously 0 for any (i, j) combination.

The only CCR-preserving channel satisfying this set of conditions (down to mode permutation) is described by

$$\begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \end{pmatrix} = \begin{pmatrix} \xi & \sqrt{1-\xi^2} & 0 \\ \sqrt{1-\xi^2} & -\xi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{a}_\psi \\ \hat{a}_{\mathrm{r}1} \\ \hat{a}_{\mathrm{r}2} \end{pmatrix}, \tag{111}$$

for any $\xi \in [0, 1]$, defining a class of potential tripartite state sharing protocols.

Any protocol from this class would be usable for tripartite QSS, but there is a clear benefit to selecting the symmetric $\xi = 1/\sqrt{2}$ channel. As ξ moves away from $1/\sqrt{2}$ the more of the original secret state is necessarily contained within a single share. In addition to potentially leaving more information exposed to an adversary who gained access to that share, this would result in a worse-quality reconstruction when that share was not available. By selecting $\xi = 1/\sqrt{2}$, then, we optimise for the worst-case scenario — maximising the quality of the reconstruction in the least effective of the three reconstruction permutations and minimising the potential information leakage should one share be intercepted.

4.2.1.2 *Overview of dealer protocol*

Let us now for clarity restate the dealer protocol. The dealer is passed a secret state ψ whose class may be public knowledge[6] but the specifics of which are unknown to the dealer. The dealer then constructs three shares by interfering this secret state on a balanced beam splitter with one mode of the two-mode entangled resource, as outlined in figure 5.

[6]For example, it may be known that the secret state is a coherent state, or a Fock state

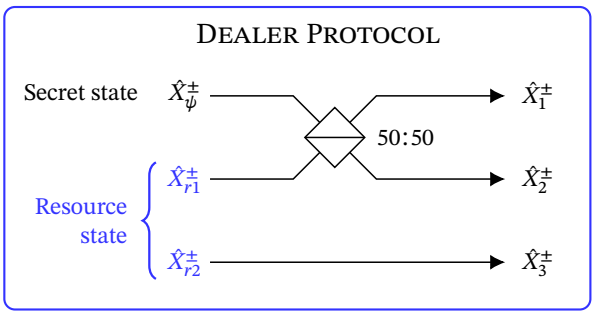


Figure 5: The selected dealer protocol for {2, 3}-threshold quantum state sharing. Three modes are produce by mixing the secret state with one mode of an entangled resource state at a balanced beamsplitter.

The three output modes are related to the input, $\hat{X}^{\pm}_{\psi}$, and resource, $\hat{X}^{\pm}_{ri}$, modes by the quadrature relation

$$\hat{X}^{\pm}_{1} = \frac{1}{\sqrt{2}}(\hat{X}^{\pm}_{\psi} + \hat{X}^{\pm}_{r1}), \tag{112}$$

$$\hat{X}^{\pm}_{2} = \frac{1}{\sqrt{2}}(\hat{X}^{\pm}_{\psi} - \hat{X}^{\pm}_{r1}), \tag{113}$$

$$\hat{X}^{\pm}_{3} = \hat{X}^{\pm}_{r2}. \tag{114}$$

Crucially, the relatively large single-mode variance of $\hat{X}^{\pm}_{r1}$ obscures the secret state so none of $\hat{X}^{\pm}_{1,2,3}$ individually contain enough information to accurately reconstruct it. Each of the three modes are then distributed to a player as labelled.

To improve the security of each individual share further, anti-correlated displacement operations could be performed on each share, with displacements drawn classically from a defined probability distribution, $\delta\mathcal{N}$. This will transform the shares as

$$\hat{X}^{\pm}_{1} = \frac{1}{\sqrt{2}}(\hat{X}^{\pm}_{\psi} + \hat{X}^{\pm}_{r1}) + \delta\mathcal{N}^{\pm}, \tag{115}$$

$$\hat{X}^{\pm}_{2} = \frac{1}{\sqrt{2}}(\hat{X}^{\pm}_{\psi} - \hat{X}^{\pm}_{r1}) - \delta\mathcal{N}^{\pm}, \tag{116}$$

$$\hat{X}^{\pm}_{3} = \hat{X}^{\pm}_{r2} \pm \delta\mathcal{N}^{\pm}. \tag{117}$$

When the shares are combined through the methods outlined in the following subsections this noise automatically falls out producing no-worse a reconstruction than in the zero-noise case. We focus exclusively here on security that can be guaranteed by the no-cloning theorem, which classical noise is not able to provide,[7] so we do not discuss this possibility further

in this thesis. The interested reader is directed to the discussion of the addition of classical noise in Ref. [48].

4.2.2 *Secret state reconstruction using shares 1 & 2*

Having established the process by which the information describing the secret quantum state is distributed between players, let us now consider their ability to retrieve this secret state. The three output shares the dealer produces are not created equally[8] — shares 1 and 2 each contain a 50% contribution from the secret state while the third share contains none at all! This asymmetry in shares means a different reconstruction sub-protocols are required for different combinations of shares, which produce reconstructions of varying quality.

The best possible reconstruction is obtained when the two shares that directly contain information about ψ are available. With access to both shares 1 & 2, the collaborating players can trivially reverse the dealer protocol by passing them through a second balanced beam splitter, leaving in one beam-splitter output the state described by

$$\hat{X}^{\pm}_{\text{out}} = \frac{1}{\sqrt{2}}\left(\hat{X}^{\pm}_{1} + \hat{X}^{\pm}_{2}\right) = \hat{X}^{\pm}_{\psi}. \tag{118}$$

In the ideal case with no transmission or component losses, this will revert the system to its original separable state. With ψ completely separated from the resource modes it is reconstructed perfectly regardless of the resource state used, even in the wholly classical case in which no entanglement is present.

This trivial form of reconstruction represents the exceptional, best-case scenario however, and is only possible for one of the three permutations of collaborating players.

[7] As an example, the security we analyse in section 4.3 is based on the physically-limited availability of quantum information so is secure against an eavesdropper within the dealer protocol itself. Security reliant on the classical noise would not provide this, as an eavesdropper in the classical-noise-generation step would be able to obtain perfect knowledge of this noise and later remove it from their share. Classical noise is therefore a good additional step to reduce the information available to eavesdroppers *outside* the dealer protocol but should not be relied upon for the core security analysis.

[8] Indeed, returning to the allowed class of dealer protocols shown in equation (111), we can see that there is no QSS protocol under definition 1 which allows the secret state information to be distributed equally.

4.2.3 *Secret state reconstruction using shares 1 & 3 or 2 & 3*

When the first and second shares are not both available, though, the secret cannot be directly reconstructed through classical optics alone. Instead, a more complex process is required to disentangle ψ from the resource state contribution. We described the sub-protocol which can achieve this here.

Recalling that the resource state is entangled such that the modes cancel maximally when mixed with ratio $\hat{X}^{\pm}_{r1} \mp g\hat{X}^{\pm}_{r2}$ for some $g \in \mathbb{R}$, it is clear that the optimal output state will take the general form

$$\begin{aligned} \hat{X}^{\pm}_{\text{out},\{1,3\}} &= \eta\left[\hat{X}^{\pm}_{\psi} + (\hat{X}^{\pm}_{r1} \mp g\hat{X}^{\pm}_{r2})\right] \qquad (119)\\ &= \eta\left[\sqrt{2}\hat{X}^{\pm}_{1} \mp g\hat{X}^{\pm}_{3}\right], \\ \hat{X}^{\pm}_{\text{out}\{2,3\}} &= \eta\left[\hat{X}^{\pm}_{\psi} - (\hat{X}^{\pm}_{r1} \mp g\hat{X}^{\pm}_{r2})\right] \qquad (120)\\ &= \eta\left[\sqrt{2}\hat{X}^{\pm}_{1} \pm g\hat{X}^{\pm}_{2}\right], \end{aligned}$$

for some amplification factor $\eta \in \mathbb{R}$. For a Gaussian input, this produces an output state with mean vector $\bar{\boldsymbol{r}} = \eta\bar{\boldsymbol{r}}_{\psi}$ and covariance matrix $V = \eta^2 V_{\psi} + \eta^2 E_{1|2}(g) I_2$: an amplified, generally-noisy copy of the input state.

We will show in the next subsection that for this transformation to represent a physical channel, it must impose an amplification of

$$\eta = \frac{1}{\sqrt{2 - g^2}} \qquad (121)$$

on the output. The reconstruction channel is entirely characterised by the collaborating players free choice of g, which we will henceforth term the *reconstruction parameter*. Notably, for $g \neq 1$, this amplification is not unity and the protocol distorts the original state. Selecting the reconstruction protocol for a specific resource state is not simply a matter of choosing that g that maximises its entanglement. The optimal reconstruction parameter g will depend on the specific setup, the resource state used and the envelope of input states for which the protocol is expected to be used. Consequently, in all our analysis here we will focus on the impact of this parameter in the abstract, and not on any specific features of the entanglement structure within the resource states.

4.2.3.1 *Derivation of reconstruction amplification, η*

Before we go on to discuss the mechanisms we will use to correct for the amplification this reconstruction introduces, let us prove the channel does

indeed impart the amplification we claimed it would in equation (121). The {1,3} and {2,3} reconstruction sub-protocols follow the same general process, with only minor parameter changes to account for the phase difference between the $\hat{X}^{\pm}_{r1}$ contributions in equations (112) and (113). They each therefore impose the same amplification factor η for any given reconstruction parameter g and so we present the derivation only for {1,3} reconstruction — the {2,3} case can be derived by following the same process.

The reconstruction channel described by equation (119) can be represented as the symplectic matrix $T \in \mathrm{Sp}(4;\mathbb{R})$, given by

$$T = \begin{pmatrix} \sqrt{2}\eta & 0 & -g\eta & 0 \\ 0 & \sqrt{2}\eta & 0 & g\eta \\ a & 0 & c & 0 \\ 0 & b & 0 & d \end{pmatrix}, \tag{122}$$

where $a, b, c, d \in \mathbb{R}$ represent the form of the second (unused) output mode, acting on a Gaussian input state such that

$$\bar{\boldsymbol{r}} \to T\bar{\boldsymbol{r}}, \qquad V \to TVT^{\mathrm{T}}. \tag{123}$$

For this matrix to represent a physical quantum unitary, it must be symplectic and so must preserve the symplectic form Ω described in section 2.1.4 as

$$T\Omega T^{\mathrm{T}} = \Omega. \tag{124}$$

This condition is satisfied for T only if

$$\sqrt{2}b - dg = 0, \tag{125}$$

$$\sqrt{2}a + cg = 0, \tag{126}$$

$$ab + cd = 1, \tag{127}$$

$$\eta^2(2 - g^2) = 1 \tag{128}$$

are all simultaneously satisfied.

The first three of these conditions define the second output of the channel as

$$b = \frac{g^2}{a(g^2-2)}, \tag{129}$$

$$c = -\frac{\sqrt{2}}{g}a, \tag{130}$$

$$d = \frac{\sqrt{2}g}{a(g^2-2)}, \tag{131}$$

for $a \in \mathbb{R}$ representing only the local form of the output.

It is the final condition that gives the amplification the channel must impart on the secret state as

$$\eta = \frac{1}{\sqrt{2-g^2}}. \tag{132}$$

4.2.4 *Correcting for reconstruction amplification*

As we have seen, when the {1,3} or {2,3} reconstruction protocol is implemented for any $g \neq 1$ the output state is necessarily a (de)amplified copy of the input state. In the case of a Gaussian secret, the mean vector characterising the output state will be a scaled copy of the input mean as $\bar{\boldsymbol{r}}_{\text{out}} = \eta \bar{\boldsymbol{r}}_{\text{in}}$. To ensure the secret state is reproduced accurately, we augment the protocol with an additional pre-amplification or post-attenuation step that corrects for this (de)amplification. These corrections are similar to those used by He *et al.* [21] in their study of quantum teleportation, so we borrow their descriptions here and term the two cases *late-stage attenuation (lsatt)* and *early-stage amplification (esa)* QSS.

As the effect of a (de)amplification is proportional to the magnitude of the secret state's mean vector, and as the amplification correction is sometimes noise-increasing (when $g < 1$), this may not always result in an improvement in output quality. Indeed, if one were certain the protocol would only be used for very low amplitude states it may be preferable to leave a mild de-amplification uncorrected for. However, to analyse the protocol's effectiveness across the spectrum of input states, we will analyse here primarily the corrected version of the protocol; although we will additionally consider the uncorrected output when we go on to discuss the sharing of Fock states in section 7.1.

4.2.4.1 *Late-stage attenuation (lsatt)*

When the output state would otherwise be an amplified copy of the input state (when $g > 1$, so $\eta > 1$), the optimal correction is simple: the output state is passed through a beam splitter with a vacuum environment to reduce its amplitude. For an ideal beam splitter with transmissivity selected such that $\tau = 1/\eta^2$, the state is transformed as

$$\hat{X}^{\pm}_{\text{out}} \to \frac{1}{\eta}\hat{X}^{\pm}_{\text{out}} + \sqrt{1 - \frac{1}{\eta^2}}\hat{X}^{\pm}_{\text{vac}}, \tag{133}$$

leaving the output of the protocol in the form

$$\hat{X}^{\pm}_{\text{out}} = \hat{X}^{\pm}_{\psi} + \hat{X}^{\pm}_{r1} \mp g\hat{X}^{\pm}_{r2} + (g^2 - 1)\hat{X}^{\pm}_{\text{vac}}. \tag{134}$$

As this correction both brings the mean vector closer to the input state and reduces the variance of the output (as the vacuum is minimum-uncertainty) this is strictly a fidelity-improving correction for a coherent-state secret.[9]

4.2.4.2 *Early-stage amplification (esa)*

When the output state is instead a de-amplified copy of the input state (when $g < 1$ so $\eta < 1$), we take a slightly different approach. Consider the form of the output state for $\eta < 1$,

$$\hat{X}^{\pm}_{\text{out},\{1,3\}} = \eta\left[\hat{X}^{\pm}_{\psi} + (\hat{X}^{\pm}_{r1} \mp g\hat{X}^{\pm}_{r2})\right]. \tag{135}$$

In this form, both the secret state contribution and resource contributions have been deamplified equally. If one were to naïvely amplify this output, one would not only amplify the desired ψ contribution but also the residue resource-state variance.

Instead, as the properties of the resource state (and therefore what reconstruction parameter g will be used) are considered public knowledge and are known in advance, the secret state can be amplified prior to passing it to the dealer.[10] In this way, only the desired information is amplified — any noise picked up during the QSS protocol remains in its de-amplified form.

We model this process as an ideal amplifying channel; in practice such an amplification could be achieved by a phase-insensitive amplifier [24].

[9] This is not more-generally true, notably for the Fock-state secrets we consider in section 7.1. In that case, the inclusion of vacuum noise degrades the state further.

[10] We consider here the amplification as a separate correction made by the state-owner before passing the state to the dealer. In a practical implementation this could also be absorbed into the dealer protocol.

Denoting the original secret state as ψ, the amplified input to the QSS protocol will become

$$\hat{X}^{\pm}_{\text{in}} = \frac{1}{\eta}\hat{X}^{\pm}_{\psi} + \sqrt{\frac{1}{\eta^2} - 1}\hat{X}^{\pm}_{\text{vac}}, \tag{136}$$

where $1/\eta > 1$. This amplification will then be undone by the reconstruction protocol, ultimately leaving the reconstructed output state in the form

$$\hat{X}^{\pm}_{\text{out}} = \hat{X}^{\pm}_{\psi} + \eta(\hat{X}^{\pm}_{r1} \mp g\hat{X}^{\pm}_{r2}) + \sqrt{1-\eta^2}\hat{X}^{\pm}_{\text{vac}}, \tag{137}$$

in which the secret state is reproduced with unity-amplification while the resource contributions are de-amplified by $\eta < 1$.

Note, though, that unlike the *lsatt* correction, pre-amplifying the secret state is not side-effect-free. First, the amplification step will introduce additional noise into the secret state, which for coherent states with very small mean amplitude may be more destructive than the de-amplification.

Second, as the amplification occurs prior to the dealer protocol it will additionally impact the $\{1, 2\}$ reconstruction sub-protocol, which will now reconstruct an amplified copy of the state. This will require a further attenuation step to correct for the amplification. However, the output of an ideal attenuation is necessarily better quality than *any other process producing a de-amplified state* so this output will remain strictly better than the {1,3} or {2,3} reconstructions. Introducing this pre-amplification step will not, then, impact our analysis of the worst-case effectiveness of the protocol or its security which we base entirely on the {1,3}/{2,3} protocols.

4.3 SECURE STATE RECONSTRUCTION

Throughout this part of the thesis, we will frequently refer to this protocol as providing 'guaranteed security', due to its basis in the immutable laws of physics. Indeed, one of the major questions we look to answer is: under what conditions can we be certain this protocol is secure? Let us pause for a moment here, then, to formalise what we take security to mean for a quantum communication protocol.

Definition 2. *A quantum communication protocol is considered secure when the authorised players can guarantee they have access to more information about the original state than is accessible to any adversary.*

This definition of security is useful in its simplicity and ability to provide a direct comparison between quantum protocols. In practice, though, simply requiring any better-quality reconstruction than an adversary may not

provide sufficient security — for example this allows security when the adversary has access to only marginally-less information than the authorised players! A stronger security condition in which the information available to an adversary is guaranteed to be below some value ϵ may be desired instead, as is used in the cryptography field [51, 52]. Such a definition of security could be derived simply by adjusting the fidelity threshold used — a higher reconstruction fidelity requirement will automatically impose that less information be accessible to the adversary. In keeping with the quantum communication field [45, 53, 54], in this thesis we will focus on the security definition of definition 2.[11]

An obvious, if somewhat tedious, approach to answering this question might be to collate the information available to any adversaries and directly compare it against the secret state reconstructed by the collaborating players. In the perfectly ideal case — in which the only information available to an adversary consists of the sole lost share — this approach might be fairly simple. When one moves away from this ideal scenario, though, and begins to discuss noise and loss the situation becomes more complex. Even in the event the collaborating players have access to an accurate accounting of the loss, they have no way to distinguish true loss from malicious eavesdropping. To be confident of security, then, *all* sources of loss would have to be quantified as information obtainable by an adversary and added to the calculation of information available to them — dramatically increasing the work necessary to guarantee security.

Fortunately, a much simpler method to certify this form of security can be derived from the uncertainty theorem. A natural consequence of the limits imposed on the amount of information that can exist about a quantum state is that quantum states cannot be cloned [55]; if cloning were possible, one could obtain more information than is allowed by making complementary measurements on multiple copies. What is possible, though, is to create imperfect clones of a single quantum state that do not collective contain more information than the original. The limits on these copies has been studied extensively for a variety of classes of input state [54–58]. For example, Grosshans and Grangier found in 2001 [54] that, assuming no prior knowledge of the range of coherent amplitudes, the maximum cloning fidelity that could be achieved for an unknown coherent state is $\mathcal{F} = 2/3$.

These same limits apply immediately to any quantum communication protocol. One could consider our protocol to be a (somewhat over-

[11] It is common in the cryptography field [52] to refer to both security and to 'correctness' as measures of how much information is accessible to adversaries and authorised players respectively. The latter is in this thesis represented by the reconstruction fidelity.

complicated) cloning machine with two outputs: one 'clone' constructed by the collaborating players, and a second 'clone' representing the information available to an adversary. Should the collaborating players achieve a reconstruction fidelity above the maximum shared cloning fidelity, the no-cloning theorem requires an equivalent reduction in the quality of the adversary's copy. By exceeding this threshold, then, the collaborating players can have full confidence that they possess the best available copy of the state regardless of how many sources of information their adversary has access to.

In general, we assume that the class of quantum state shared is public knowledge. An adversary who knows whether the state is coherent or thermal, for example, is able to tailor their attacks accordingly and potentially gain more information. To be safe, we must assume the adversaries are familiar with the state generation process and so are privy to this information. By using no-cloning limits specific to the type of state, this potential adversary-knowledge is built-in to our security analysis.

In most discussions of quantum communication security, it is assumed no other prior knowledge is available about the distribution of input states [45, 59]. This general view of security allows us to analyse the resource requirements for the protocol in its idealised form, and enables a direct comparison between protocols. This does not always represent a physically feasible scenario, however. For example, the coherent amplitude is effectively limited by available energy and equipment; by accepting any coherent input state we are allowing for potential inputs well beyond what can be readily generated.

A complete discussion of security, then, will require a knowledge or estimation of the probability distribution function $P(\psi)$ describing the set of input states. Any reduction in the domain of the input states corresponds to an increase in information classically available about the state and thus in the ability to clone them. Consider, as an extreme example, the case in which the input distribution consists of only one state with probability $P = 1$ — clearly this state can always be cloned perfectly simply by generating new ones to the known specification! When input states are drawn from a limited set, we account for the adversary using their knowledge of the input distribution by certifying security based on the *average* reconstruction fidelity, weighted by the probability distribution, exceeding the *average* cloning fidelity [60].

We perform this fuller security analysis under the assumption that our coherent state inputs are described by a Gaussian distribution in section 5.2.5 and in our consideration of the Fock superpositions in section 7.2. When we

consider coherent states drawn equally from the full spectrum of coherent-amplitudes in section 5.2.4, however, we need not consider this average as it is assumed there is no information available about which states are more likely to be shared.

Let us also pause at this moment to consider the conditions under which we assume the protocol to operate when considering security. The most notable of these is the question of trust: the dealer is handed an unobscured copy of the original state and therefore has access to the information and therefore must be trusted to implement the protocol properly. For simplicity, in this thesis we consider the dealer to be fully trustworthy and consider only attacks from dishonest players. However, we note also that the no-cloning theorem provides a method through which the dealer's trustworthiness is auditable: any siphoning of information from the dealer would impact the quality of the reconstruction. By sending and reconstructing decoy states, the players are therefore able to test the dealer's implementation and detect dishonesty at that stage.

5

GAUSSIAN QUANTUM STATE SHARING

Having established our quantum state sharing protocol in chapter 4, let us now start to consider its performance. As this protocol is designed to exploit continuous-variable Gaussian entanglement, it is natural that the first question we ask be: how well does it work for Gaussian secret states?

Gaussian states are those whose Wigner function takes the form of a Gaussian

$$W(\boldsymbol{q}) = \frac{1}{\pi^n \sqrt{\det V}} \exp\left[-(\boldsymbol{q}-\bar{\boldsymbol{r}})^T V^{-1} (\boldsymbol{q}-\bar{\boldsymbol{r}})\right], \tag{138}$$

and so are fully characterised by its mean vector $\bar{\boldsymbol{r}} \in \mathbb{R}^{2n}$ and covariance matrix $V \in \mathbb{R}^{2n\times 2n}$ for n modes. We will primarily explore in this chapter what impact the QSS protocol has on these two properties.

Initially, in section 5.1, we will leave the exact form of the state unspecified and take a more general view of the effect our protocol has on Gaussian states. In section 5.2 we will then consider specifically its use for the state of primary interest in Gaussian quantum information: the coherent state. In this section, we will see that the protocol outperforms teleportation for equivalent resources, and that QSS can offer guaranteed security when provided with *any* steerable resource state. We will then expand the input domain in section 5.3 to include first coherent states exhibiting quadrature-squeezing before looking at the completely general single-mode Gaussian state. We will see that the protocol remains effective for any such input state given enough entanglement resources, and that it can provide provable security for any *pure* input. Finally, in section 5.4, we will briefly consider the potential for this protocol to be used to share multi-mode Gaussian states.

5.1 GENERAL GAUSSIAN STATE OUTPUT

Let us begin by considering the form of the modes at each point in the process. As the protocol is defined to be mean-preserving, we will be primarily interested here in its effect on the covariance matrix.

5.1.1 *Dealer protocol*

The three shares constructed by the dealer are given by equations (112) to (114) as

$$\hat{X}_1^{\pm} = \frac{1}{\sqrt{2}}(\hat{X}_{\psi}^{\pm} + \hat{X}_{r1}^{\pm}), \tag{139}$$

$$\hat{X}_2^{\pm} = \frac{1}{\sqrt{2}}(\hat{X}_{\psi}^{\pm} - \hat{X}_{r1}^{\pm}), \tag{140}$$

$$\hat{X}_3^{\pm} = \hat{X}_{r2}^{\pm}. \tag{141}$$

As the initial secret state ψ is wholly separable from the resource state, we can directly read-off the single-mode variances for each of these shares as

$$V_1 = \frac{1}{2}(V_{\psi} + V_{r1}), \tag{142}$$

$$V_2 = \frac{1}{2}(V_{\psi} + V_{r1}), \tag{143}$$

$$V_3 = V_{r2}, \tag{144}$$

for V_{ri} representing the single-mode variance of mode i of the resource state.

This intermediate result demonstrates the basis on which quantum state sharing works. For entanglement to be present within a state, it must be accompanied by an increase in the uncertainty of a measurement on one mode alone. By mixing the secret state with only one mode of the resource state, then, the details of the secret state are drowned out by this comparatively large variance; the information obtainable from a single share alone is shown in figure 6. The secret state only becomes accessible when the entanglement with the second resource mode in share 3 is exploited to remove the resource contribution to mode 1 or 2.

5.1.2 *{1,2} reconstruction protocol*

Let us now turn our attention to the potential for these shares to be recombined into a copy of the original state. Recall from section 4.2.2 that when the first two modes are available the dealer protocol can be trivially reversed, perfectly reproducing its input. When the protocol is implemented for $g \geq 1$ — when the input to the dealer protocol is simply the secret state — the original state is universally reconstructed with ideal fidelity $\mathcal{F} = 1$. When the protocol is used for $g < 1$, however, it is the pre-amplified copy of the secret state that is reconstructed perfectly and so a subsequent

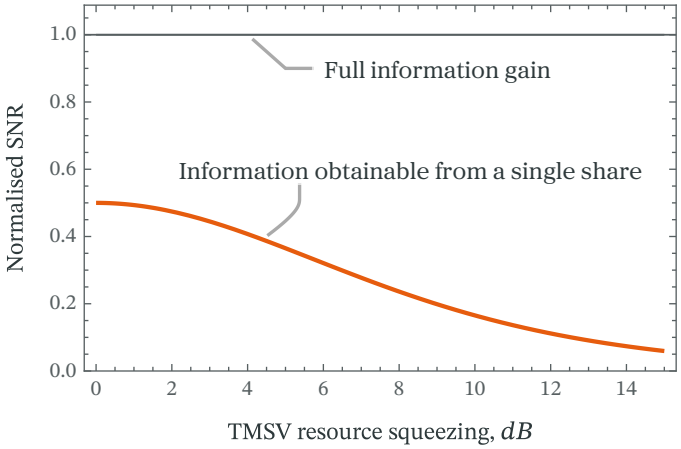


Figure 6: The information available from each quadrature of share 1 or 2, quantified by the normalised signal-to-noise ratio given by $\mathrm{SNR}_{\mathrm{out}}/\mathrm{SNR}_{\mathrm{in}}$. [1] A value of 1 would indicate full information accessibility, while the metric tends to 0 for a fully uncertain state.

de-amplification is required. Each of these (de)amplifications introduce additional noise, so the final output will be a generally-noisy copy of the original state given by

$$\hat{X}^{\pm}_{\mathrm{out}} = \hat{X}^{\pm}_{\psi} + \sqrt{1-\eta^2}\,(\hat{X}^{\pm}_{\mathrm{env1}} + \hat{X}^{\pm}_{\mathrm{env2}}), \tag{145}$$

where $\hat{X}^{\pm}_{\mathrm{evn1}}$ and $\hat{X}^{\pm}_{\mathrm{env2}}$ are distinct, uncorrelated environment modes. This output will be characterised by the original state's covariance matrix with added noise, as

$$\bar{\boldsymbol{r}}_{\mathrm{out}} = \bar{\boldsymbol{r}}_{\psi}, \tag{146}$$

$$V_{\mathrm{out}} = V_{\psi} + 2(1-\eta^2)V_{\mathrm{env}}, \tag{147}$$

for V_{env} representing environment noise; ordinarily $V_{\mathrm{env}} = I_2$ for a vacuum (ideal) environment. As environment noise is assumed to be purely random, with no bias towards a given average, it has no impact on the mean vector.

The output from an ideal attenuation must, by definition, be at least as good as *any other process resulting in a de-amplified state*, though, and so it is immediately true that even after this additional de-amplification step, the {1,2} output will always be of a higher quality than the {1,3} reconstruction. [2] We are primarily interested in the worst-case reconstruction, and so we will not consider this reconstruction much in our discussion of the protocol's effectiveness in sections 5.2 and 5.3.

[1] *SNR* here represents the signal-to-noise ratio, a measure of how accessible information is within the state that quantifies the extent to which the signal is drowned out by noise.

[2] That this is indeed the case is not immediately obvious from comparing the form of equation (147) to the output covariance matrix we will go on to find for {1,3} reconstruc-

5.1.3 *{1,3} and {2, 3} reconstruction protocols*

Let us now consider the altogether more interesting cases of {1,3} and {2,3} reconstruction, in which entanglement between the resource modes is exploited to disentangle ψ from one of the shares. The quality of these reconstructions consequently rests on the quality of the entanglement in the resource state. As perfect entanglement is not obtainable for continuous-variable states, this reconstruction will always be imperfect, including in the ideal noiseless case. In the remainder of this chapter we will consider the impact the chosen resource state has on this reconstruction.

Recall from section 4.2.4 that the specific form of the output state will depend on whether a post-attenuation or pre-amplification step is necessary. We can combine equations (133) and (136) to write the output as

$$\hat{X}^{\pm}_{\text{out}} = \begin{cases} \hat{X}^{\pm}_{\psi} + \eta(\hat{X}^{\pm}_{r1} \mp g\hat{X}^{\pm}_{r2}) + \sqrt{1-\eta^2}\hat{X}^{\pm}_{\text{vac}} & g < 1 \quad (\eta < 1) \\ \hat{X}^{\pm}_{\psi} + \hat{X}^{\pm}_{r1} \mp \hat{X}^{\pm}_{r2} & g = 1 \\ \hat{X}^{\pm}_{\psi} + \hat{X}^{\pm}_{r1} \mp g\hat{X}^{\pm}_{r2} + \sqrt{1-\frac{1}{\eta^2}}\hat{X}^{\pm}_{\text{vac}} & 1 < g < \sqrt{2} \quad (\eta > 1), \end{cases} \tag{149}$$

where $\eta = 1/\sqrt{2-g^2}$. Noting that the resource state can always be designed to be zero-mean, we can see already that the mean vector describing the secret state is perfectly reproduced in every case.

Let us consider, then, how noisy this copy of the input state is. As there is now entanglement between the $\hat{X}^{\pm}_{r1}$ and $\hat{X}^{\pm}_{r2}$ modes, we can no longer simply read off the variance of our output state as the sum of the variances of its components. Let us instead return to the definition of the single-mode quadrature variance given in equation (25) as

$$v^{\pm}_{i,i} = \frac{1}{2}\left\langle \{\Delta\hat{X}^{\pm}_i ; \Delta\hat{X}^{\pm}_i\} \right\rangle, \tag{150}$$

for $\{A;B\}$ the anti-commutator of A and B and $\Delta\hat{X}^{\pm}_i = \hat{X}^{\pm}_i - \langle\hat{X}^{\pm}_i\rangle$.

---

tion in equation (154). However, it is true for the physically-allowed range of steering parameters,

$$E_{1|2}(g) \geq |1-g^2| \tag{148}$$

imposed by the uncertainty theorem, recalling that $\eta \geq 1/\sqrt{2}$ (as $g \geq 0$).

This steering limit is also the reason that we need only consider values up to the limit of $g < \sqrt{2}$ — steering cannot exist outside this range!

Applying this formula to equation (149) the single-mode variance of the output can be found as

$$v_{\text{out}}^{\pm} = \frac{1}{2}\left\langle\left\{\Delta\hat{X}_{\text{out}}^{\pm}\,;\,\Delta\hat{X}_{\text{out}}^{\pm}\right\}\right\rangle \tag{151}$$

$$= \begin{cases} \frac{1}{2}\left\langle\left\{\Delta\hat{X}_{\psi}^{\pm}\,;\,\Delta\hat{X}_{\psi}^{\pm}\right\}\right\rangle + \eta^2\frac{1}{2}\left\langle\left\{\Delta\hat{X}_{r1}^{\pm} \mp g\Delta\hat{X}_{r2}^{\pm}\,;\,\Delta\hat{X}_{r1}^{\pm} \mp g\Delta\hat{X}_{r2}^{\pm}\right\}\right\rangle & \\ \qquad + (1-\eta^2)\frac{1}{2}\left\langle\left\{\Delta\hat{X}_{\text{vac}}^{\pm}\,;\,\Delta\hat{X}_{\text{vac}}^{\pm}\right\}\right\rangle & g \le 1 \\ \frac{1}{2}\left\langle\left\{\Delta\hat{X}_{\psi}^{\pm}\,;\,\Delta\hat{X}_{\psi}^{\pm}\right\}\right\rangle + \frac{1}{2}\left\langle\left\{\Delta\hat{X}_{r1}^{\pm} \mp g\Delta\hat{X}_{r2}^{\pm}\,;\,\Delta\hat{X}_{r1}^{\pm} \mp g\Delta\hat{X}_{r2}^{\pm}\right\}\right\rangle & \\ \qquad + (1-\frac{1}{\eta^2})\frac{1}{2}\left\langle\left\{\Delta\hat{X}_{\text{vac}}^{\pm}\,;\,\Delta\hat{X}_{\text{vac}}^{\pm}\right\}\right\rangle & g \ge 1, \end{cases} \tag{152}$$

for

$$\frac{1}{2}\left\langle\left\{\Delta\hat{X}_{r1}^{\pm} \mp g\Delta\hat{X}_{r2}^{\pm}\,;\,\Delta\hat{X}_{r1}^{\pm} \mp g\Delta\hat{X}_{r2}^{\pm}\right\}\right\rangle \coloneqq E_{1|2}(g) \tag{153}$$

the steering parameter describing the resource. The covariance matrix describing this output state is therefore

$$V_{\text{out}} = \begin{cases} V_{\psi} + \eta^2 E_{1|2}(g) I_2 + (1-\eta^2)V_{\text{vac}} & g < 1 \quad (\eta < 1) \\ V_{\psi} + E_{1|2}(g) I_2 & g = 1 \\ V_{\psi} + E_{1|2}(g) I_2 + (1-\frac{1}{\eta^2})V_{\text{vac}} & 1 < g < \sqrt{2} \quad (\eta > 1), \end{cases} \tag{154}$$

again for $\eta^2 = 1/(2-g^2)$ and where $V_{\text{vac}} = I_2$.

In the special case in which $g = 1$ and no amplification correction is required, the covariance matrix describing the output can be written simply as

$$V_{\text{out}} = V_{\psi} + E_{1|2}(g=1) I_2 = \begin{pmatrix} v_{\psi} + E_{1|2}(g=1) & 0 \\ 0 & v_{\psi} + E_{1|2}(g=1) \end{pmatrix}. \tag{155}$$

We can begin to see here the impact resource state entanglement has on the reconstructed state. The base output (at $g = 1$) essentially consists of the original secret state plus some residue of the resource state corresponding to its imperfect entanglement. As the entanglement strength increases this residue will reduce in size, and in the limit of perfect entanglement disappear.

One might think, then, that the optimal approach would be to select the reconstruction parameter g that maximises the entanglement in the resource state so minimising this residue. However, we have also seen that moving away from $g = 1$ necessitates an amplification correction, adding additional noise to the output. The optimal reconstruction parameter will

therefore be a tradeoff between maximising usage of the entanglement and minimising the distance from $g = 1$.

5.2 QUANTUM STATE SHARING OF COHERENT STATES

The core information carrier in continuous-variable quantum communication, and in much of classical communication, is the coherent state: the unsqueezed Gaussian state that saturates the uncertainty limit.

Coherent states have a defined covariance matrix $V_{\text{coherent}} = I_2$ and are entirely characterised by their mean vector, $\bar{\boldsymbol{r}} \in \mathbb{R}$, which can be modulated to encode information. Quantum communication protocols involving coherent states aim to preserve this mean vector while introducing the minimum possible noise.

Our quantum state sharing protocol has been designed to preserve the mean in every case so this first aim is automatically met. However, with the limited exception of {1,2} reconstruction, the protocol necessarily adds noise stemming from the imperfect nature of continuous-variable entanglement. The reconstructed output state will therefore no longer be a minimum-uncertainty coherent state, instead taking the form of a thermal state with the same mean. In assessing the effectiveness of this protocol, a natural first question to ask would be: how thermal is this output state?

5.2.1 *Output state purity*

We can already answer this question simply by reading off the elements from the covariance matrix in equation (154), which for coherent-state inputs reduces to

$$V_{\text{out}} = \begin{cases} (2 + \eta^2 E_{1|2}(g) - \eta^2) I_2 & g < 1 \quad (\eta < 1) \\ (1 + E_{1|2}(g)) I_2 & g = 1 \\ (2 + E_{1|2}(g) - \frac{1}{\eta^2}) I_2 & 1 < g < \sqrt{2} \quad (\eta > 1). \end{cases} \tag{156}$$

By comparing to the known form of a thermal state, given in section 2.2.1, we can immediately say that this output state will have average thermal photon number

$$\bar{n} = \frac{1}{2} \begin{cases} 1 + \eta^2 E_{1|2}(g) - \eta^2 & g < 1 \quad (\eta < 1) \\ E_{1|2}(g) & g = 1 \\ 1 + E_{1|2}(g) - \frac{1}{\eta^2} & 1 < g < \sqrt{2} \quad (\eta > 1). \end{cases} \tag{157}$$

This result immediately translates into another relevant measure of the reconstruction 'thermality': the purity of this output state, defined as the trace of the square of the density matrix, $\mathcal{P}(\hat{\rho}) = \mathrm{Tr}\,\hat{\rho}^2$. For a thermal state, this measure reduces simply to the determinant of the covariance matrix as

$$\mathcal{P}(\hat{\rho}_{\mathrm{out}}) = \frac{1}{\sqrt{\det V_{\mathrm{out}}}} = \begin{cases} 1/(2+\eta^2 E_{1|2}(g)-\eta^2) & g<1 \quad (\eta<1) \\ 1/(1+E_{1|2}(g)) & g=1 \\ 1/(2+E_{1|2}(g)-\frac{1}{\eta^2}) & 1<g<\sqrt{2} \quad (\eta>1), \end{cases} \tag{158}$$

and is shown in figure 7. As is expected, in the perfect-entanglement limit

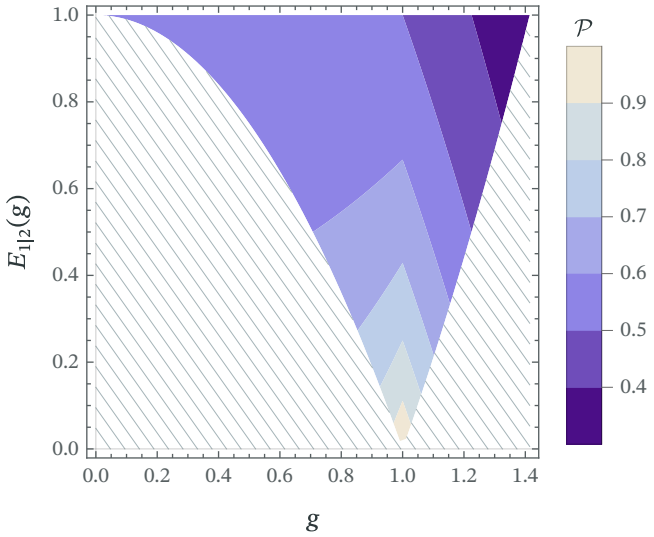


Figure 7: The purity of the output of a QSS protocol acting on a coherent-state secret, for reconstruction parameter and resource-state steering parameter as shown. A purity of 1 indicates a wholly pure state, while any other value indicates the state is mixed.

that $E_{1|2}(g) \to 0$ at $g = 1$ the output state is again a coherent state with purity 1.

5.2.2 *Reconstruction fidelity*

For coherent-state inputs to a mean-preserving protocol, purity is already a good proxy for output quality — it measures the decoherence resulting from the entanglement residue and environment noise, which are the only sources of imperfection introduced to coherent inputs. This is not true of all classes of secret state, however; a squeezed vacuum input, for example, may be attenuated and replaced with any amount of vacuum noise with no

impact on its purity. To enable a comparison across different secret states, let us consider instead a more direct measurement of the output's likeness to the input: the fidelity, a mixed-state generalisation of state overlap.

For a pure input state, fidelity reduces simply to the Hilbert-Schmidt overlap which for Gaussian states can be written as [23]

$$\begin{aligned}\mathcal{F} &= \langle\psi|\hat{\rho}_{\text{out}}|\psi\rangle \\ &= \frac{2}{\sqrt{\det(V_\psi + V_{\text{out}})}} \exp\left[-(\bar{\boldsymbol{r}}_\psi - \bar{\boldsymbol{r}}_{\text{out}})^T (V_\psi + V_{\text{out}})^{-1} (\bar{\boldsymbol{r}}_\psi - \bar{\boldsymbol{r}}_{\text{out}})\right]. \end{aligned} \tag{159}$$

As our QSS protocol preserves the mean vector, the exponential component to the fidelity vanishes, leaving it a function of the covariance matrices alone as

$$\mathcal{F} = \frac{2}{\sqrt{\det(V_\psi + V_{\text{out}})}}. \tag{160}$$

Applying this formula first to the output state reconstructed from shares 1 & 2 given in equation (154), we get a reconstruction fidelity given by

$$\mathcal{F}_{\{1,2\}} = \begin{cases} 1/(2-\eta^2) & g < 1 \quad (\eta < 1) \\ 1 & 1 \leq g < \sqrt{2}, \end{cases} \tag{161}$$

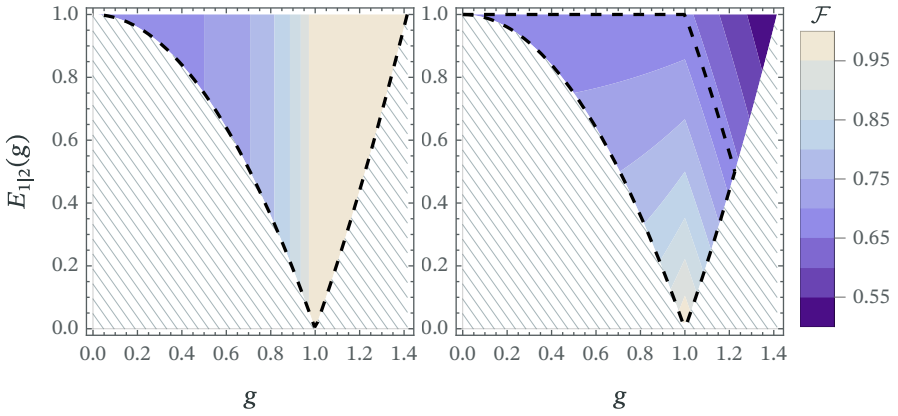
for $\eta(g) = 1/\sqrt{2-g^2}$. As expected, the reconstructed fidelity has no dependence on the properties of the resource state used and is only impacted by the degree to which the secret state has been amplified prior to the dealer. In the extreme limit, in which $g \to 0$, the worst-case reconstruction fidelity for {1,2} reconstruction minimises to $\mathcal{F}_{\{1,2\}} = 2/3$ so the no-cloning threshold is always satisfied.

Applying the fidelity formula instead to the covariance matrix in equation (156) for {1,3} or {2,3} reconstruction, we get

$$\mathcal{F}_{\{1,3\}} = \begin{cases} 2/(3+\eta^2 E_{1|2}(g) - \eta^2) & g < 1 \quad (\eta < 1) \\ 2/(2+E_{1|2}(g)) & g = 1 \quad (\eta = 1) \\ 2/(3+E_{1|2}(g) - 1/\eta^2) & 1 < g < \sqrt{2} \quad (\eta > 1), \end{cases} \tag{162}$$

again for $\eta(g) = 1/\sqrt{2-g^2}$. The achievable reconstruction fidelity for each of these cases is shown in figure 8.

As the achievable fidelity from {1,2} reconstruction is always better than {1,3} reconstruction, let us focus now solely on the latter, shown in figure 8b. As might be expected, the reconstruction fidelity improves significantly



(a) {1,2} reconstruction (b) {1,3} / {2,3} reconstruction

Figure 8: The fidelity between the reconstructed output state and the original input state for a coherent-state secret when the output is constructed from (a) shares 1 & 2, (b) shares 1 & 3 or 2 & 3. Dashed black line indicates the region in which secure QSS can be achieved; see section 5.2.4. Shaded region denotes $E_{1|2}(g) < |1 - g^2|$ in which quantum steering is not physically possible.

with increasing resource entanglement and as $E_{1|2}(g) \to 0$, approaches perfect reconstruction.

There is a clear bias towards implementing the protocol for $g = 1$. Not only is this the point at which the greatest levels of quantum steering exist, but for a constant level of resource steering it can be seen that fidelity drops off gently as $g < 1$ and rapidly as $g > 1$. This effect is due to the amplification that occurs as part of the protocol for non-unity reconstruction parameters. Both amplification and de-amplification introduce additional noise into the state, so the best possible reconstruction is obtained when these corrections are not required. The asymmetry around the unity-amplification point — with fidelity dropping off more rapidly above $g = 1$ than below it — is explained by the different points in which the amplification correction is introduced. Fundamentally, each case involves both an amplification and a deamplification. When $g < 1$, though, this amplification occurs prior to the dealer protocol so only impacts the secret state, while the attenuation imparted by the reconstruction protocol affects both secret and resource states alike, leaving the latter ultimately de-amplified. For $g > 1$, on the other hand, the amplification is an unavoidable part of the reconstruction step so acts on the resource state contributions alongside the desired secret state — leaving both parts ultimately non-amplified after the de-amplification correction. Despite requiring the same degree of amplification correction a worse outcome is achieved for reconstruction parameters above 1. This asymmetry is discussed further in section

section 5.2.3, where the potential to improve fidelity by converting a $g > 1$ protocol into a $g < 1$ protocol is considered.

Let us finally pause here to compare these results to the fidelity achievable from quantum teleportation protocols. He *et al.* [45] found that, when teleported using a generally-asymmetric resource state, coherent states were amplified by a factor of g, and so the optimal fidelity was

$$\mathcal{F}_{\text{teleportation}} = 2/\begin{cases} 3 + g^2 E_{1|2}(g) - g^2 & g < 1, \\ 2 + E_{1|2}(g) & g = 1, \\ 3 + E_{1|2}(g) - \frac{1}{g^2} & 1 < g < \sqrt{2}. \end{cases} \tag{163}$$

Comparing this result to the worst-case state reconstruction fidelity derived in equation (162) for QSS, we can see that when implemented for $g \neq 1$, teleportation imposes a higher amplification on the state ($\eta > g$ for $g < 1$, $1/\eta > 1/g$ for $g > 1$) so quickly becomes more destructive than quantum state sharing. In addition to providing a better quality output for any given resource state, quantum state sharing also reduces the entanglement required to achieve a certain threshold fidelity and certify security, which we discuss in section 5.2.4.

The increased fidelity obtainable from QSS is compounded by the fact that we have only considered here the worst-case reconstruction option. One third of the expected state reconstructions will be performed using the {1,2}-reconstruction sub-protocol so on average the fidelity from QSS will outperform teleportation by a greater margin.

5.2.3 *Optimising fidelity by swapping resource modes*

Let us now consider in further detail the asymmetry around $g = 1$, where the impact of the amplification correction increases more rapidly for $g > 1$ than for $g < 1$.

It is trivial to show that every state steerable in one direction for some $g > 1$ is also necessarily steerable in the opposite direction for $\bar{g} := 1/g < 1$. We will show here that, in fact, it is *always* preferable to swap the order of the resource modes such that the quantum steering is utilised in the direction that permits a $g < 1$ reconstruction.

Let us first note that, given a state with steering parameter $E_{1|2}(g) < 1$ for some $g > 1$, the steering parameter in the opposite direction for $\bar{g}$ can be found as

$$E_{2|1}(\bar{g}) = m + \bar{g}^2 n - 2\bar{g}c = \frac{1}{g}^2 (n + g^2 m - 2gc) = \frac{1}{g^2} E_{1|2}(g), \tag{164}$$

for n, m, c the symplectic invariants described in section 2.1.3.

Let us now consider the case in which QSS is implemented directly for $g > 1$ for the given resource; the output state will have fidelity given by equation (162) as

$$\mathcal{F}_{1|2} = \frac{2}{3 + E_{1|2}(g) - 1/\eta^2}, \tag{165}$$

for, as usual, $\eta = 1/\sqrt{2 - g^2}$.

If the dealer had instead swapped the resource modes such that the quantum steering was exploited in the opposite direction for $\bar{g} = 1/g < 1$, the reconstruction state would have fidelity

$$\mathcal{F}_{2|1} = \frac{2}{3 + \bar{\eta}^2 E_{2|1}(\bar{g}) - \bar{\eta}^2}, \tag{166}$$

for $\bar{\eta} = 1/\sqrt{2 - \bar{g}^2}$. It is preferable, therefore, to swap the order of the resource modes whenever

$$\mathcal{F}_{2|1} = \frac{2}{3 + \bar{\eta}^2 E_{2|1}(\bar{g}) - \bar{\eta}^2} > \frac{2}{3 + E_{1|2}(g) - 1/\eta^2} = \mathcal{F}_{1|2}, \tag{167}$$

simplifying to the condition that

$$3 + \bar{\eta}^2 \frac{1}{g^2} E_{1|2}(g) - \bar{\eta}^2 < 3 + E_{1|2}(g) - 1/\eta^2, \tag{168}$$

$$\implies \quad (\bar{\eta}^2 \frac{1}{g^2} - 1) E_{1|2}(g) < \bar{\eta}^2 - 1/\eta^2, \tag{169}$$

$$\implies \quad E_{1|2}(g) > 1 - g^2. \tag{170}$$

This is trivially satisfied for all $g > 1$, so it is in every case preferable to reorder the resource modes such that the quantum steering is utilised for $g < 1$ when this is an option.

This reordering relies on the dealer's ability to select which resource mode is mixed with the secret state and so in which direction the resource is quantum steered. In some instances this option may not be available to the dealer. For example, it may be desirable to distribute the second resource mode to player 3 prior to the dealer being given the secret state.

For this reason, we will continue to analyse the protocol's effectiveness for the full $0 < g < \sqrt{2}$ range in the remainder of this part of the thesis.

Finally, although this proof is presented here for coherent states, it equally applies to the full class of mixed Gaussian state that we will study later in section 5.3.2.[3]

5.2.4 *General security*

Having established the expected reconstruction fidelity for a given resource state (characterised by $E_{1|2}(g)$) and reconstruction protocol (characterised by choice of g) in the previous subsection, let us now consider the conditions under which the security of state transmission can be guaranteed.

Recall from section 4.3 that we have defined our quantum state sharing protocol to be secure when the collaborating players can be certain their copy of the secret state contains more information than the adversary has access to, a condition we certify using the no-cloning theorem. Under the assumption that the input states are equally drawn from the full spectrum of coherent amplitudes, Grosshans and Grangier [54] derived the maximum allowed cloning fidelity as $\mathcal{F} = 2/3$.

Recalling from section 5.2.2 that the worst-case fidelity obtainable from {1,2} reconstruction, in the $g \to 0$ limit, is $\mathcal{F} = 2/3$ this security condition is automatically satisfied in that case. As we must certify security independently for every reconstruction option, the ease of achieving security for {1,2} QSS does not reduce our security requirements, though, so we do not consider it further here.

Comparing this cloning threshold to the achievable fidelity for {1,3} or {2,3} reconstruction given in equation (162) allows us to state our first security condition.

Result 3. *A sufficient condition for a two-mode resource state to be useful for secure* $(2,3)$*-threshold QSS of a coherent-state secret is that a* $g \in (0, \sqrt{2})$ *exist such that its steering parameter satisfies*

$$E_{1|2}(g) < \begin{cases} 1 & g \leq 1, \\ 2 - g^2 & g > 1. \end{cases} \tag{171}$$

Notably, while this result shows that any resource state exhibiting EPR-steering for some $g \leq 1$ is useful for secure QSS, a larger degree of quantum

[3] This can be seen by first noting that the general single-mode Gaussian fidelity expression from result 8 decreases monotonically with increasing χ. That χ is minimised by swapping resource modes is equivalent to the conditions we have expressed in equations (168) to (170).

steering is required when the protocol is implemented for $g > 1$. This asymmetry in resource requirements reflects what we saw in section 5.2.3: it is in fact always preferable to exploit the entanglement in such a way that $g < 1$. Recalling that every resource steerable for $g > 1$ is also steerable for some $\bar{g} = 1/g < 1$, accounting for the dealer's ability to swap the direction in which they utilise the resource entanglement by relabelling modes $1 \leftrightarrow 2$, allows us to restate this result to take a more general view on the requirements for secure QSS.

Result 4. *All EPR-steerable states (one way and two way) are useful for the secure sharing of a coherent-state secret with a suitable dealer allocation of resource modes.*

Again, though, we caution the reader that this result relies on the assumption that the dealer has access to both resource modes and so is able to choose the direction in which the quantum steering is exploited. This should always be possible according to a strict reading of the specification of this protocol; in practice, however, as the dealer only needs access to a single resource mode it is possible to operate a version of this protocol in which the second resource mode is not provided to them. In such a case result 4 would no longer apply.

5.2.5 *Security for limited codebooks*

The security analysis we have presented in the previous subsection is based on the standard assumption that secret states are drawn from the full range of coherent states. Although useful for a first-level analysis of protocol security, this is not an assumption that holds in practice as a coherent state's amplitude is directly proportional to its energy. In any real-world setup, there will naturally be an upper limit on the coherent states that could form the secret state based on the energy and equipment resources available. To guarantee security for any quantum communication protocol, then, a bespoke security analysis should be performed taking into account the specific codebook available and an assumption of the relative likelihood of any given state being present.

Let us consider an example of such a limited codebook now. The optimal way to encode information in coherent states is to draw their amplitudes from a Gaussian distribution [61] with probability distribution function

$$P(\bar{\boldsymbol{r}}) = \frac{1}{\pi\sigma^2} \exp\Big(\frac{-\bar{r}_x^2 - \bar{r}_p^2}{\sigma^2}\Big), \tag{172}$$

for some distribution variance σ relative to the standard quantum limit.

The ability to clone coherent states drawn from known probability distributions was studied by Cochrane *et al.* in 2004 [61]. They found that when the coherent amplitudes were drawn from a Gaussian probability distribution, the average cloning fidelity depended solely on the width of the distribution.[4] In this case, they found that the states could be optimally cloned with fidelity

$$\mathcal{F}_{\text{threshold}} = \begin{cases} \frac{2}{(3-2\sqrt{2})\sigma^2+2} & \sigma^2 \leq 1+\sqrt{2} \\ \frac{2\sigma^2+2}{3\sigma^2+1} & \sigma^2 \geq 1+\sqrt{2}, \end{cases} \tag{173}$$

for σ the width of the Gaussian relative to the standard quantum limit.[5] For example, coherent states drawn from an input distribution with variance of two standard quantum limits, $\sigma = 2$, can be cloned with fidelity $\mathcal{F}_{\text{clone}} \approx 0.77$, compared to $\mathcal{F}_{\text{clone}} = 2/3$ for states drawn from the full spectrum.

To ensure security for such an input distribution, the average reconstruction fidelity across the distribution must exceed this new fidelity threshold. As the reconstruction fidelity for our protocol does not depend on the coherent amplitude, this limit can be directly applied to the reconstruction fidelity given in equation (162), allowing us to state the following updated security condition.

Result 5. *A resource state is useful for secure tripartite quantum state sharing of coherent states drawn from a Gaussian distribution of width σ relative to the standard quantum limit if some $g \in (0, \sqrt{2})$ exists such that its steering parameter satisfies*

$$E_{1|2}(g) < \begin{cases} 1 - \frac{1}{\eta^2} + \frac{1}{\eta^2}(3-2\sqrt{2})\sigma^2 & \sigma^2 \leq 1+\sqrt{2},\ g \leq 1 \\ \frac{1}{\eta^2} - 1 + (3-2\sqrt{2})\sigma^2 & \sigma^2 \leq 1+\sqrt{2},\ g \geq 1 \\ 1 - \frac{2}{\eta^2(1+\sigma^2)} & \sigma^2 \geq 1+\sqrt{2},\ g \leq 1 \\ \frac{1}{\eta^2} - \frac{2}{1+\sigma^2} & \sigma^2 \geq 1+\sqrt{2},\ g \geq 1. \end{cases} \tag{174}$$

In the $\sigma \to \infty$ limit, that is the limit in which there is no information about the input state distribution, this result reduces to that found in result 3. The minimum amount of quantum steering required to be present to

[4] As details of the distribution are public knowledge, any distribution with non-zero mean can be converted to a zero-mean distribution with a blanket displacement operation, cloned, and then converted back with the same output fidelity as a distribution that was originally zero-mean

[5] One might be surprised to see that this no-cloning threshold is given by a piecewise function. This is because the mathematically-optimal cloning machine is not physical for distributions narrower than $\sigma = 1+\sqrt{2}$! The cloner involves an amplifier with, optimally, an amplification gain of $G = 8\sigma^4/(2\sigma^2+1)^2$, but any amplifier must also have a gain $G > 1$.

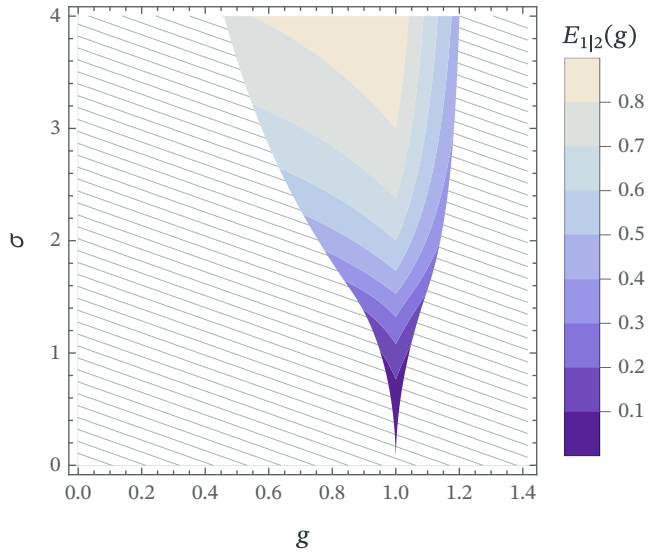


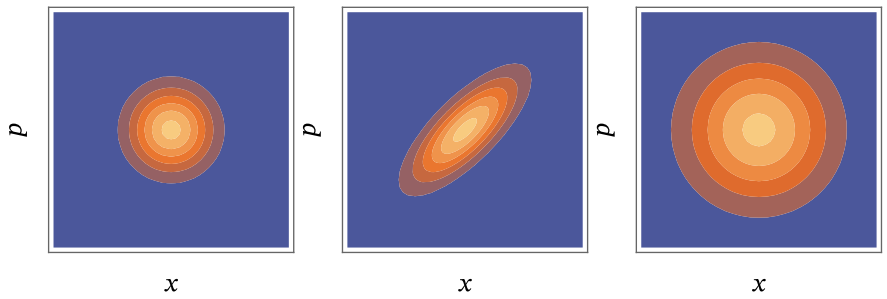
Figure 9: Theoretical minimum resource steering (maximum steering parameter $E_{1|2}(g)$) necessary to securely share coherent states drawn from a Gaussian distribution with width σ relative to the standard quantum limit. Only steering parameters physically allowed are shown; shaded region denotes that theoretically-required steering is not physical for that g value, based on the requirement that $E_{1|2}(g) \leq |1 - g^2|$.

share states drawn from such a distribution is shown in figure 9 for varying Gaussian width. The protocol's usefulness for very tight codebooks (those drawn from a Gaussian distribution of width $\sigma \lesssim 1$ standard quantum limit) is limited, with such distributions nearly impossible to share securely. This is not unexpected, as such small distributions of states are very easy to clone. As the distribution increases in width, however, we quickly enter a region in which even low levels of entanglement are able to provide guaranteed security. This more rapid reduction in reconstruction fidelity as g drops below 1 (when coupled with the fundamental limits on asymmetric steering that $E_{1|2}(g) > |1 - g^2|$) imposes a limit on secure QSS that was not previously present. The region for which the entanglement level to achieve secure QSS can exist is denoted by the dashed line in figure 9 — as σ increases, the region of allowed reconstruction parameters also increases but there remains a limit on how asymmetric the entanglement may be for security to be possible. For practical coherent-state QSS, then, one is limited to the broadly-symmetric region around $g = 1$.

Finally, let us note here that even this is a somewhat over-pessimistic analysis. The QSS protocol presented here is optimised assuming the full range of coherent states; recall that we have defined it to be mean-preserving in every case. This is a necessary choice when the states have fully unknown coherent amplitude as the increasingly-large disparity in input/output

mean as $\bar{\boldsymbol{r}}$ increases would otherwise destroy the reconstruction fidelity. However, when input states are limited to a region with small coherent amplitude, a better reconstruction fidelity may be achieved by allowing for a small uncorrected amplification. The question as to how to maximise fidelity for a given input distribution is beyond the scope of this thesis, but may be worth investigating should one look to implement this protocol in the future.

5.3 QUANTUM STATE SHARING OF OTHER SINGLE-MODE GAUSSIAN STATES



(a) Coherent state (b) Squeezed state (c) Thermal state

Figure 10: Wigner functions representing (a) a coherent state, (b) a squeezed state with 4 dB squeezing at $\theta = \pi/4$, (c) an unsqueezed thermal state with average thermal photon number $\bar{n} = 1$. The coherent and squeezed states saturate the uncertainty limit, with the uncertainty reduced in one quadrature of the squeezed state and increased in the other, while the thermal state has additional variance represented by the presence of thermal photons in addition to those due to the coherent amplitude of the state. All state representations presented on the same scale.

Although coherent states are broadly the most important of the Gaussian states, they represent only one subset of possible single-mode states. In this section, we will consider the wider classes of Gaussian state available, and so analyse this protocol's effectiveness more generally.

We will begin in section 5.3.1 by expanding on our analysis of coherent states to allow for an unknown squeezing in be present within the secret state. These squeezed coherent states, shown in figure 10(b) represent the most general possible *pure* Gaussian state, and so complete our first-level analysis of this protocol. We will show that, subject to a sufficient but not strictly necessary security condition, our QSS protocol can be made secure for the sharing of states subject to *any* degree of squeezing.

We will then go on in section 5.3.2 to relax the assumption that the input state be pure. By considering squeezed displaced thermal states, the most general possible single-mode Gaussian state, we will show that this

protocol remains effective for the entire class of Gaussian states and that arbitrarily high fidelity can be achieved regardless of the input state. An example of a thermal state is shown in figure 10(c).

5.3.1 *Squeezed-state quantum state sharing*

A squeezed state is a Gaussian state in which the uncertainty in one quadrature is reduced below the standard quantum limit at the expense of a corresponding increase in the other quadrature. The squeezed states we will consider in this subsection, squeezed coherent states, are those which continue to saturate the uncertainty limit ($\Delta x \Delta p = 1$). These states have arbitrary mean $\bar{\boldsymbol{r}} \in \mathbb{R}^2$ and covariance matrix defined by squeezing parameter $\zeta \in \mathbb{R}$ and squeezing angle θ as

$$V = \begin{pmatrix} \mathrm{e}^{-2\zeta}\cos\theta + e^{2\zeta}\sin\theta & 2\sinh(2\zeta)\cos\theta\sin\theta \\ 2\sinh(2\zeta)\cos\theta\sin\theta & \mathrm{e}^{2\zeta}\cos\theta + e^{-2\zeta}\sin\theta \end{pmatrix}, \tag{175}$$

which continues to have determinant 1.

In general, these states may be squeezed along any angle $\theta \in [0, 2\pi]$ in phase space. As our protocol is phase-independent, however, any phase-rotation of the input state will always be reflected perfectly in the output state so ultimately have no effect on the reconstruction quality. Without losing generality, then, we are free to assume that the state is squeezed along the $\hat{X}^+$ and $\hat{X}^-$ quadratures and take, for simplicity, $\theta = 0$. The input covariance matrix then reduces to

$$V = \begin{pmatrix} \mathrm{e}^{-2\zeta} & 0 \\ 0 & \mathrm{e}^{2\zeta} \end{pmatrix}, \tag{176}$$

for squeezing parameter ζ.

As no change is made to the protocol to account for the change in input state, the output will continue to be described by the covariance matrix given in equation (154),

$$V_{\text{out}} = V_\psi + \begin{cases} (\eta^2 E_{1|2}(g) + 1 - \eta^2) I_2 & g \leq 1 \\ (E_{1|2}(g) + 1 - 1/\eta^2) I_2 & g \geq 1 \end{cases}, \tag{177}$$

with V_ψ now the squeezed state covariance matrix from equation (176).

5.3.1.1 *Impact of QSS on squeezing parameter*

Before we go on to quantify the quality of the output through the usual fidelity measure, let us first briefly consider what impact this protocol has on the degree to which the secret state is squeezed.

Let us consider first *esa* QSS, when $g \leq 1$. In this case, from equation (177), we know the reconstructed secret state will be described by covariance matrix

$$V_{\text{out}} = \begin{pmatrix} e^{-2\zeta} & 0 \\ 0 & e^{2\zeta} \end{pmatrix} + \begin{pmatrix} \eta^2 E_{1|2}(g) + 1 - \eta^2 & 0 \\ 0 & \eta^2 E_{1|2}(g) + 1 - \eta^2 \end{pmatrix}, \quad (178)$$

representing the original squeezed state mixed with unsqueezed thermal noise.

Ultimately, this output state will take the form of a thermal state that has been squeezed to some degree $\zeta' \in \mathbb{R}$ along the same angle as the input state and be described by

$$V_{\text{out}} = (2\tilde{n} + 1) \begin{pmatrix} e^{-2\zeta'} & 0 \\ 0 & e^{2\zeta'} \end{pmatrix}, \quad (179)$$

for $\tilde{n}$ the average number of thermal photons present in the output state.

Equating these two representations we find the degree of squeezing in the output state to be related to the squeezing in the input state and the usual protocol parameters as

$$\zeta' = \frac{1}{4} \ln \left[\frac{e^{2\zeta} + \eta^2 E_{1|2}(g) + 1 - \eta^2}{e^{-2\zeta} + \eta^2 E_{1|2}(g) + 1 - \eta^2} \right] \leq \zeta, \quad (180)$$

which is strictly less than the squeezing of the input state ζ except in the perfect entanglement limit.

Similarly, for *lsatt* QSS when $g \geq 1$ we find the output squeezing parameter to be

$$\zeta' = \frac{1}{4} \ln \left[\frac{e^{2\zeta} + E_{1|2}(g) + 1 - \frac{1}{\eta^2}}{e^{-2\zeta} + E_{1|2}(g) + 1 - \frac{1}{\eta^2}} \right] \leq \zeta, \quad (181)$$

again strictly less than the input squeezing except in the perfect entanglement limit.

We can see that in addition to increasing the noise present within the state, performing quantum state sharing on a squeezed state will always result in a reduction in its squeezing, as shown in figure 11. Recalling that in the absence of perfect entanglement, one effect of the protocol is to mix

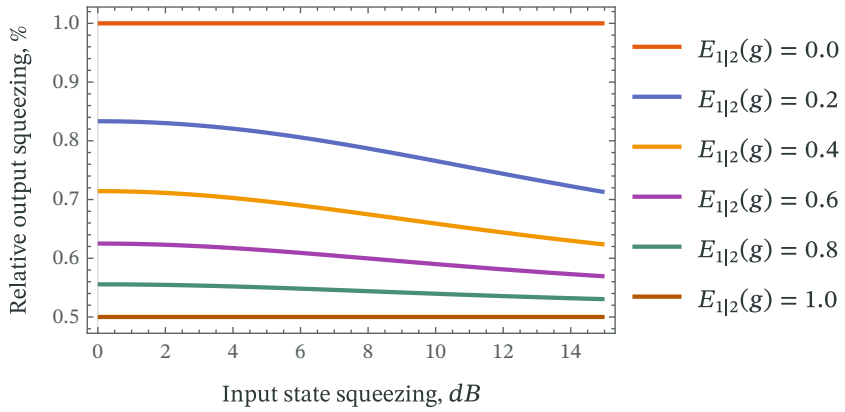


Figure 11: Percentage retention of squeezing level after application of the QSS protocol at $g = 1$ using a resource with steering parameter as labelled.

unsqueezed noise in to the reconstructed state, it should not be surprising that this has the effect of diluting the squeezing level.

In both *lsatt* and *esa* cases, the relative impact on the squeezing parameter given by $1 - \zeta'/\zeta$ increases as the level of squeezing increases, indicating that the protocol is more destructive for highly squeezed states. Even in the highly-squeezed limit, though, the squeezing reduction tends towards a halving in the degree of squeezing present, with

$$\lim_{\zeta \to \infty} \frac{\zeta'}{\zeta} = \frac{1}{2}, \tag{182}$$

regardless of reconstruction parameter g and resource squeezing $E_{1|2}(g) \neq 0$.[6] We can be confident, therefore, that at least half of the squeezing will survive at $g = 1$, so long as the resource state exhibits any form of steering.

5.3.1.2 *Fidelity*

Let us now consider how close this reconstructed output state is to the original input in a more general sense. Applying the general formula for fidelity for a single-mode pure Gaussian input to the output covariance matrix given in equation (177), we can see that a squeezed secret sate will be reconstructed with a fidelity of

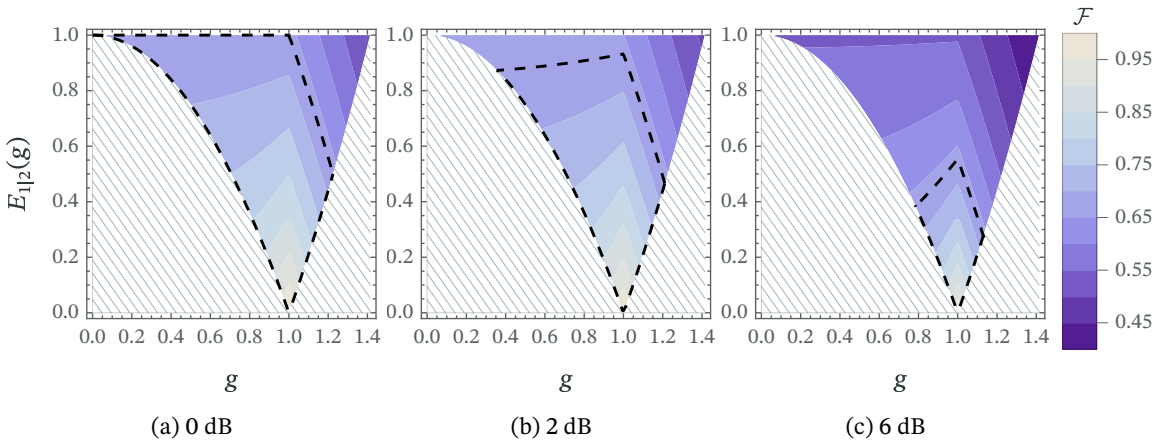
$$\mathcal{F} = \frac{2}{\sqrt{\det(V_\psi + V_{\text{out}})}} = \frac{2}{\sqrt{(2\,\mathrm{e}^{2\zeta} + \chi)(2\,\mathrm{e}^{-2\zeta} + \chi)}}, \tag{183}$$

[6] This $E_{1|2}(g) \neq 0$ condition could otherwise be stated as: for any *physical* resource state.

for

$$\chi = \begin{cases} \eta^2 E_{1|2}(g) + 1 - \eta^2 & g < 1 \quad (\eta < 1) \\ E_{1|2}(g) & g = 1 \\ E_{1|2}(g) + 1 - \frac{1}{\eta^2} & g > 1 \quad (\eta > 1) \end{cases} \tag{184}$$

representing the g-dependent component introduced by the amplification correction and resource residue. This single-shot reconstruction fidelity is shown for a variety of squeezed state inputs in figure 12.



(a) 0 dB (b) 2 dB (c) 6 dB

Figure 12: Reconstruction fidelity for increasing resource steering for pure Gaussian states with (a) 0 dB (no squeezing), (b) 2 dB, (c) 6 dB of squeezing. Contours represent the reconstruction fidelity when QSS is implemented for g using a resource state with steering parameter $E_{1|2}(g)$. The dashed lines denote the secure region for which the reconstruction exceeds the asymptotic no-cloning theorem. As the squeezing in the input state increases, the achievable reconstruction fidelity decreases and this 'secure' region shrinks. Only shown is the region in which $E_{1|2}(g) > |1 - g^2|$ which is physically allowed.

It is clear that increasing the squeezing ζ in the secret state reduces the effectiveness of this protocol — a greater entanglement will be required to maintain a consistent reconstruction fidelity as input squeezing increases. The impact on fidelity of the choice of reconstruction parameter seen in section 5.2.2 for coherent states continues to apply here: the best reconstruction quality is found in the region around $g = 1$, while a good reconstruction is broadly achievable in the $g < 1$ region. As the reconstruction parameter g increases above 1, though, a rapid drop-off in fidelity is seen.

5.3.1.3 *Broad sufficient security bound*

Let us now consider the security requirements for this class of input state, based on our security condition that we must be able to guarantee that no adversaries have access to a greater amount of information than the collaborating parties.

Ordinarily, in producing a security condition one would need to compare the *average* reconstruction fidelity to the *average* possible cloning fidelity, weighted by the probability of each input state for the full range of input states. We were able to skip this process in section 5.2.4 as neither the reconstruction nor the cloning fidelity depended on the input state. This is not the case for squeezed states, however, where we know from equation (183) that the reconstruction fidelity is strongly dependent on the degree of squeezing present in the input state. It has also been shown [56] that Gaussian states become more difficult to clone as squeezing increases, so both the achievable cloning fidelity and the state reconstruction fidelity will contain a ζ dependence.

We will go on consider this approach fully for a specific distribution of input states and obtain a tight security bound in section 5.3.1.4. For now, though, let us draw a more generally applicable security condition by first noting two facts about squeezed states.

1. Squeezed states are strictly more difficult to clone than unsqueezed coherent states; the maximum cloning fidelity for a state with *any* squeezing is $\mathcal{F} = 2/3$ [56].

2. The reconstruction fidelity decreases monotonically with increasing squeezing parameter; for any distribution of states the average reconstruction fidelity is greater than the reconstruction fidelity for the most-squeezed state.

Consequently, to prove security it is *sufficient* (if not necessary) that the fidelity for the most-squeezed state in the distribution exceed $\mathcal{F} = 2/3$. Let us now ask, then, how much resource squeezing is required for a squeezed state with a particular squeezing parameter ζ to be reconstructed with fidelity $\mathcal{F} = 2/3$?

Applying this threshold to equation (183), we can restate our security condition as

Result 6. *A QSS protocol for the sharing of a pure Gaussian secret state with squeezing of up to ζ_{max} is secure if the resource state used has steering of*

$$E_{1|2}(g) < \begin{cases} 1 - \frac{1}{\eta^2}\Gamma(\zeta_{max}) & g < 1 \quad (\eta < 1) \\ 1 - \Gamma(\zeta_{max}) & g = 1 \\ \frac{1}{\eta^2} - \Gamma(\zeta_{max}) & g > 1 \quad (\eta > 1), \end{cases} \tag{185}$$

for some $g \in (0, \sqrt{2})$ where $1/\eta^2 = 2 - g^2$ and

$$\Gamma(\zeta) = 1 + 2\cosh(2\zeta) - \sqrt{4\cosh^2(2\zeta) + 5} \geq 0 \tag{186}$$

is a monotonically increasing function of ζ with $\Gamma(0) = 0$.

As the entanglement requirements strictly increase with increasing squeezing parameter, satisfying the condition for any $\zeta_{\max}$ automatically implies the security of *any* distribution of squeezed states that do not exceed $\zeta_{\max}$.

This security condition is illustrated in figure 13. To securely share states with an increasing degree of squeezing requires increasingly strong entanglement in the resource state. Indeed, as the input state squeezing in-

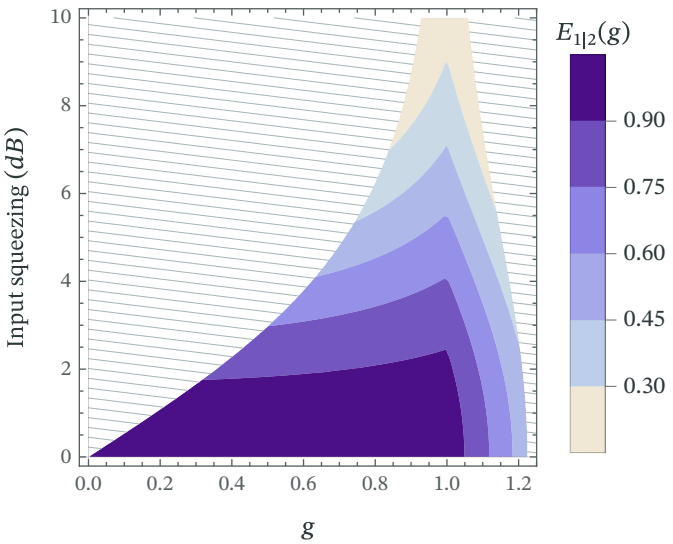


Figure 13: Minimum resource steering that would required to share a coherent state squeezed to the given degree below the standard quantum limit with fidelity exceeding $\mathcal{F} = 2/3$, for a QSS protocol with varying g value. Black dashed line denotes the region within which this required resource steering is physical for the given g. Shaded region indicates that the required level of resource steering would be unphysical.

creases the set of reconstruction parameters g that can support the required steering narrows. To share very highly squeezed states, then, resources exhibiting broadly symmetric $g \approx 1$ steering are required. This reflects a common theme we will find throughout this part of the thesis:[7] as the loosely-defined 'quantumness' of a state increases, the effectiveness of this QSS protocol decreases. It is notable, though, that the required extra steering tends to $\Gamma = 1$ as $\zeta \to \infty$ so even highly-squeezed states (and, in the limit, quadrature states) can be shared securely with a suitably-entangled resource state.[8]

5.3.1.4 *Tight security bound*

The security results in the previous section act as a sufficient guarantor of security, but do not represent a tight bound on the resource requirement. Let us now consider the possibility that a much more permissive result may exist: that squeezed input states may be able to be securely shared with the same resource requirement as coherent states if they are drawn from a suitable distribution, $P(\zeta)$.

It is believed [56] (although not yet rigorously proven) that for a Gaussian distribution of squeezed states the optimal cloning strategy coincides with the optimal coherent-state cloning machine.[9] Although the optimality of this cloning approach has not been proven, let us proceed for this subsection under the assumption that it is the case with the appropriate caveat in mind.

They found that this cloning machine is able to reproduce individual squeezed states with fidelity

$$\mathcal{F}_{\text{clone}}(\zeta) = \frac{2}{\sqrt{9 + 8\sinh^2(\zeta)}}, \tag{187}$$

for some squeezing $\zeta \in \mathbb{R}$.[10]

[7] And indeed one that is found throughout quantum communication.

[8] Although such states are unphysical, and so unlikely to form the secret state, the fact that *in theory* they could be shared in the limit of perfect entanglement demonstrates that any physical secret state is sharable.

[9] This result is believed to be true for Gaussian distributions centred around zero-squeezing because any change to the cloning machine that accommodates squeezing in one quadrature would result in a worse clone for the opposite squeezing. This can then be used for distributions centred on any known mean, $\bar{\zeta}$ by applying a blanket squeezing of $-\bar{\zeta}$ to the states prior to cloning (rebasing the input distribution to a zero-mean Gaussian). The original distribution can then be reproduced by applying an equivalent $\bar{\zeta}$ to the clones.

[10] Strictly speaking, as squeezing can be along any angle, we should consider any $\zeta \in \mathbb{C}$ and perform the integral in equation (188) over the complex plane. However, as neither the cloning nor our QSS protocol is phase sensitive we can neglect the complex phase and consider the complex number wholly real.

To find the average cloning limit for a given input spectrum, then, one would look to evaluate the integral

$$\mathcal{F}_{\text{threshold}} = \int_{-\infty}^{\infty} d\zeta P(\zeta)\mathcal{F}_{\text{clone}}(\zeta). \tag{188}$$

Due to the $\sinh|\zeta|$ term in the denominator of equation (187) this is a highly non-trivial integral to evaluate analytically even for simple distributions such as a Gaussian P. Let us leave it unevaluated then, and move on to the average reconstruction fidelity. Recall, for $g \leq 1$ our reconstruction fidelity for a single squeezed state with unknown squeezing ζ is given by equation (183) as

$$\mathcal{F} = \frac{2}{\sqrt{(2\,\mathrm{e}^{2\zeta} + \eta^2 E_{1|2}(g) + 1 - \eta^2)(2\,\mathrm{e}^{-2\zeta} + \eta^2 E_{1|2}(g) + 1 - \eta^2)}}, \tag{189}$$

which we can turn into an average reconstruction fidelity by again integrating over the probability distribution as

$$\mathcal{F}_{\text{avg}} = \int_{-\infty}^{\infty} d\zeta P(\zeta)\mathcal{F}(\zeta). \tag{190}$$

Knowing though that we aim to show that *any* steering is sufficient, let us cheat somewhat and simplify this expression by considering already only the case for which $E_{1|2}(g) = 1$, the boundary at which steering is certified. At this point, the fidelity expression reduces to

$$\mathcal{F} = \frac{2}{\sqrt{(2\,\mathrm{e}^{2\zeta} + 1)(2\,\mathrm{e}^{-2\zeta} + 1)}} = \frac{2}{\sqrt{9 + 8\sinh^2(\zeta)}}, \tag{191}$$

and the g-dependence vanishes. This is exactly the fidelity achievable by the optimal state cloning machine shown in equation (187)!

The security condition that the average reconstruction fidelity exceed the average achievable cloning fidelity then becomes

$$\mathcal{F}_{\text{avg}} = \int_{-\infty}^{\infty} d\zeta P(\zeta)\mathcal{F}(\zeta) > \int_{-\infty}^{\infty} d\zeta P(\zeta)\mathcal{F}_{\text{clone}}(\zeta) = \mathcal{F}_{\text{threshold}} \tag{192}$$

$$\implies \int_{-\infty}^{\infty} d\zeta P(\zeta)\Big[\mathcal{F}(\zeta) - \mathcal{F}_{\text{clone}}(\zeta)\Big] > 0. \tag{193}$$

Noting that $\mathcal{F}(\zeta) = \mathcal{F}_{\text{clone}}(\zeta)$ for $E_{1|2}(g) = 1$, and that $\mathcal{F}(\zeta)$ only increases with decreasing $E_{1|2}(g)$, this condition is trivially satisfied for any $E_{1|2}(g) < 1$ (and $g \leq 1$) regardless of the input distribution $P(\zeta)$. Consequently, security can be guaranteed for any $E_{1|2}(g) < 1$ and $g \leq 1$ for

any input distribution for which the cloning fidelity in equation (187) is optimal.

Accounting, as we did for coherent states, for our ability to swap modes and ensure $g \leq 1$ for any resource state exhibiting steering we can now restate a familiar security condition.

Result 7. *A sufficient condition for a two-mode resource state to be useful for the secure* $(2, 3)$*-threshold QSS of squeezed secret states drawn from any distribution for which the cloning machine described in Ref. [56] is optimal, is that it exhibit any degree of quantum steering for any* $g \in (0, \sqrt{2})$.

We would again remind the reader here, though, that this result relies upon the likely but unproven assumption that the cloning machine outlined in Ref. [56] is optimal for a Gaussian distribution of squeezed states. Should that optimality be proven at any point in the future, then this result would apply immediately. However, in the absence of a proof of optimality care should be taken in relying upon this result alone for security.

Let us also note that this security analysis again assumes states are drawn from the full range of coherent amplitudes $\bar{\boldsymbol{r}}$. A bespoke security analysis for a real-world implementation of this protocol would have to be performed for the specific distribution of secret state coherent amplitudes, as performed for unsqueezed states in section 5.2.4, to guarantee security.

5.3.2 *Thermal-state quantum state sharing*

Finally, let us consider the use of this protocol for real-world continuous-variable states: (squeezed) displaced thermal states. Until now, we have exclusively considered pure states: idealised quantum states that saturate the uncertainty limit. Although useful for the core analysis of the protocol's effectiveness, these minimum-uncertainty states do not exist in the real world — any state initially created as a pure coherent state would immediately begin to interact with the environment and decohere into a mixed thermal state. Often, this decoherence can be limited through the use of controlled cryogenic environments such that the pure-state approximation may be close enough [53, 62]. Nonetheless, it is worth considering the effect this protocol has on states of varying mixedness.

A single-mode squeezed thermal state is characterised by the covariance matrix

$$V = (2\bar{n} + 1)\begin{pmatrix} \exp(-2\zeta) & 0 \\ 0 & \exp(2\zeta) \end{pmatrix}, \tag{194}$$

where $\bar{n} \geq 0$ represents the average number of thermal photons in the state prior to displacement, and ζ again represents a quadrature squeezing imposed on the state. We will also assume here that the state is additionally displaced such that information may be encoded within the mean vector $\bar{\boldsymbol{r}} \in \mathbb{R}^2$.

5.3.2.1 *Reconstruction fidelity for mixed states*

As our input state is no longer pure, the simple formula for fidelity from equation (159) is no longer valid. Instead, we must consider the full form of the Uhlmann fidelity for Gaussian states given by [24]

$$\mathcal{F} = \frac{2}{\sqrt{\Delta + \delta} - \sqrt{\delta}} \exp\Big[-(\bar{\boldsymbol{r}}_\psi - \bar{\boldsymbol{r}}_{\text{out}})^T (V_\psi + V_{\text{out}})^{-1} (\bar{\boldsymbol{r}}_\psi - \bar{\boldsymbol{r}}_{\text{out}})\Big], \tag{195}$$

reducing to

$$\mathcal{F} = \frac{2}{\sqrt{\Delta + \delta} - \sqrt{\delta}} \tag{196}$$

for equal-mean states, where

$$\Delta = \det(V_\psi + V_{\text{out}}), \tag{197}$$

$$\delta = (\det V_\psi - 1)(\det V_{\text{out}} - 1). \tag{198}$$

The output state will again be described by the covariance matrix

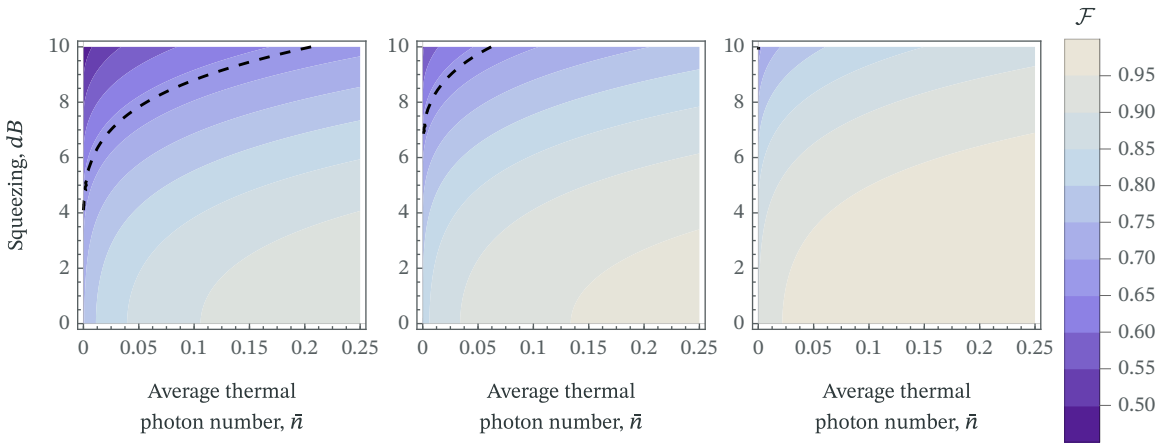
$$V_{\text{out}} = V_\psi + \chi I_2, \tag{199}$$

where χ represents both noise from the use of an imperfect resource state and from the (de)amplification, and is given by

$$\chi = \begin{cases} \eta^2 E_{1|2}(g) + 1 - \eta^2 & g \leq 1 \quad (\eta < 1) \\ E_{1|2}(g) & g = 1 \\ E_{1|2}(g) + 1 - \frac{1}{\eta^2} & g > 1 \quad (\eta > 1) \end{cases}. \tag{200}$$

The two covariance-matrix-dependent components to the fidelity expression will then take the form

$$\begin{aligned} \Delta &= \det(V_\psi + V_{\text{out}}) \\ &= \left(2\tilde{n}\, \mathrm{e}^{-2\zeta} + \chi\right)\left(2\tilde{n}\, \mathrm{e}^{2\zeta} + \chi\right) \\ &= 4\tilde{n} + \chi^2 + 4\tilde{n}\chi \cosh(2\zeta), \end{aligned} \tag{201}$$



(a) $E_{1|2}(g) = 0.75$ (b) $E_{1|2}(g) = 0.5$ (c) $E_{1|2}(g) = 0.25$

Figure 14: Reconstruction fidelity achievable at $g = 1$ for a single-mode Gaussian secret with given thermal photon number $\bar{n}$ and squeezing in decibels. Black dashed line denotes $\mathcal{F} = 2/3$ as a benchmark only, as this does not guarantee security for general Gaussian states.

and

$$\begin{aligned}\delta &= (\det V_\psi - 1)(\det V_{\text{out}} - 1)\\ &= (\tilde{n}^2 - 1)\big((\tilde{n}\,\mathrm{e}^{-2\zeta} + \chi)(\tilde{n}\,\mathrm{e}^{2\zeta} + \chi) - 1\big)\\ &= (\tilde{n}^2 - 1)\left(\tilde{n}^2 + \chi^2 + 2\tilde{n}\chi\cosh(2\zeta) - 1\right),\end{aligned} \tag{202}$$

for $\tilde{n} = 2\bar{n} + 1$ representing the single-mode variance arising from the $\bar{n}$ thermal photons in the input state. We can then directly calculate the reconstruction fidelity for a thermal input state from equation (196).

Result 8. *An arbitrary single-mode Gaussian state with squeezing parameter ζ and thermal photon number $\bar{n}$ can be shared and reconstructed with fidelity*

$$\mathcal{F} = 2\Big/\left\{\begin{array}{c}\sqrt{\left(\tilde{n}\chi + (\tilde{n}^2+1)\,\mathrm{e}^{2\zeta}\right)\left(\tilde{n}\chi + (\tilde{n}^2+1)\,\mathrm{e}^{-2\zeta}\right)}\\ -\sqrt{(\tilde{n}^2-1)\left(\tilde{n}^2 + \chi^2 + 2\tilde{n}\chi\cosh(2\zeta) - 1\right)}\end{array}\right\}, \tag{203}$$

for $\tilde{n} = (2\bar{n} + 1)$ and

$$\chi = \begin{cases}\eta^2 E_{1|2}(g) + 1 - \eta^2 & g \le 1 \quad \textit{(esa)}\\ E_{1|2}(g) + 1 - \frac{1}{\eta^2} & g \ge 1 \quad \textit{(lsatt)}\end{cases} \tag{204}$$

when a resource state exhibiting steering of $E_{1|2}(g)$ is used.

The pure-input-state case is recovered from result 8 by setting $\tilde{n} = 1$, whereupon the second root in the denominator vanishes and the fidelity reduces to that precisely previously found in equation (183). The achievable reconstruction fidelity for thermal states is shown in figure 14. Although

these states continue to be more difficult to share as the squeezing increases, this effect is very quickly overwhelmed by the increase in reconstruction fidelity obtained as the thermality of the input state increases. As seen most obviously in figure 14c for $E_{1|2}(g) = 0.25$, equivalent to 9 dB of resource-squeezing for a TMSV state, a large proportion of squeezed states are able to be reconstructed with near-perfect fidelity as the average thermal photon number increases. This is not an unexpected result — in addition to being more classical than uncertainty-saturating states, their existing thermality makes these states resilient to added thermal noise and so less impacted by both the resource residue and amplification corrections.

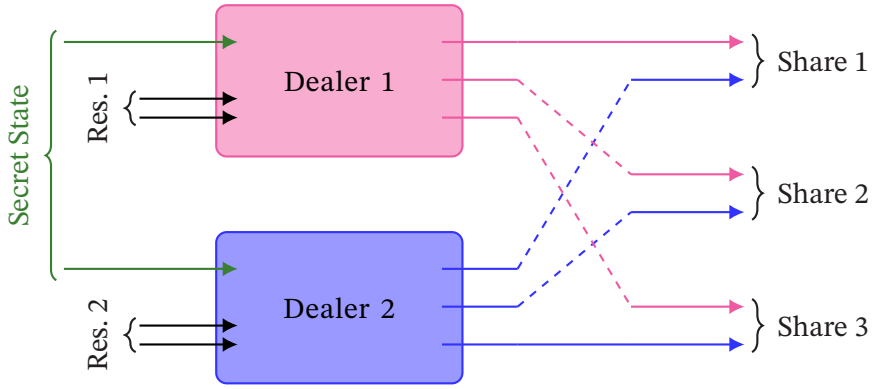
5.3.2.2 *Security*

In contrast to previous sections, we do not present a security analysis for thermal state QSS here. As they do not saturate the uncertainty limit, thermal states can be cloned with significantly better fidelity than their coherent state counterparts; indeed, in the limit of infinite thermal photon number thermal states can be cloned perfectly. Some work has been done studying the cloning potential of thermal states. In 2006, Olivares *et al.* analysed the effectiveness of the standard coherent-state cloning machine for thermal states [56], but did not argue that it was optimal. A cloning strategy proven to be optimal for minimising norm distance was presented by Guta *et al.* in 2006 [63]; it may be possible to derive a security condition based on norm distance, but the details of this are beyond the scope of this thesis. That this cloning strategy is optimal under the norm-distance metric does not mean that it is also optimal for maximising clone fidelity, so this cloning machine is not suitable for a fidelity-based security condition. In the event that one of these strategies be proven to optimise cloning fidelity, a security condition could easily be derived for a subset of thermal states by a similar method to that presented in sections 5.2.4 and 5.3.1.4 for coherent and squeezed states.

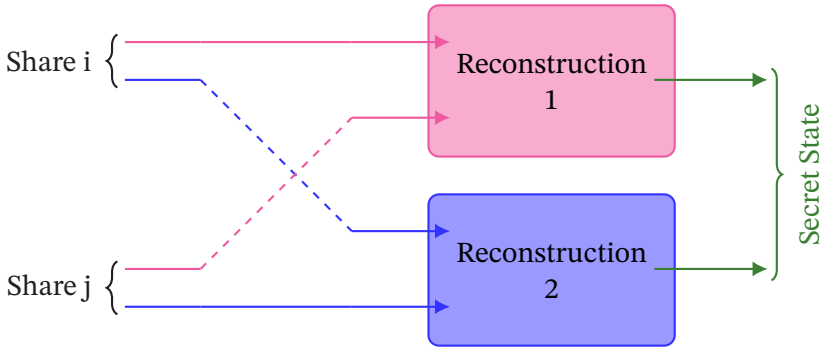
5.4 MULTI-MODE QUANTUM STATE SHARING

Now that we have characterised this quantum state sharing protocol for the full range of single-mode Gaussian states, a natural next question arises: can it be adapted to allow for the sharing of multi-mode states? The desire to share these states may arise, for example, in distributed quantum computing stacks in which one may wish to transmit an entangled multi-mode state that has arisen mid-computation.

We consider in this section a 'local' adaptation to our QSS protocol in which each mode of the secret is shared with its own single-mode dealer protocol. The $3n$ outputs from these dealer protocols are then grouped into three shares, each containing a single output mode from each dealer as shown in figure 15a.



(a) Each mode of the secret state is passed to an independent dealer protocol and locally converted into three shares. The two sets of output modes are then grouped into three shares such that each share contains one mode from each dealer.



(b) Each mode of the original state can be reconstructed individually, by passing the relevant parts of each share through a local reconstruction setup.

Figure 15: An example of a 'local' quantum state sharing protocol acting on a two-mode secret state.

Should any two players collaborate to reconstruct the state they will then have access to two components from each input mode and can reproduce the input state as a whole by reconstructing each mode individually, as shown in figure 15b.

Having studied this protocol's use for single-mode Gaussian states in sections 5.2 and 5.3, we already know it to be effective at reproducing the intra-mode features of a quantum state. We are be primarily interested here, then, in how well the inter-mode features such as entanglement are preserved. In this section, we begin to answer this question by analysing the simplest multi-mode case: a 2-mode secret state.

We are unable to discuss the question of security for such states as the optimal cloning process is not known. Some research into this area has occurred, with cloning machines proposed by Weedbrook *et al.* [22] (which assumes pre-existing knowledge of the entanglement structure) and by Ge *et al.* [64] (which does not, but nonetheless has not been proven optimal). However, a provably optimal cloning strategy for multi-mode entangled Gaussian states remains unknown — as does, consequently, a threshold fidelity against which we can certify guaranteed security from the no-cloning theory.

Nonetheless, we will show here that this protocol remains effective for multi-mode states with arbitrarily high fidelity reachable in the limit of perfect entanglement. Should a no-cloning threshold be derived for multi-mode entangled Gaussian states in the future, the security requirements for this protocol can be derived through the same technique as for coherent states in section 5.2.4.

5.4.1 *Two-mode quantum state sharing*

To allow us to consider this protocols effectiveness in more detail, let us focus only on a pure 2-mode secret input state. The protocol can be extended to any number of modes simply by appending additional single-mode QSS protocols to this 2-mode case; if of interest, these could then be analysed in the same way.

Recalling that every 2-mode Gaussian state can be put into normal form through local operations alone, to better study the protocols impact on entangled states let us take as our input the state with arbitrary mean vector $\bar{\boldsymbol{r}} \in \mathbb{R}^4$ and covariance matrix [22]

$$V = \begin{pmatrix} n & 0 & c & 0 \\ 0 & n & 0 & -c \\ c & 0 & m & 0 \\ 0 & -c & 0 & m \end{pmatrix}, \tag{205}$$

for n, $m \in \mathbb{R}$ and $c \leq \sqrt{nm-1}$.

As we are interested in the worst-case scenario, let us consider that the collaborating players have access to the 1st and 3rd shares for each mode,

$$\hat{X}_1^{\pm} = \frac{1}{\sqrt{2}}(\hat{X}_{\psi}^{\pm} + \hat{X}_{r1}^{\pm}), \tag{206}$$

$$\hat{X}_3^{\pm} = \hat{X}_{r2}^{\pm}. \tag{207}$$

Consequently, each of the modes will be reconstructed as

$$\hat{X}^{\pm}_{\text{out}} = \begin{cases} \hat{X}^{\pm}_{\psi} + \hat{X}^{\pm}_{r1} \mp g\hat{X}^{\pm}_{r2} + \sqrt{\frac{1}{\eta^2} - 1}\hat{X}^{\pm}_{\text{vac}} & g < 1 \quad (\eta < 1) \\ \hat{X}^{\pm}_{\psi} + \hat{X}^{\pm}_{r1} \mp \hat{X}^{\pm}_{r2} & g = 1 \\ \hat{X}^{\pm}_{\psi} + \hat{X}^{\pm}_{r1} \mp g\hat{X}^{\pm}_{r2} + \sqrt{1 - \frac{1}{\eta^2}}\hat{X}^{\pm}_{\text{vac}} & g > 1 \quad (\eta > 1) \end{cases}. \tag{208}$$

Assuming there are no correlations between the two resource states, the output state will have covariance matrix

$$V_{\text{out}} = V_{\psi} + \begin{pmatrix} (\eta^2 E^{r1}_{1|2}(g) + 1 - \eta^2)I_2 & 0_2 \\ 0_2 & (\eta^2 E^{r2}_{1|2}(g) + 1 - \eta^2)I_2 \end{pmatrix}, \tag{209}$$

for $g \leq 1$ and

$$V_{\text{out}} = V_{\psi} + \begin{pmatrix} (E^{r1}_{1|2}(g) + 1 - 1/\eta^2)I_2 & 0_2 \\ 0_2 & (E^{r2}_{1|2}(g) + 1 - 1/\eta^2)I_2 \end{pmatrix}, \tag{210}$$

for $g \geq 1$, where $E^{ri}_{1|2}(g)$ represents the steering parameter for the ith resource state.

From the form of equations (209) and (210) we can deduce that the protocol will act to reduce entanglement. While it increases the intra-mode noise (the diagonal elements of the covariance matrix), there is no compensating increase in the inter-mode correlations, and so the overall quality of the entanglement is decreased. Let us put this thought aside for now, though, and instead ask how good this representation of the input state is overall.

5.4.1.1 *Fidelity for pure 2-mode states*

The fidelity between a pure 2-mode state and a generally-mixed output with equal mean vector can be found as the overlap

$$\mathcal{F} = \langle\psi|\hat{\rho}_{\text{out}}|\psi\rangle = \frac{4}{\sqrt{\det(V_{\psi} + V_{\text{out}})}}. \tag{211}$$

For the output states described in equations (209) and (210), then, a reconstruction fidelity of

$$\mathcal{F} = \begin{cases} 4/[(2n + \eta^2 E^{r1}_{1|2}(g) + 1 - \eta^2)(2m + \eta^2 E^{r2}_{1|2}(g) + 1 - \eta^2) - 4c^2] & g \leq 1 \\ 4/[(2n + E^{r1}_{1|2}(g) + 1 - 1/\eta^2)(2m + E^{r2}_{1|2}(g) + 1 - 1/\eta^2) - 4c^2] & g \geq 1, \end{cases} \tag{212}$$

is achievable, where $c = \sqrt{nm-1}$ for pure input states and as usual $\eta = 1/\sqrt{2-g^2}$.

As might be expected, each mode's contribution to the fidelity depends only on the resource state used in its single-mode protocol. With no entanglement between the resource states, the protocol has no effect on the inter-mode covariance parameter c and it is reproduced with no change.[11] As we will see in section 5.4.1.2, though, this will still have the effect of reducing the entanglement.

It is possible, and perhaps in some scenarios desirable,[12] to use resource states with different properties to share each mode; in this case the resource state used for a given mode will of course impact *only* that mode and so the reconstruction quality of each mode may differ.

Let us focus on the case in which the resource states used to share each mode are equivalent such that $E^{r1}_{1|2}(g) = E^{r2}_{1|2}(g) := E_{1|2}(g)$. As the dealer should have full control over the properties of each resource state, it is imagined that this would be the ordinary case. This assumption simplifies the fidelity to the form

$$\mathcal{F} = \begin{cases} 4\Big/\Big(3 + \big[1 - E_{1|2}(g)\big]^2\eta^4 + 2\big[n+m+1\big]\big[\eta^2 E_{1|2}(g) + 1 - \eta^2)\big]\Big) & g \leq 1 \\ 4\Big/\Big(3 + \big[\frac{1}{\eta^2} - E_{1|2}(g)\big]^2 + 2\big[n+m+1\big]\big[E_{1|2}(g) + 1 - \frac{1}{\eta^2}\big]\Big) & g \geq 1, \end{cases} \tag{213}$$

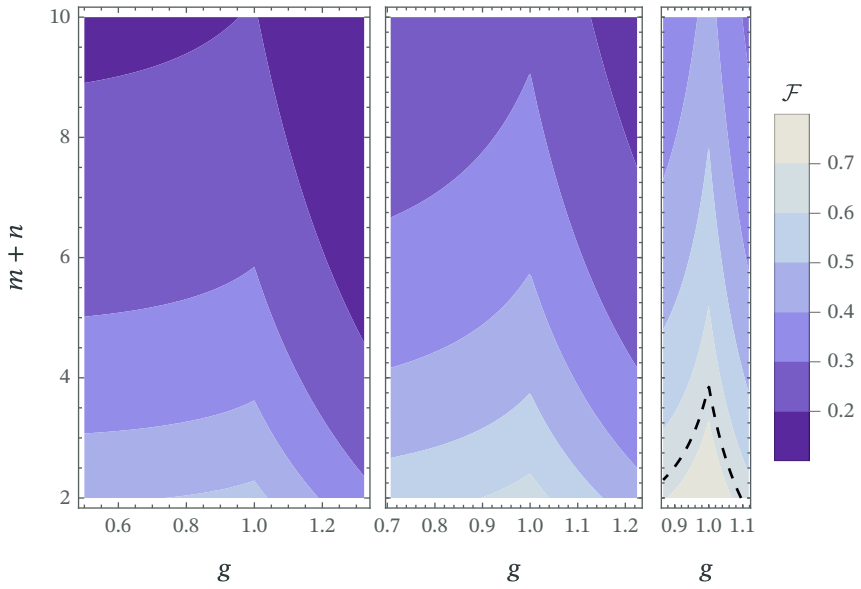
where we have also accounted for the purity condition that $c^2 = nm - 1$. Aside from the choice of resource state, this reconstruction quality depends solely on the sum of the individual mode variances, $n+m$ — any asymmetry in them does not appear to impact the quality of the reconstruction. This fidelity is shown for varying $E_{1|2}(g)$ in figure 16.

Much of this figure reflects what was seen when we studied single-mode states in section 5.3. Although the fidelity is predominantly dependent on the properties of the secret state, there remains a strong peak in achievable fidelity for $g = 1$. As was previously the case, excellent reconstruction fidelity can continue to be found across the $g < 1$ region, while it drops much more rapidly as g rises above 1.

The achievable fidelity depends strongly on the degree of entanglement present within the secret state. Recalling that we are considering here only pure input states, in which inter-mode entanglement is proportional to the size of the intra-mode variances, this dependence on the sum $m + n$

[11] This does not need to be the case. If the entanglement structure within the secret state is public knowledge the resource states could be designed with a complimentary entanglement structure that better preserves the secret.

[12] Although one must admit no such scenarios immediately spring to mind.



(a) $E_{1|2}(g) = 0.75$ (b) $E_{1|2}(g) = 0.5$ (c) $E_{1|2}(g) = 0.25$

Figure 16: Reconstruction fidelity achievable for QSS of a pure two-mode Gaussian state using a Gaussian resource state with steering parameter (a) $E_{1|2}(g) = 0.75$, (b) $E_{1|2}(g) = 0.5$, (c) $E_{1|2}(g) = 0.25$ for varying reconstruction parameter g. Here, $n + m$ is the sum of the symplectic invariants describing each mode of the input state, which acts as a proxy for the amount of entanglement present in the state. Only those g values for which the indicated steering is physical are shown. Dashed black line indicates $\mathcal{F} = 2/3$ as a benchmark only; exceeding this threshold does not ensure security for multi-mode states.

acts as a rough proxy for the total entanglement present in the state. We can immediately conclude then that higher-entangled multi-mode states require greater resource entanglement to successfully be shared.

This is not a surprising result. We saw when we studied squeezed states in section 5.3.1.2 and thermal states in section 5.3.2.1 that increasing the 'quantumness' of the secret state (through, for example, increasing squeezing) increases the resource requirements, while decreasing the 'quantumness' (through, for example, increasing thermality) decreases it. That increasing secret-state entanglement comes with increased resource entanglement requirements is simply a continuation of this trend.

Let us now focus on the optimal case in which the protocol is performed for reconstruction parameter $g = 1$; that is, when no amplification correc-

tion is required. In this case, the achievable reconstruction fidelity reduces to

$$\mathcal{F} = \frac{4}{[2n + E^{r1}_{1|2}(g)][2m + E^{r2}_{1|2}(g)] - 4c^2} \tag{214}$$

$$= \frac{4}{4 + 2(m+n)E_{1|2}(g) + E_{1|2}(g)^2}, \tag{215}$$

where the two resource states used are assumed to be different in the first line and identical in the second. This fidelity is shown in figure 17.

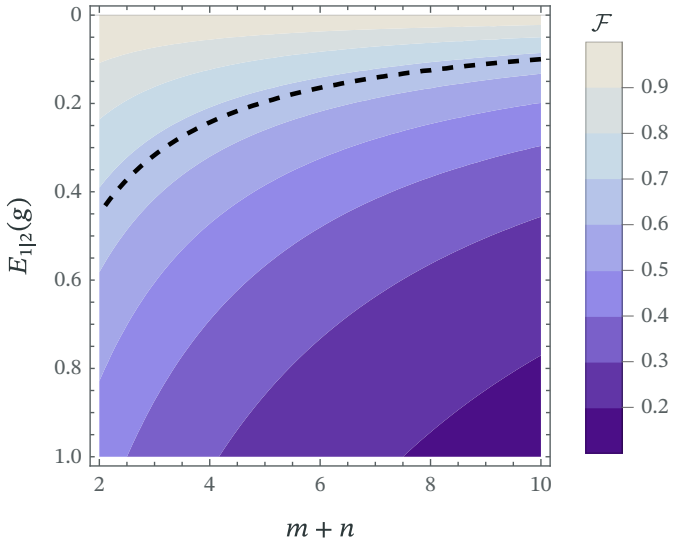


Figure 17: Achievable reconstruction fidelity for a 2-mode Gaussian secret state with single-mode variances n and m using a QSS protocol with reconstruction parameter $g = 1$ and two resource states with steering $E_{1|2}(g)$ as shown. Dashed black line indicated $\mathcal{F} = 2/3$ as a benchmark only; exceeding this threshold does not guarantee security for multi-mode states.

As we have already seen, the achievable fidelity decreases rapidly as the entanglement present in the state increases, with greater resource steering required to maintain a good fidelity. In the limit of perfect entanglement (as $E_{1|2}(g) \to 1$), though, we can see that this protocol always approaches perfect reconstruction $\mathcal{F} \to 1$. Although this is not an achievable limit, it does indicate that reconstruction fidelity is limited only by the quality of the resource available: arbitrarily high reconstruction fidelity can be achieved for *any* pure 2-mode secret state given suitably good-quality entanglement.

5.4.1.2 *Entanglement preservation*

Reconstruction fidelity is a very blunt measure of success, however, as it represents the average reconstruction across all features of the state. While

this is a useful measure of overall reconstruction quality, in the context of multimode secrets it is worth additionally asking a more specific question: how much of the inter-mode entanglement structures survive the process?

To match the measure of resource entanglement we have used throughout our QSS discussion, let us quantify the secret state correlations through their EPR steering parameter, defined by

$$E_{1|2}(\gamma) = \Delta^2(\hat{X}_1^+ - \gamma\hat{X}_2^+) = \Delta^2(\hat{X}_1^- + \gamma\hat{X}_2^-), \tag{216}$$

which for a two-mode Gaussian state in the standard form given by equation (205) is defined as

$$E_{1|2}(\gamma) = n + \gamma^2 m - 2\gamma c, \tag{217}$$
$$E_{2|1}(\gamma) = m + \gamma^2 n - 2\gamma c, \tag{218}$$

where we have used γ instead of the customary g to avoid confusion with the equivalent parameter describing the resource state.

Applying this formula to the output states we found in equations (209) and (210), the output state will exhibit steering in each direction of

$$E_{1|2}(\gamma) = \begin{cases} E^{\text{in}}_{1|2}(\gamma) + \eta^2\left[E^{r1}_{1|2}(g) + \gamma^2 E^{r2}_{1|2}(g)\right] + (1+\gamma^2)(1-\eta^2) & g \leq 1 \\ E^{\text{in}}_{1|2}(\gamma) + E^{r1}_{1|2}(g) + \gamma^2 E^{r2}_{1|2}(g) + (1+\gamma^2)(1-\frac{1}{\eta^2}) & g \geq 1, \end{cases} \tag{219}$$

and

$$E_{2|1}(\gamma) = \begin{cases} E^{\text{in}}_{2|1}(\gamma) + \eta^2\left[\gamma^2 E^{r1}_{1|2}(g) + E^{r2}_{1|2}(g)\right] + (1+\gamma^2)(1-\eta^2) & g \leq 1 \\ E^{\text{in}}_{2|1}(\gamma) + \gamma^2 E^{r1}_{1|2}(g) + E^{r2}_{1|2}(g) + (1+\gamma^2)(1-\frac{1}{\eta^2}) & g \geq 1. \end{cases} \tag{220}$$

For better clarity, let us consider the special case in which $g = 1$; even in this optimal case, the level of steering present in the output state is described by

$$E_{1|2}(\gamma) = E^{\text{in}}_{1|2}(\gamma) + E^{r1}_{1|2}(g) + \gamma^2 E^{r2}_{1|2}(g). \tag{221}$$

The steering present in our input state has been reduced to the sum of the steering parameters of all three states — the input state plus the two resource states! Recalling that a lower steering parameter $E_{1|2}(\gamma) \in [0,1]$ implies a greater degree of entanglement, this is a discouraging result. It is clear that the application of this protocol effectively destroys entanglement

between the modes, even when both resource states are presumed to be relatively high-entanglement.

Without knowledge of the underlying entanglement structures, this extension of quantum state sharing to multi-mode states is not suitable for the sharing of entanglement.

5.4.2 *Worst-case fidelity improvement through permutation of shares*

Although we do not consider the question here of whether more bespoke schemes for multi-mode state sharing exist, a remarkably simple administrative change can drastically improve the reconstruction fidelity for secret states of three or more modes.

Recall that in section 4.2 we found that different reconstruction protocols were required for different modes, and further that when one had access to the first two output shares perfect reconstruction could always be achieved regardless of the resource state. We have until now neglected to analyse this case in great detail as it is only available to a single pair of shares.

This asymmetry exists because one of the three output modes, and hence the share owned by one of the players, does not directly contain any information about ψ and only consists of the auxiliary resource mode. When we have a 3-mode secret state, however, we need not select a single 'unlucky' player who is given the poor-quality share for every mode. Rather, we can allocate each player two higher-quality shares and a single poorer-quality share. For a 3-mode secret this ensures that every possible pair of players will reconstruct a single mode perfectly and two modes imperfectly. Although this will not impact the average fidelity across potential reconstructions, it will increase the worst-case reconstruction.

By iterating which player is given the 'bad' share each mode, then, the protocol effectively shares every third mode 'for free'!

5.4.3 *Further study in this area*

As well as improving the achievable reconstruction fidelity through permuting shares, there remain other open questions regarding the possibility for multi-mode QSS. The first is simply whether a better protocol can be found: here, we have designed a protocol for single-mode states and simply applied it twice to multi-mode states. It has previously been shown that for the task of cloning entangled states a 'global' protocol which acts on both modes at once is superior to performing two 'local' protocols on each mode [22]. This is likely to be true for QSS also, so it may well be that reimagining

the protocol from the ground up results in improvements in reconstruction effectiveness.

Even without changing the protocol, there additionally remains the question of the choice of resource state. We have here used two wholly independent 2-mode resource states to share our 2-mode secret. However, there is significant research interest at present in the properties and usefulness of multi-mode entanglement [65, 66]. It may well be the case that, by optimising the fidelity over the space of all possible 4-mode resource states, a better fidelity can be achieved. This could be investigated, for example, by running a numerical optimisation over the space of 4-mode resource states.

5.5 CONCLUSION

In this chapter we have considered in detail the use of our quantum state sharing protocol for the sharing of Gaussian state secrets. We have shown that for a coherent-state secret, security can be guaranteed when the resource state used demonstrates *any* degree of steering, including states only steerable in one direction. Notably, this is a much looser requirement than that for secure quantum teleportation, where bidirectional steering is required, and reflects the generally better-quality output from QSS with asymmetric resource states. In addition to cases in which splitting secrets into multiple parts is directly desirable, then, QSS may find uses as a 'cheaper' alternative to teleportation for secure communication when a security level of *secure under the assumption an attacker has access to only one of three transmission mediums* is acceptable. While perhaps not useful for high-risk scenarios against highly-capable bad actors, say in military or diplomatic communication, this level of security may be considered a reasonable tradeoff for increasing output state fidelity in less risky areas.

Looking at this protocol's use for squeezed coherent states — and thus *any* pure single-mode Gaussian state — we have found that the protocol remains highly effective. Indeed, we have shown that it is provably secure for any set of squeezed input states given a suitably-entangled resource.

Finally, we have shown that this protocol is effective for the sharing of mixed and multi-mode states. Although we are not able to derive a security condition for such states due to the lack of a provably-optimal cloning strategy, the arbitrarily-good reconstruction fidelity achievable for them indicates that secure QSS is possible here regardless of what the no-cloning threshold is, so long as a suitably-entangled resource state is available.

6

INTERLUDE: MODELLING HYBRID FOCK-GAUSSIAN PROCESSES

The use of Gaussian entanglement for the distribution of Fock states offers the potential to sidestep the difficult process of generating discrete-variable entanglement [67]. We will consider the use of our quantum state sharing protocol for such states in the next chapter. Let us for now, though, pause our discussion of quantum state sharing and consider the effect of a general Gaussian channel on Fock-like states.

The standard approach to the modelling of interaction between Gaussian and Fock modes has previously been to either operate in the Fock formalism [68–70], or to fall back on numerical models. Neither of these approaches are particularly satisfying; although continuous-variable states can be well-represented as infinite summations, for example as

$$|\alpha\rangle = \mathrm{e}^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \tag{222}$$

for a coherent state [16], the inclusion of more than one or two auxiliary Gaussian modes (or, indeed, more complex states) quickly makes this a non-trivial task. The use of numerical integrations, while an extremely powerful tool when all state parameters are specified, do not allow for general analytic results.

Some prior work has occurred investigating the impact of a Gaussian channel on Fock inputs analytically. Ivan *et al.* in 2011 [71] described an algorithmic method to convert a Gaussian channel from its phase space description to a set of Fock-space Kraus operators. However, applying this process to a specific Gaussian channel is in-general nontrivial, as shown by the examples given in that paper. Similarly, Caves *et al.* in 2004 [72] discussed the output fidelity achievable from an arbitrary Gaussian channel given knowledge of the thermal noise it imparts. The approach taken in that paper, however, is only applicable to the study of the output fidelity and does not describe the form of the output state itself.

In this chapter, we outline the approach we will take to model the impact of our QSS protocol on Fock states. In section 6.1 we briefly recap some useful properties of Gaussian channels, and define the subset of such channels that our approach works for. In section 6.2 we present a novel

approach to the study of Gaussian channels acting on Fock eigenstates, working entirely in the phase space formalism. By solving the integral for the partial trace over an arbitrary Gaussian channel, a general description of the output state is presented in theorem 10 for both the Wigner and density-matrix formalisms. Although we do not present a similar general solution for the Fock superposition states, in section 6.3 we discuss the modelling of such states and outline an algorithm to model the impact of a Gaussian channel on them.

6.1 GAUSSIAN CHANNELS AND $(x-p)$-BALANCE

Consider a Gaussian channel $\hat{\Lambda}$ that maps Gaussian states to Gaussian states but that does not preserve state purity. The standard way to model such a purity-decreasing channel is to include an additional noise contribution as

$$\bar{\boldsymbol{r}} \mapsto T\bar{\boldsymbol{r}}, \qquad V \mapsto TVT^T + N, \tag{223}$$

for some T, $N \in \mathbb{R}^{2\times 2}$ representing $\hat{\Lambda}$.

This channel can also, however, be represented as some equivalent unitary, $\hat{U}_\Lambda$, acting on an enlarged system comprising the input mode alongside some N auxiliary modes representing the environment. The single-mode output can then found by tracing out these inaccessible environment modes, as

$$\hat{\Lambda}(\hat{\rho}) = \mathrm{Tr}_{2\dots N+1}\Big[\hat{U}_\Lambda\big(\hat{\rho}\otimes|G(0,V)\rangle\langle G(0,V)|\big)\Big]. \tag{224}$$

This unitary is characterised by a symplectic matrix, $\Lambda \in \mathrm{Sp}(2(N+1);\mathbb{R})$, describing the transformation of the system as [73]

$$W_{\text{out}}(\boldsymbol{q}) = W_{\text{in}}(\Lambda\cdot\boldsymbol{q}), \tag{225}$$

for $\boldsymbol{q} = (x_1, x_2, \dots, x_{N+1}, p_1, p_2, \dots p_{N+1})^T$ the quadrature coordinates. In contrast to the rest of thesis, for the remainder of this chapter we will assume that the quadratures are 'xp-ordered' such that the x quadrature variables for all modes precede the p quadratures.

Let us here restrict our discussion to those unitaries that act on these two classes of quadrature separately, such that the transformation matrix can be written

$$\Lambda = \begin{pmatrix} \Lambda_x & 0 \\ 0 & \Lambda_p \end{pmatrix}. \tag{226}$$

Assuming the environment system is zero-mean,[1] the initial quantum channel $\hat{\Lambda}$ is wholly characterised by these transformation matrices[2] Λ_x and Λ_p and the initial state of the environment system, V. The solutions we present in this chapter are valid only for Gaussian channels displaying some degree of symmetry between the quadrature dynamics, which we term '$(x-p)$-balanced' channels and specify in definition 9.

Definition 9. *A quantum protocol $\hat{\Lambda}$ acting on one information mode, ψ, and $N-1$ auxiliary modes is* $(x-p)$-balanced *if, after the protocol, the two quadratures representing the output mode have*

1. *the same variance overall, and*
2. *identical contributions from the respective quadratures of the input mode, $\hat{X}^{\pm}_{\psi}$.*

When the information mode is a Fock state or unsqueezed coherent state and the auxiliary modes a Gaussian state with covariance matrix $V = V_x \oplus V_p$, this is equivalent to the conditions

$$(\Lambda_x)_{1,1} = (\Lambda_p)_{1,1}, \textit{and} \tag{227}$$

$$[\Lambda_x^{-1}(1 \oplus V_x)(\Lambda_x^{-1})^T]_{1,1} = [\Lambda_p^{-1}(1 \oplus V_p)(\Lambda_p^{-1})^T]_{1,1}, \tag{228}$$

for $\Lambda_{x/p}$ the matrices describing the evolution of the x/p quadratures under the unitary, and V_x/V_p the covariance matrix for the auxiliary Gaussian modes.

6.2 GAUSSIAN CHANNELS ACTING ON FOCK EIGENSTATES

Let us imagine now that the input to the channel $\hat{\Lambda}$ consists of a Fock state of known particle number,

$$\hat{\rho}_{\text{in}} = |n\rangle\langle n|, \tag{229}$$

[1] which noise sources tend to be

[2] In fact, we need only one of the transformation matrices to characterise the unitary as the symplectic nature of the channel dictates that $\Lambda_p^T = \Lambda_x^{-1}$.

which will again be transformed by the channel to produce an output state,

$$\hat{\rho}_{\text{out}} = \text{Tr}_{2\ldots N+1}[\hat{\Lambda}(|n\rangle\langle n| \otimes |G(0,V)\rangle\langle G(0,V)|)]. \tag{230}$$

The change in the system from the application of this unitary is well understood and trivial to implement — the Wigner function evolves as a coordinate transform exactly as described in equation (225). Tracing out the environment modes to understand the state of the single-mode output, however, is not such a simple task when the system is comprised of non-Gaussian components. It is this latter task we primarily consider here.

6.2.1 *Overview of approach*

Although the details underpinning this framework are left to appendix B, let us here give a brief overview of the approach we take.

The Wigner function describing the Fock state with known particle number n that forms the input to the channel, $|n\rangle\langle n|$ is given by [74]

$$W_n(x,p) = \frac{(-1)^n}{\pi}\exp\Big[-(x^2+p^2)\Big]L_n\Big[2(x^2+p^2)\Big], \tag{231}$$

for $L_n(x)$ the nth Laguerre polynomial[3]

$$L_n(x) = \sum_{k=0}^{n}\binom{n}{k}\frac{(-1)^k}{k!}x^k. \tag{232}$$

The N-mode auxiliary Gaussian state, meanwhile, has Wigner Function [73]

$$W_G(\boldsymbol{x},\boldsymbol{p}) = \frac{1}{\pi^N\sqrt{\det V}}\exp\Big(-\boldsymbol{x}^T V_x^{-1}\boldsymbol{x} - \boldsymbol{p}^T V_p^{-1}\boldsymbol{p}\Big), \tag{233}$$

for $V = V_x \oplus V_p$ the matrix direct sum of the x and p covariance matrices. The collective system acting as input to the Gaussian unitary, consisting of the tensor product of these states, is given simply by the product of the Wigner functions,

$$\begin{aligned} W_{\text{in}}(\boldsymbol{x},\boldsymbol{p}) =& W_n(x_1,p_1)W_G(x_2,\ldots,x_N,p_2,\ldots,p_N) && (234)\\ =& \frac{(-1)^n}{\pi^{N+1}\sqrt{\det V}}L_n\Big[2\boldsymbol{x}^T(1\oplus 0_N)\boldsymbol{x} + 2\boldsymbol{p}^T(1\oplus 0_N)\boldsymbol{p}\Big] \\ & \times\exp\Big[-\boldsymbol{x}^T(1\oplus V_x)^{-1}\boldsymbol{x} - \boldsymbol{p}^T(1\oplus V_p)^{-1}\boldsymbol{p}\Big], && (235)\end{aligned}$$

[3] See equation (18.5.E12) in Ref. [75]

for 0_n the $n \times n$ matrix with all zero elements.[4]

6.2.1.1 *State of the wider system post-operation*

The Wigner function describing this system after the application of the unitary is given by the coordinate transform,

$$W_{\text{out}}(\boldsymbol{x}, \boldsymbol{p}) = W_{\text{in}}(\Lambda_x \cdot \boldsymbol{x}, \Lambda_p \cdot \boldsymbol{p}) \tag{236}$$

$$= \frac{(-1)^n}{\pi^{N+1}\sqrt{\det V}} L_n\Big[2\boldsymbol{x}^T \Lambda_x^T (1 \oplus 0_N)\Lambda_x \boldsymbol{x} + 2\boldsymbol{p}^T \Lambda_p^T (1 \oplus 0_N)\Lambda_p \boldsymbol{p}\Big] \times \exp\Big[-\boldsymbol{x}^T \Lambda_x^T (1 \oplus V_x)^{-1}\Lambda_x \boldsymbol{x} - \boldsymbol{p}^T \Lambda_p^T (1 \oplus V_p)^{-1}\Lambda_p \boldsymbol{p}\Big], \tag{237}$$

which we can write simply as

$$W_{\text{out}}(\boldsymbol{x}, \boldsymbol{p}) = \frac{(-1)^n}{\pi^{N+1}\sqrt{\det V}} L_n\Big[2(\lambda_x \cdot \boldsymbol{x})^2 + 2(\lambda_p \cdot \boldsymbol{p})^2\Big] \times \exp\Big[-\boldsymbol{x}^T \tilde{V}_x^{-1} \boldsymbol{x} - \boldsymbol{p}^T \tilde{V}_p^{-1} \boldsymbol{p}\Big] \tag{238}$$

by noting that

$$\boldsymbol{x}^T \Lambda_x^T (1 \oplus 0_N)\Lambda_x \boldsymbol{x} = \sum_{i,j} x_i (\Lambda_x^T)_{i,1} (\Lambda_x)_{1,j} x_j = (\boldsymbol{\lambda}_x \cdot \boldsymbol{x})^2, \tag{239}$$

for $\boldsymbol{\lambda}_{x/p}$ the first row of $\Lambda_{x/p}$, and defining

$$\tilde{V} = \Lambda^{-1}(I_2 \oplus V)(\Lambda^{-1})^T. \tag{240}$$

This Wigner function represents a full characterisation of the wider state-environment system. In this form, though, it tells us very little about the state of the single output mode of the Gaussian channel that we have access to, which is likely to the be the object of most interest. To find the single-mode output of the channel, we must trace out the remaining auxiliary modes; a task we consider in the next subsection.

6.2.1.2 *Tracing over a joint Fock-Gaussian system*

The form of a single subsystem of a wider multi-mode state, $W_{\text{sys}}(\boldsymbol{q})$, can be found by tracing out any unwanted modes. In the Wigner formalism, this is achieved by integrating over those modes' quadratures as [16]

$$W_{\text{subssys}}(\boldsymbol{q}_1) = \int_{\mathbb{R}^{2N}} d\boldsymbol{q}_{2,\ldots,N+1} W_{\text{sys}}(\boldsymbol{q}), \tag{241}$$

[4] Note that $(A \oplus B)^{-1} = A^{-1} \oplus B^{-1}$ so $1 \oplus V^{-1} = (1 \oplus V)^{-1}$.

where $\boldsymbol{q}_i$ are the quadratures representing modes i. In the wholly Gaussian case, this is a trivial and well-known Gaussian integral that reduces to the matrix transforms we have discussed in section 2.1.2.

For a system consisting of both Gaussian and Fock modes, mixed according to some unitary $\hat{U}_\Lambda$, the integral is complicated by the additional presence of a Laguerre polynomial,

$$W_{\psi'} = \frac{(-1)^n}{\pi^{N+1}\sqrt{\det V}} \int_{\mathbb{R}^{2N}} d\boldsymbol{x}' d\boldsymbol{p}' \left\{ \begin{array}{c} L_n\Big[2(\boldsymbol{\lambda}_x \cdot \boldsymbol{x})^2 + 2(\boldsymbol{\lambda}_p \cdot \boldsymbol{p})^2\Big] \\ \times \exp\Big[-\boldsymbol{x}^T \tilde{V}_x^{-1} \boldsymbol{x} - \boldsymbol{p}^T \tilde{V}_p^{-1} \boldsymbol{p}\Big] \end{array} \right\} \tag{242}$$

and is not immediately trivial to evaluate.

We tackle this integral in appendix B by first solving the foundational integral,[5]

$$\int_{\mathbb{R}^N} d^N x \left(\prod_{i \in \beta} x_i \right) \exp(-\boldsymbol{x}^T V^{-1} \boldsymbol{x} + \boldsymbol{a}^T \boldsymbol{x}), \tag{243}$$

for some arbitrary vector of indices β_i in theorem B.3. By exploiting the symmetry inherent to Laguerre polynomials and the powers of vector-products, this integral can be used to construct the full integral of equation (242) through a number of stepping-stone integrals.

The reader is directed to appendix B for a full derivation of these integrals; let us here instead simply note the structure of this derivation. By expanding the Laguerre polynomial to the summations

$$\begin{aligned} & L_n\Big[2(\boldsymbol{\lambda}_x \cdot \boldsymbol{x})^2 + 2(\boldsymbol{\lambda}_p \cdot \boldsymbol{p})^2\Big] \\ &= \sum_{a=0}^{n} \binom{n}{a} \frac{(-1)^a}{a!} 2^a \Big[(\boldsymbol{\lambda}_x \cdot \boldsymbol{x})^2 + (\boldsymbol{\lambda}_p \cdot \boldsymbol{p})^2\Big]^a \qquad (244) \\ &= \sum_{a=0}^{n} \binom{n}{a} \frac{(-1)^a}{a!} 2^a \sum_{b=0}^{a} \binom{a}{b} (\boldsymbol{\lambda}_x \cdot \boldsymbol{x})^{2b} (\boldsymbol{\lambda}_p \cdot \boldsymbol{p})^{2a-b}, \qquad (245) \end{aligned}$$

we can build up to the solution to the Laguerre polynomial integral by using theorem B.3 to solve the integral of each of these 'blocks' in turn. We

[5] This integral can be recognised as a generalisation of the known integral underpinning Wick's theorem given by [76]

$$\int_{\mathbb{R}^N} d^N x \left(\prod_{i \in \beta} x_i \right) \exp(\boldsymbol{x}^T V^{-1} \boldsymbol{x}) = \begin{cases} 0 & n \text{ odd} \\ \dfrac{\pi^{N/2}}{2^{n/2}} \sqrt{\det V} \displaystyle\sum_{\sigma \in P(\beta)} V_{\sigma_0,\sigma_1}, \ldots, V_{\sigma_{n-2},\sigma_{n-1}} & n \text{ even,} \end{cases}$$

for $n = |\beta|$ and $P(\beta)$ the set of all pairings of β. To enable us to perform the integral in equation (242) over only a subset of the x_i indices and not the whole $\boldsymbol{x}$ space (i.e. to perform a *partial* trace), we must introduce the additional linear term $\boldsymbol{a}^T \boldsymbol{x}$.

first solve for the integral of a Gaussian multiplying the power of a vector dot product,

$$\int_{\mathbb{R}^N} d^N x (\boldsymbol{\lambda}_x \cdot \boldsymbol{x})^n \exp(-\boldsymbol{x}^T V^{-1} \boldsymbol{x} + \boldsymbol{a}^T \boldsymbol{x}), \tag{246}$$

in theorem B.5 before extending this to polynomials of both quadratures of the form $[(\boldsymbol{\lambda}_x \cdot \boldsymbol{x})^2 + (\boldsymbol{\lambda}_p \cdot \boldsymbol{p})^2]^n$ in theorem B.7. This foundation allows us to finally solve for an integral of the general form

$$\int_{\mathbb{R}^{2N}} d\boldsymbol{x}' d\boldsymbol{p}' \, L_n\Big[2(\boldsymbol{\lambda}_x \cdot \boldsymbol{x})^2 + 2(\boldsymbol{\lambda}_p \cdot \boldsymbol{p})^2\Big] \exp\Big[-\boldsymbol{x}^T V_x^{-1} \boldsymbol{x} - \boldsymbol{p}^T V_p^{-1} \boldsymbol{p})\Big], \tag{247}$$

under some x-p symmetry conditions, in theorem B.8. This integral is immediately applicable to our use, but in the interest of simplicity we use the knowledge that V is a symmetric covariance matrix and Λ is symplectic to specify the result further in corollary B.11 and theorem B.14.

6.2.2 *Statement of theorem*

The integrals solved in appendix B allow us to present the following general theorem for the output of a Gaussian quantum protocol acting on a Fock state.

Theorem 10. *The first output mode of a symplectic, $(x-p)$-balanced (see definition 9) quantum unitary acting on a Fock state $|n\rangle$ and zero-mean Gaussian auxiliary state $|G(0, V_r)\rangle$ as*

$$\hat{\rho}_{out} = \mathrm{Tr}_{2,\ldots}\Big[\hat{U}_\Lambda(\hat{\rho}_n \otimes \hat{\rho}_{G(0,V_r)})\Big], \tag{248}$$

is given in the Wigner formalism by

$$W_{out}(x,p) = \frac{(-1)^n}{\pi} \frac{(\eta-\upsilon)^n}{(\eta+\upsilon)^{n+1}} L_n\Big[\frac{2\eta}{\eta^2-\upsilon^2}(x^2+p^2)\Big] \exp\Big[-\frac{1}{\eta+\upsilon}(x^2+p^2)\Big], \tag{249}$$

or by the diagonal density matrix in Fock space

$$\hat{\rho}_{out} = \sum_m C_m |m\rangle\langle m| \tag{250}$$

for

$$C_m = \frac{2}{(\upsilon+\eta+1)^{n+m+1}} \sum_{b=0}^{\min(m,n)} \binom{n}{b}\binom{m}{b} (4\eta)^b (\upsilon+\eta-1)^{m-b} (\upsilon-\eta+1)^{n-b} \tag{251}$$

Here,

$$\upsilon = \Lambda_{pIJ}^T \cdot V_{rx} \cdot \Lambda_{pIJ} = \Lambda_{xIJ}^T \cdot V_{rp} \cdot \Lambda_{xIJ}, \tag{252}$$

$$\eta = [(\Lambda_x)_{1,1}]^2 = [(\Lambda_p)_{1,1}]^2, \tag{253}$$

for Λ the transformation matrix representing the unitary.

Proof. This theorem is simply a restatement of theorem B.14 from appendix B in ket notation, noting that $\Lambda_{\mathcal{J}} = \sqrt{\eta}$ and $V_{\mathcal{J}} = \eta + \upsilon$.

□

Finally, noting that the Hilbert-Schmidt overlap, or fidelity, is given simply by the $m = n$ element of the density matrix as

$$\mathcal{F} = \langle n|\hat{\rho}_{\text{out}}|n\rangle = \sum_m C_m \langle n|m\rangle\langle m|n\rangle = C_n, \tag{254}$$

we can immediately state the following corollary of theorem 10.

Corollary 11. *The fidelity between the first output mode of a symplectic, $(x-p)$-balanced (see definition 9) quantum unitary acting on a Fock state and a zero-mean Gaussian auxiliary state,*

$$\text{Tr}_{2\ldots N}\Big[\hat{U}_\Lambda\big(|n\rangle|G(0, V_r)\rangle\big)\Big], \tag{255}$$

and the original Fock state is given by

$$\mathcal{F} = 2\frac{\left[\upsilon^2 - (1-\eta)^2\right]^n}{(\eta+\upsilon+1)^{2n+1}} \sum_{b=0}^{n} \binom{n}{b}^2 \left[\frac{4\eta}{\upsilon^2 - (1-\eta)^2}\right]^b, \tag{256}$$

for υ, η as defined in theorem 10.

6.2.3 *Example: thermal attenuating channels*

To illustrate this method, let us now consider the action of a canonical Gaussian channel: linear attenuation in the presence of thermal noise. This channel represents some portion $\epsilon < 1$ of the signal — here the Fock state — being lost and replaced by thermal environment noise, and can be modelled simply as a beamsplitter with a thermal environment.

The coordinate transform for an attenuating channel is

$$\Lambda_x = \Lambda_p = \begin{pmatrix} \sqrt{1-\epsilon} & \sqrt{\epsilon} \\ -\sqrt{\epsilon} & \sqrt{1-\epsilon} \end{pmatrix}, \tag{257}$$

acting on the system $\hat{\rho}_n \otimes \hat{\rho}_{\text{th}}$ for $\hat{\rho}_{\text{th}}$ a Gaussian thermal state with covariance matrix $V_{\text{x}} = V_{\text{p}} = (1 + 2\bar{n})I_2$ for $\bar{n}$ representing the temperature of the environment state. In the language of theorem 10, then,

$$\eta = 1 - \epsilon, \quad \upsilon = \epsilon(2\bar{n} + 1). \tag{258}$$

When the environment mode is traced out, theorem 10 states that the output state will be described by the Wigner function

$$\begin{aligned} W_{\text{out}}(x, p) = &\frac{(-1)^n}{\pi(1 + 2\bar{n}\epsilon)}\Big[\frac{2(1-\epsilon)}{1 + 2\bar{n}\epsilon} - 1\Big]^n \\ &\times L_n\Big[\frac{2(1-\epsilon)}{(1-\epsilon)^2 - (2\bar{n}+1)^2\epsilon^2}(x^2 + p^2)\Big] \\ &\times \exp\Big[-\frac{1}{1 + 2\bar{n}\epsilon}(x^2 + p^2)\Big], \end{aligned} \tag{259}$$

or by the density matrix $\hat{\rho}_{\text{out}} = C_m|m\rangle\langle m|$ with diagonal elements

$$C_m = \frac{1}{(1 + \bar{n}\epsilon)^{n+m+1}} \sum_{b=0}^{\min(m,n)} \binom{n}{b}\binom{m}{b}(1-\epsilon)^b(\bar{n}\epsilon)^{m-b}((1+\bar{n})\epsilon)^{n-b}. \tag{260}$$

In the special case of ideal de-amplification (for which $\bar{n} = 0$), the output simplifies to

$$W_{\text{out}} = \frac{(-1)^n(1-2\epsilon)^n}{\pi} L_n\Big[\frac{1-\epsilon}{1-2\epsilon}(2x^2 + 2p^2)\Big] \exp\Big[-(x^2 + p^2)\Big], \tag{261}$$

and

$$C_m = \sum_{b=0}^{\min(m,n)} \binom{n}{b}\binom{m}{b}(1-\epsilon)^b[0]^{m-b}\epsilon^{n-b} \tag{262}$$

$$= \begin{cases} \binom{n}{m}(1-\epsilon)^m\epsilon^{n-m} & m \le n \\ 0 & m > n \end{cases}. \tag{263}$$

The output of an ideal amplitude de-amplification operation on a Fock state, then, is a weighted distribution of all Fock states of lower energy than the original as might be expected from a purely de-amplying process. When the environment consists of thermal excitations above the vacuum, the de-amplification no longer results in a probability distribution of strictly lower-energy states. The thermal noise acts on each of the potential outputs by randomly displacing their phase-space amplitude according to a zero-mean Gaussian distribution. Consequently, although the output state probability remains centred around the pure deamplified Fock state, there exists a possibility that even very high-energy states may be found.

Although these are not novel results, the application of theorem 10 makes the output of this channel immediately accessible.

6.3 GAUSSIAN CHANNELS ACTING ON FOCK SUPERPOSITION STATES

Finally, before returning to our discussion of quantum state sharing, let us consider the modelling of the impact of Gaussian channels on superpositions of the Fock eigenstates. These are states of the form

$$|\psi\rangle = \sum_n \alpha_n |n\rangle, \tag{264}$$

with $\sum |\alpha_n|^2 = 1$.

In the density matrix formalism, these states consist not only of diagonal $|n\rangle\langle n|$ eigenstate components but also of coherences between the eigenstates of the form $|n\rangle\langle m|$. We have studied in the previous section the impact of a Gaussian channel on the diagonal components, so it remains here to consider the impact on the coherences.

Recall from section 2.1.2 the Wigner function describing such superposition states is given by the sum of the eigenstate Wigner functions and coherence terms as

$$W_\psi(x,p) = \sum_n |\alpha_n|^2 W_n(x,p) + \sum_{n\neq m} \alpha_n^\dagger \alpha_m I_{n,m}(x,p). \tag{265}$$

For two Fock states, these coherence terms are given by the expression [77]

$$I_{n,m}(x,p) = \frac{(-1)^n}{\pi}\sqrt{\frac{n!}{m!}}\left[\sqrt{2}(x+\mathrm{i}\,p)\right]^{m-n} L_n^{(m-n)}\left[2(x^2+p^2)\right]\mathrm{e}^{-(x^2+p^2)} \tag{266}$$

for $n > m$ and by its complex conjugate for $n < m$; here, $L_n^{(\alpha)}(x)$ is the generalised Laguerre polynomial given by

$$L_n^{(\alpha)}(x) = \sum_{j=0}^{n} \frac{(-1)^j}{j!}\binom{n+\alpha}{n-j} x^j. \tag{267}$$

Despite the presence of an imaginary part in each of the coherence terms, the Wigner function remains wholly real as the imaginary components of $I_{n,m}$ and $I_{m,n}$ cancel fully.

After the application of the unitary, $\hat{U}_\Lambda$, the output system will again be dictated by the coordinate transformation,

$$W_{\text{out}}(\boldsymbol{q}) = W_{\text{in}}(\Lambda \cdot \boldsymbol{q}) \tag{268}$$

$$= \sum_n |\alpha_n|^2 W_n((\Lambda \cdot \boldsymbol{q})_1) W_G((\Lambda \cdot \boldsymbol{q})_{2\ldots N}) + \sum_{n \neq m} \alpha_n^\dagger \alpha_m I_{n,m}((\Lambda \cdot \boldsymbol{q})_1) W_G((\Lambda \cdot \boldsymbol{q})_{2\ldots N}), \tag{269}$$

for $\boldsymbol{q}$ the quadrature coordinates. We already know the form of the partial trace of this first summation from theorem 10. Integrating over the second summation, however, poses some challenge due to the presence of the additional $(x \pm \mathrm{i}\, p)^n$ component that breaks the previous symmetry.

To tackle these coherence term integrals, we return to the fundamental integral from theorem B.3, following the general process

1. expand the polynomial recursively to reach a sum of integrals of the form given in equation (243);
2. solve each integral individually using theorem B.3;
3. sum these integral solutions to get the result.

The results presented in section 7.2 on the use of quantum state sharing for Fock superposition states are derived following algorithm 1, implemented using Wolfram Mathematica [78], an algebraic computing library which produces an analytic output. Although this process must be performed separately for every superposition combination, the symbolic nature of the

Algorithm 1 Integration of a superposition Wigner function

To trace out the auxiliary modes from a superposition containing eigenstates drawn from the set $\mathcal{N}_\psi$.

```
SET solution = 0
▷ Solve for the eigenstate terms
FOR n ∈ N_ψ DO
    APPLY theorem 10 TO |n⟩⟨n| TO GET result
    ADD |α_n|² × result TO solution
▷ Solve for the coherence terms
FOR n ∈ N_ψ, m ∈ N_ψ WHERE m ≠ n DO
    expand I_{n,m}((Λ · q)_1) W_G((Λ · q)_{2…N}) to a sum of integrals
    SET result_sum = 0
    FOR EACH integral IN I_{n,m}((Λ · q)_1) W_G((Λ · q)_{2…N}) DO
        APPLY theorem B.2 TO integral TO GET result
        ADD result TO result_sum
    ADD α_n† α_m × result_sum TO solution
RETURN solution
```

computations ensures that a full characterisation of the output for each superposition class is achieved. In particular, the analysis covers an arbitrary weighting of the eigenstates; the specific superposition is not specified.

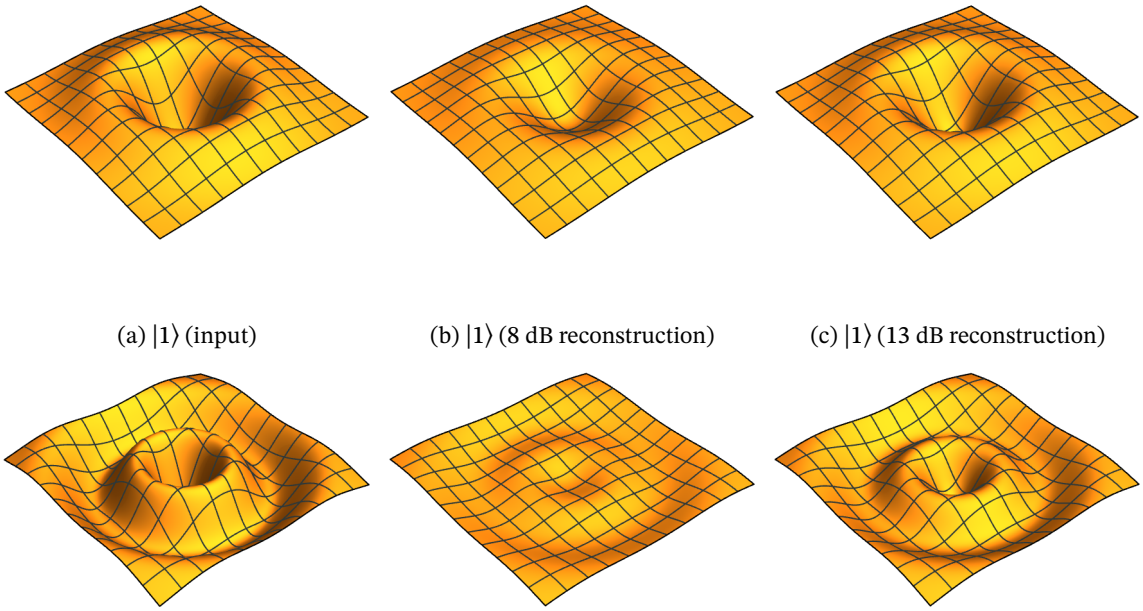
We also use this approach to find the reconstruction fidelity by solving the integral

$$\mathcal{F} = 2\pi \int_{\mathbb{R}^2} dx dp W_{\psi'}(x, p) W_{\text{in}}(x, p), \tag{270}$$

using an approach analogous to algorithm 1.

7

HYBRID FOCK-GAUSSIAN QUANTUM STATE SHARING



(a) $|1\rangle$ (input) (b) $|1\rangle$ (8 dB reconstruction) (c) $|1\rangle$ (13 dB reconstruction)

(d) $|3\rangle$ (input) (e) $|3\rangle$ (8 dB reconstruction) (f) $|3\rangle$ (13 dB reconstruction)

Figure 18: Input and reconstructed-output states for $|1\rangle$ and $|3\rangle$ Fock-eigenstate secrets shared using a TMSV resource state with 8 dB or 13 dB squeezing as labelled. In every case, the output state is a smoothed, noisy copy of the input state due to the added thermal noise from the resource state residue, but the essential features of the original Wigner function are clearly visible. As resource-state entanglement increases, the similarity of the reconstruction to the original input state increases.

Discrete-variable systems, in which the quantum states are represented as eigenstates of a countable variable such as particle number, are widely used within quantum technology and form the basis of many quantum computation stacks [79–82]. However, their usefulness is tempered by the relative difficulty in the distribution of such states. In contrast to continuous-variable communication schemes, which operate fully deterministically, discrete-variable protocols are inherently probabilistic due to their reliance on the outcome of a probabilistic Bell measurement [68, 83, 84]. A certain proportion of attempted transmissions will fail, posing a problem for the communication of quantum information that cannot be readily copied. On the other hand, continuous-variable communica-

tion can provide a 100% success rate at the cost that every transmission will be degraded in some way [70]. Further, continuous-variable entanglement can be readily generated and distributed in a deterministic way simply by interfering squeezed states [85], while the generation of discrete-variable entanglement requires sophisticated quantum equipment. The development of so-called *hybrid protocols* which allow for the use of such continuous-variable entanglement and communication protocols for the transmission of discrete-variable states is therefore of great interest.

In this chapter we will consider the potential for our quantum state sharing protocol to be used as such a hybrid protocol, sharing discrete variable states using Gaussian entanglement. In section 7.1 we will first consider the sharing of particle-number eigenstates. Using the framework established in chapter 6, we will demonstrate that this protocol remains useful for the full range of Fock eigenstates given a suitably-entangled resource state. As the eigenstates are perfectly distinguishable using a particle-number measurement, and thus perfectly clonable, we do not present a security condition for such states here. However, we will discuss briefly the potential for security for this type of state.

We will then, in section 7.2, go on to consider two-level superpositions of number states, most notably the $|0\rangle/|1\rangle$ and $|1\rangle/|2\rangle$ qubit states.

7.1 SHARING FOCK EIGENSTATES

Let us begin our discussion of the use of our protocol for discrete-variable states by considering the photon-number eigenstates.

In section 4.2.4 we introduced pre-amplification or post-attenuation steps, as required, to our quantum state sharing protocol to correct for the amplification it imposes on the secret state and thus preserve the phase-space mean. The Fock states we will consider in this chapter, though, are always zero-mean so it is worth asking at this stage whether this correction remains useful. For much of our discussion of Fock eigenstate QSS, we will consider both the uncorrected generally-amplifying protocol and the corrected nonamplifying version.

7.1.1 *Output state*

Let us begin by considering what the output of this protocol will look like. We know from theorem 10 that Gaussian protocols induce no superposition so the output density matrix will be wholly diagonal and take the form

$$\hat{\rho}_{\text{out}} = \sum_m C_m |m\rangle\langle m|, \tag{271}$$

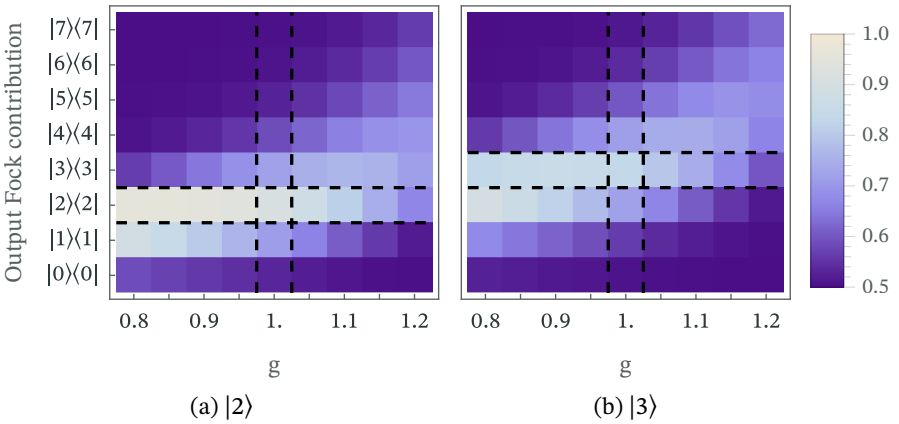
for a set of coefficients C_m summing to 1.

Let us initially consider what we will term the 'raw' output from the protocol, where no correction is made for the amplification. Applying theorem 10 to the protocol outlined in section 4.2.3 yields an output state described by the density matrix elements

$$C_m = \frac{2}{\eta^2} \sum_{b=0}^{\min(m,n)} 4^b \binom{n}{b}\binom{m}{b} \frac{(E_{1|2}(g) - 1 + 1/\eta^2)^{n-b}(E_{1|2}(g) + 1 - 1/\eta^2)^{m-b}}{(E_{1|2}(g) + 1 + 1/\eta^2)^{m+n+1}}, \tag{272}$$

for $\eta = 1/\sqrt{2 - g^2}$ representing the amplification of the output state.

These contributions to the output state are shown for varying reconstruction parameter g in figure 19. Unsurprisingly, the form of the output state shows a clear gradient in photon number as g (and thus the amplification) increases. For values of g below 1, the state becomes increasingly likely to be found with a number of photons below the photon number of the



(a) $|2\rangle$ (b) $|3\rangle$

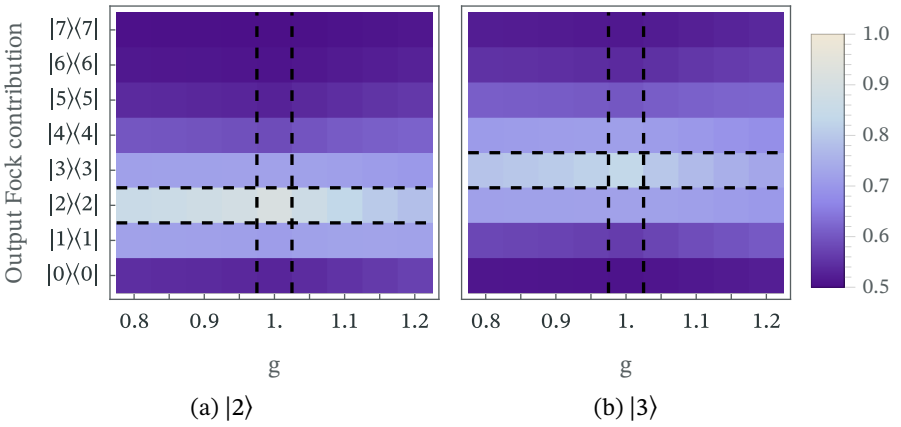
Figure 19: Density matrix contributions to the uncorrected (generally-amplified) reconstruction of (a) $|2\rangle$, (b) $|3\rangle$ input secret state using a TMSV resource state exhibiting 6 dB of squeezing, and so a steering parameter of $E_{1|2}(g) \approx 0.5$. Vertical dashed lines outline the $g = 1$ non-amplifying case; horizontal dashed lines outline the output contribution that matches the input state.

original input state. Conversely, as g increases above 1 the state becomes more sparsely distributed across a distribution of higher-energy eigenstates. We also begin to see the emergence of a curious result, particularly in the $|2\rangle$ case — some of the deamplifying protocols (those for which $g < 1$) produce an output result in which there is a greater likelihood of reproducing the input state than in the nonamplifying ($g = 1$) protocol. We will discuss this result more in section 7.1.2.

Let us now turn our attention to the case in which this amplification has been corrected for through either an additional pre-amplification (for $g < 1$) or post-attenuation (for $g > 1$) step. Applying theorem 10 again to this corrected protocol produces an output state described by the density matrix $C_m|m\rangle\langle m|$ with elements

$$C_m = \begin{cases} 2\frac{(1+\eta^2 E_{1|2}(g)-\eta^2)^{n+m}}{(3+\eta^2 E_{1|2}(g)-\eta^2)^{n+m+1}} \sum_{b=0}^{\min(n,m)} \binom{n}{b}\binom{m}{b}\left[\frac{2}{1+\eta^2 E_{1|2}(g)-\eta^2}\right]^{2b} & g \leq 1 \\ 2\frac{(1+E_{1|2}(g)-1/\eta^2)^{n+m}}{(3+E_{1|2}(g)-1/\eta^2)^{n+m+1}} \sum_{b=0}^{\min(n,m)} \binom{n}{b}\binom{m}{b}\left[\frac{2}{1+E_{1|2}(g)-1/\eta^2}\right]^{2b} & g \geq 1, \end{cases} \tag{273}$$

for an input state with known photon number n. These contributions are shown for varying g and the same example resource state in figure 20.



(a) $|2\rangle$ (b) $|3\rangle$

Figure 20: Density matrix contributions to the ampification-corrected reconstruction of (a) $|2\rangle$, (b) $|3\rangle$ input secret state using a TMSV resource state exhibiting 6 dB of squeezing, and so a steering parameter of $E_{1|2}(g) \approx 0.5$. Vertical dashed lines outline the $g = 1$ non-amplifying case; horizontal dashed lines outline the output contribution that matches the input state.

Comparing to figure 19, the effect of this amplification correction becomes apparent. Much of the g-dependence in the spread of output contributions has vanished, with the output state much more closely bunched around the input state for all values of g. There remains potential for the output

state to be a different Fock state to the input state, but this probability now presents in a broadly uniform way and is limited to the region immediately around the input state.

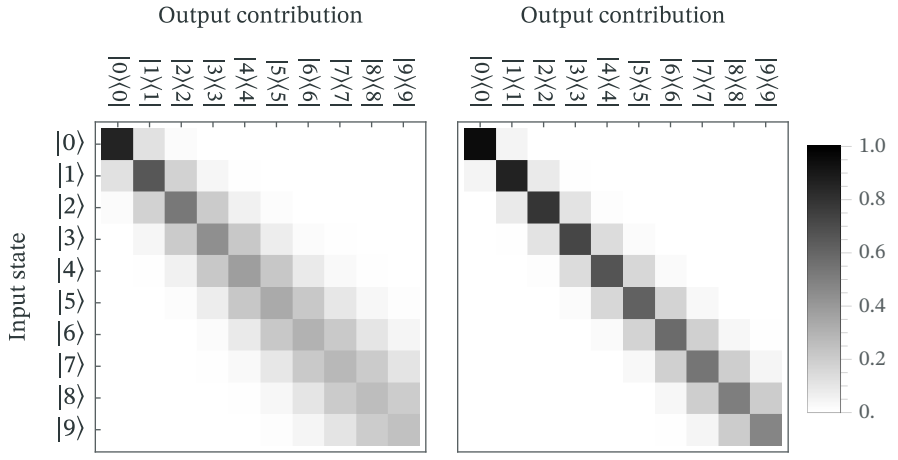
On the other hand, however, the improved likelihood of finding the output state in the correct eigenstate seen for $g < 1$ in figure 19 has been lost, so for some g values correcting for the amplification may produce a worse result. We will consider this trade-off more when we discuss the benefits of correcting for amplification in more detail in section 7.1.3.

Finally, let us consider for a moment the special case in which $g = 1$ and no phase-space amplification is imparted by the protocol. In addition to gaining the benefits of a non-amplifying protocol without requiring the use of noise-inducing correction steps, $g = 1$ represents the space in which the best-quality entanglement can be found. When $g = 1$, the density matrix elements reduce to

$$C_m = \frac{2}{(E_{1|2}(g)+2)^{m+n+1}} \sum_{b=0}^{\min(m,n)} \binom{n}{b}\binom{m}{b} 4^b [E_{1|2}(g)]^{m+n-2b} \quad (274)$$

Notably, the summation runs only to the smaller of m and n — only when $m = n$ does the $2b = m + n$ case, in which the term does not directly scale with $E_{1|2}(g)$, exist. In the perfect entanglement limit, when $E_{1|2}(g) \to 0$ only the C_n term will survive therefore and the original secret state will always be reproduced perfectly.

The makeup of this output state is shown for a non-amplifying protocol for a number of example input states in figure 21. Clearly, increasing the



(a) 8 dB TMSV resource (b) 13 dB TMSV resource

Figure 21: Output state contributions for the first 10 Fock input states at $g = 1$, using the labelled resource states. Each row in the figure represents a the noted input state, with each box in the row representing the contributions to the output density matrix, as labelled above. Tighter horizontal distributions represent better reconstructions.

energy level of the input state increases the photon-number variance in the output. Low particle-number states can be shared and reconstructed relatively easily with limited introduction of neighbouring states; as the particle number increases, though, the spread in potential photon number increases and the output state becomes less pure. Although these higher-photon-number states decohere quickly for the worse-entangled resource state in figure 21a by increasing the resource squeezing even the high-energy states return to the reasonably tight distribution seen in figure 21b.

This output state can be described by a classical probability distribution with variance strongly dependent both on the resource steering parameter and original input state as[1]

$$\sigma^2 = E_{1|2}(g) + \frac{1}{2}nE_{1|2}(g) + \frac{1}{4}E_{1|2}(g)^2, \tag{275}$$

indicating that although higher photon number states are more susceptible to the noise added by the protocol this can indeed be offset by increasing entanglement. In the perfect entanglement limit, the input state is reproduced perfectly with no variance in photon number. One can see intuitively why this might be the case by returning to the Wigner functions shown in figure 18: as the photon number increases, the state contains more complex quantum features and is thus more susceptible to the impact of added Gaussian noise. This represents a continuation of the trend we have returned to throughout this thesis, that the more 'quantum' an input state is the harder it is to share.

Although this distribution appears to be centred on the input state, in fact even in the non-amplifying case the output state has a mean photon number slightly above the input state given by

$$\langle \hat{n} \rangle = n_{\text{in}} + \frac{1}{2}E_{1|2}(g), \tag{276}$$

reflecting the slight photon-increasing bias inherent to thermal noise.[2] This distribution in output photon number is shown more clearly in figure 22 for selected input states.

[1] This variance and the following mean measure were found through a numerical fit that perfectly predicted the first 150 Fock state outputs. Given the natural form of the result (consisting only of simple rational coefficients), we conjecture that these results are exact and hold for all $|n\rangle$ input.

[2] This bias is easily explained by considering that the random thermal fluctuations act on phase-space amplitude, which is not bounded by 0. A large enough shift in the direction initially reducing photon number will eventually surpass the origin and begin to increase photon number again, while a shift in the opposite direction will always increase photon number. When these random fluctuations are described by a Gaussian distribution the tails will be photon-increasing in both directions, causing a small but not-insignificant bias towards photon-number increase.

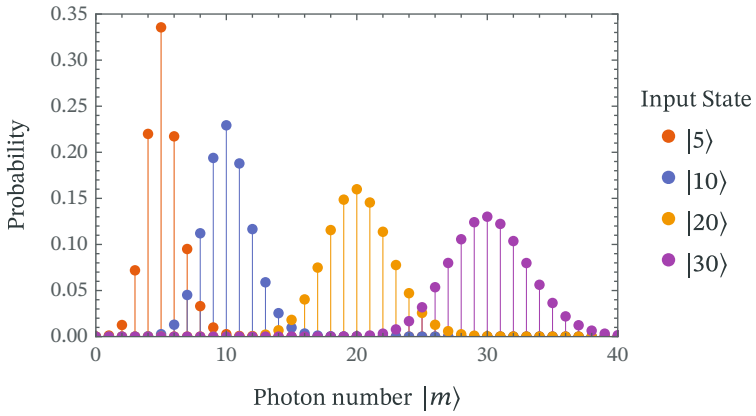


Figure 22: Photon probability distribution in the reconstructed output state for given Fock input states when shared at $g = 1$ using a two-mode squeezed vacuum resource state with 8 dB squeezing.

The output probability closely resembles a Gaussian distribution, tapering exponentially as the distance between the input and output photon numbers increases, but exhibits a degree of skewness due to the asymmetric bounds on photon number. A standard Gaussian is therefore unable to properly describe the output, although there are a number of asymmetric Gaussian distributions that allow for skewness, including the skewed-normal [86, 87] and folded normal [88] distributions, that may be able to be applied to model these results.

7.1.2 *Reconstruction quality*

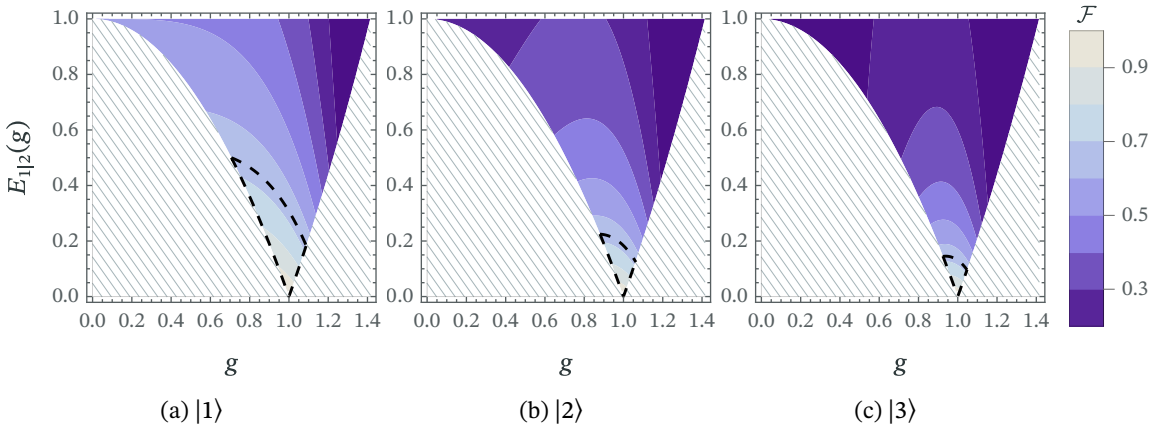
Let us now focus more closely on the quality with which this protocol reconstructs Fock input states. As the spectrum of possible output states is orthogonal, the fidelity of the output state is precisely the probability amplitude describing the likelihood that it would be found in the $|n\rangle\langle n|$ state, given by the relevant density matrix element

$$\mathcal{F} = C_n. \tag{277}$$

For the 'raw' QSS protocol, prior to the application of any amplification corrections, a Fock-state secret $|n\rangle$ is reconstructed with fidelity

$$\mathcal{F} = \frac{2}{\eta^2(E_{1|2}(g) + 1 + \frac{1}{\eta^2})^{2n+1}} \sum_{b=0}^{n} 4^b \binom{n}{b}^2 (E_{1|2}(g)^2 - (1/\eta^2 - 1)^2)^{n-b}, \tag{278}$$

for $\eta = 1/\sqrt{2 - g^2}$ as usual.



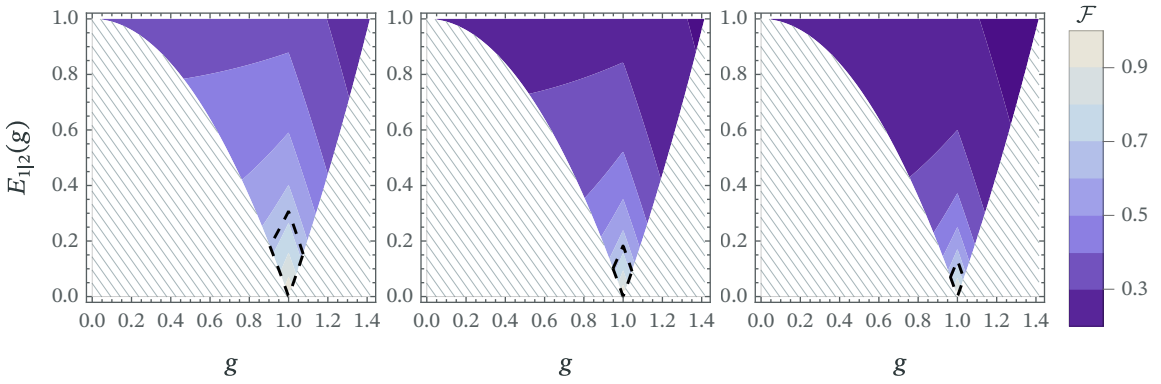
(a) $|1\rangle$ (b) $|2\rangle$ (c) $|3\rangle$

Figure 23: Achievable reconstruction fidelity using an uncorrected QSS protocol parameterised by g with a Gaussian resource state with steering parameter $E_{1|2}(g)$ to share the labelled Fock states. Only shown is the region $E_{1|2}(g) > |1 - g^2|$ for which steering is able to exist. Region enclosed by the dashed line denotes $\mathcal{F} > 2/3$ as a benchmark only as this does not guarantee security for the set of Fock eigenstate inputs.

This fidelity is shown for the first three excited Fock states for varying setup parameters in figure 23. We can see again here the curious behaviour for $g < 1$ we noted in section 7.1.1 emerge from the fidelities shown in this figure. Ordinarily, all other circumstances being equal, we would expect that the best result would be obtained at $g = 1$, when the output state is least-distorted by amplification. As can be seen most clearly in the $|2\rangle$ reconstruction profile in figure 23b, though, the optimal value of g for any set steering parameter lies some point below $g = 1$. As the photon number increases, the optimal reconstruction parameter increases towards $g = 1$ with further de-amplified reconstructions again producing a worse result.

There are two effects contributing to this phenomenon, the first of which we have previously seen affect Gaussian states in chapter 5. While a de-amplification in the protocol represents a degradation of the information describing the state, it also reduces the thermal noise added by the remaining resource state contributions. As deamplifiction is a multiplicative effect — impacting larger photon-number states more than lower energy ones — the negative impact is more fully offset by the reduced thermal contributions for lower Fock states. As the energy level of the input state rises, though, the deamplification begins to have a much more significant effect that outweighs any potential improvement from the reduced thermal contribution, so the optimal reconstruction parameter moves again towards $g = 1$.

There is, additionally, a further effect that only emerges when we consider Fock states: as we found in section 7.1.1, thermal noise is not zero-mean in photon-number space. Even at $g = 1$ this protocol has a mild



(a) $|1\rangle$ (b) $|2\rangle$ (c) $|3\rangle$

Figure 24: Achievable reconstruction fidelity using an amplification-corrected QSS protocol parameterised by g and a Gaussian resource state with steering parameter $E_{1|2}(g)$ to share the labelled Fock states. Only shown is the region $E_{1|2}(g) > |1 - g^2|$ for which steering is able to exist. Dashed region denotes $\mathcal{F} > 2/3$ as a benchmark only, as this does not guarantee security for the set of Fock eigenstates.

photon-number increasing effect and so, in contrast to the phase-space picture, the unity-gain point for this protocol will lie mildly below $g = 1$. As any deamplification acts proportionately to the photon number n, while the photon-number increase from thermal noise is constant, this is not an amplification that can be universally corrected for. When the input states are drawn from a known distribution of Fock states, a specific optimal g_{unity} can be derived that minimises the average amplification across that distribution (although some input states will still be amplified while others will be de-amplified). In the absence of such implementation-specific details, however, we will continue to use $g = 1$ as our unity-gain point to avoid an outsize impact on very high photon number states.

Let us consider now the reconstruction quality when we do make a correction such that there is no phase-space amplification at any value of g. The same protocol with this correction applied produces a reconstruction fidelity of

$$\mathcal{F} = \begin{cases} 2\frac{(1+(E_{1|2}(g)-1)\eta^2)^{2n}}{(3+(E_{1|2}(g)-1)\eta^2)^{2n+1}} \sum_{b=0}^{n} \binom{n}{b}^2 \left[\frac{2}{1+(E_{1|2}(g)-1)\eta^2}\right]^{2b} & g \leq 1 \\ 2\frac{(1+E_{1|2}(g)-1/\eta^2)^{2n}}{(3+E_{1|2}(g)-1/\eta^2)^{2n+1}} \sum_{b=0}^{n} \binom{n}{b}^2 \left[\frac{2}{1+E_{1|2}(g)-1/\eta^2}\right]^{2b} & g \geq 1 \end{cases} \tag{279}$$

which is shown in figure 24. As anticipated, applying an amplification correction has removed the fidelity bonus observed in the $g < 1$ region. Indeed, in every case the peak achievable reconstruction fidelity is now found at $g = 1$, where no amplification correction is required. However, this mild loss in fidelity for values of $g < 1$ close to 1 is accompanied by

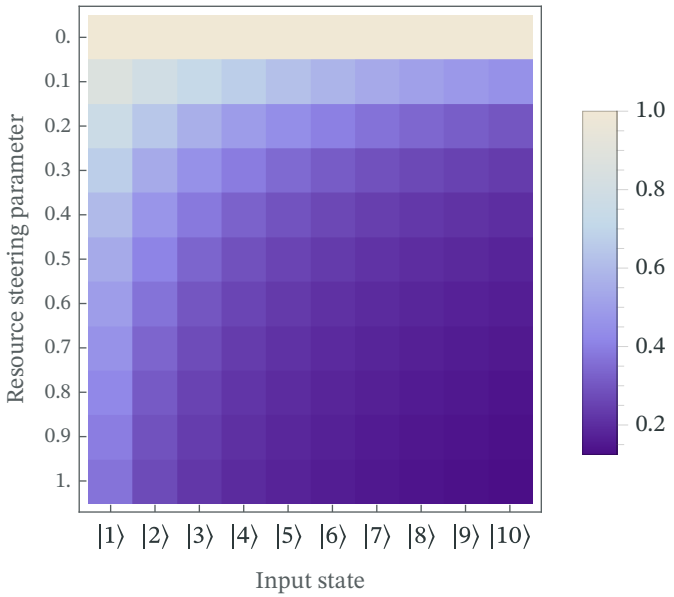


Figure 25: The achievable reconstruction fidelity using a QSS protocol implemented at $g = 1$ for Fock input states as denoted and a Gaussian resource state with steering parameter $E_{1|2}(g)$.

a significant increase in the achievable reconstruction fidelity in highly asymmetric $g \neq 1$ regions.

As the pre-amplification correction for $g < 1$ setups occurs prior to the application of the QSS protocol, this setup continues to benefit from the resource-contribution deamplification that we attributed much of the increased fidelity to. However, this amplification step is itself necessarily noisy, so imparts an additional thermal distortion to the secret state. The quality of the amplification-corrected output is a tradeoff between this additional noise added by the amplification and the improvement by undoing the deamplification on the output state. For low photon numbers and reconstruction parameters already close to $g = 1$ the negative impacts of this correction may well outweigh any gains from removing the de-amplification.

Finally, let us comment briefly on the special $g = 1$ case in which the protocol is naturally non-amplifying in phase-space and no correction is required. This is also the point in which the greatest levels of quantum steering is possible, as the level of steering present in a resource state is bounded by $|g^2 - 1|$. In this case, the reconstruction fidelity for a Fock $|n\rangle$ input is given by

$$\mathcal{F}_{g=1} = \frac{2}{(E_{1|2}(g)+2)^{2n+1}} \sum_{b=0}^{n} \binom{n}{b}^2 4^b E_{1|2}(g)^{2(n-b)}. \tag{280}$$

This obtainable reconstruction fidelity for varying degrees of resource entanglement is shown for the first 10 Fock states in figure 25. Although a deep reduction in reconstruction fidelity is observed as n increases, with extremely good entanglement necessary to achieve high fidelity for high-photon-number input states, in the limit of ideal entanglement every Fock state is reproduced perfectly. Although in practical terms the resource requirements may be prohibitively high in some cases, every Fock state can therefore in theory be shared with arbitrarily good fidelity — the only limit to protocol effectiveness is the availability of entanglement resources.

7.1.3 *To amplify or not to amplify*

Having now considered both uncorrected (amplifying) and corrected (non-amplifying) forms of the protocol, let us briefly discuss the question of which should be used. As with many questions of optimising for quantum information, there is no single answer here; the preferable protocol depends greatly on which set of Fock states are used and at which reconstruction parameter.

Consider for example figure 26, which shows the reconstruction and steering parameters for which an improvement in fidelity can be found by correcting for the amplification. Although most states show a fidelity

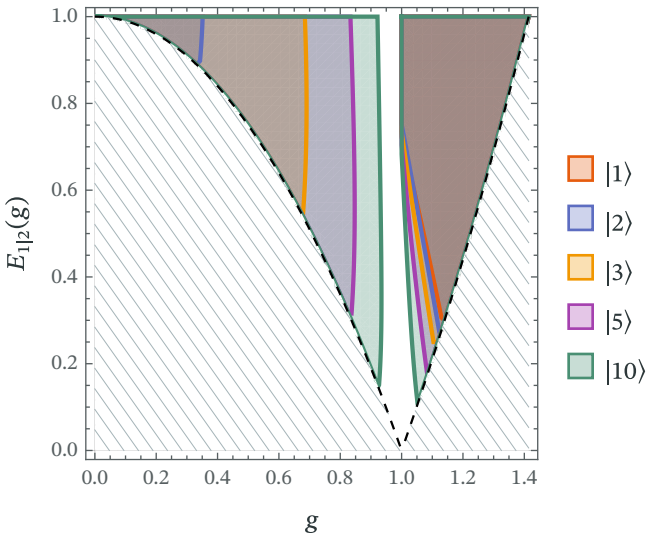
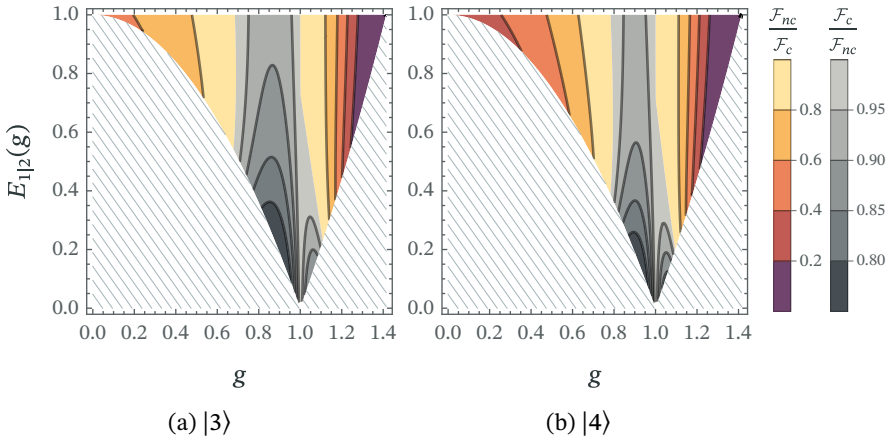


Figure 26: The regions in which reconstruction fidelity would be improved by introducing an amplification correction for the labelled Fock input states. Filled regions indicate that improvement is possible from the correction.

improvement from the amplifiction-correction for $g > 1$, the low photon-number states — those least affected by attenuation — are improved by

leaving the de-amplification uncorrected for $g < 1$. As the photon number increases, though, the region in which not correcting for the amplification is preferable shrinks.

We should consider here also the magnitude of this improvement, though. This is shown for the $|3\rangle$ and $|4\rangle$ states in figure 27 . Although there may be



(a) $|3\rangle$ (b) $|4\rangle$

Figure 27: Ratio of reconstruction fidelity using corrected $\mathcal{F}_c$ and uncorrected $\mathcal{F}_{nc}$ QSS protocols on (a) $|3\rangle$ and (b) $|4\rangle$ input secret states when using a resource state with steering parameter $E_{1|2}(g)$ and reconstruction parameter g as indicated. Coloured region indicates improvement from correcting for amplification; grey region indicates a better fidelity can be achieved by leaving the output in its amplified form.

a reduction in fidelity from correcting for the amplification in the region around $g = 1$ (shown in grey), this loss is small compared to the loss from leaving the state (de)amplified elsewhere (shown in colour). The maximal loss from correcting for the amplification is a $\sim 20\%$ reduction in fidelity, while the potential loss from not correcting for it approaches the complete loss of information about the state. Were the amplified output found to be optimal for some of the set of input states and the non-amplified output optimal for others, it is likely then that correcting for the amplification would provide the best overall fidelity.

Finally, let us note that the amplification correction opens up the use of this protocol for future hybrid CV-DV states. The displaced Fock state, in which a phase-space displacement operation has been applied to a photon number state, has the ability to carry information both in its original discrete Fock state form and through the continuous displacement of its mean [89]. Any remaining phase-space amplification would distort this mean vector as it would for the coherent state, disproportionately reducing the reconstruction fidelity for highly-displaced states. For the protocol to be useful to this class of state, then, the amplification correction is required regardless of its impact on the underlying Fock state.

7.1.4 *Security*

Finally, let us consider for a moment the matter of security. Although we have frequently highlighted the $\mathcal{F} > 2/3$ region to enable easier comparison between figures, we do not present a security analysis for the set of Fock eigenstates. This is because our usual no-cloning criteria cannot be applied here. As the Fock states consist of an orthogonal set, there is no limit on the quality with which they can be cloned — a Fock state can be readily identified through a particle-counting measurement, after which any number of perfect clones can be created directly. The lack of an upper limit on cloning fidelity for these states means we are unable to be certain of security based on the achieved reconstruction fidelity alone.

However, let us briefly note that although the input states are drawn from an orthogonal set, the intermediate shares they produce are not orthogonal — and form part of a wider entangled system — and so cannot themselves be readily cloned. The protocol therefore retains a degree of security from measure-and-resend attacks.

A security analysis for these states could be performed by considering the information available to an adversary directly. In the ideal case, this would correspond to simply comparing the amount of information contained in a single share to the information obtainable from the reconstruction. In a realistic implementation, however, a full analysis would have to consider all sources of loss as these could be due to the presence of an eavesdropper siphoning part of the signal.

The lack of a defined security condition on the Fock eigenstates may limit the use of this protocol for state transmission. However, QSS still has uses outside of secure quantum cryptography; for example, in recovery from transmission loss. The high-fidelity reconstruction of Fock states therefore remains of interest regardless of the security of the protocol. Let us now turn our attention to superpositions of such states, where security can again be guaranteed.

7.2 SHARING FOCK SUPERPOSITION STATES

In this section, we continue our analysis of the Fock states by considering superposition states of the general form

$$|\psi\rangle = \sum_n \alpha_n |n\rangle \tag{281}$$

for $\sum|\alpha|^2 = 1$.

In the previous section, in addition to looking at the non-amplifying QSS protocol we also considered the output states in their 'raw' uncorrected form. In some limited cases, such as the $|0\rangle$ / $|1\rangle$ qubit state, an improvement in fidelity may be found by allowing for such a de-amplifying QSS protocol. However, as the de-amplification acts on each eigenstate differently — and directly transforms some $|1\rangle$ states into $|0\rangle$ states — care would need to be taken when allowing for a generally-amplifying protocol to ensure the superposition was not unduly affected. To avoid this problem, we focus here exclusively on the non-amplifying case.

In any event, much of our discussion in this section will focus on the $g = 1$ case where the two cases coincide.

7.2.1 *Overview of process and security condition*

In our discussion of superposition QSS, we will initially consider the single-shot fidelity obtainable for a fully specified state — that is to say, for specific values of α_i — as

$$\mathcal{F}(\boldsymbol{\alpha}) = 2\pi \int_{\mathbb{R}^2} dx\, dp\, W_{\text{in}}(x, p, \boldsymbol{\alpha}) W_{\text{out}}(x, p, \boldsymbol{\alpha}), \tag{282}$$

which can be solved using algorithm 1 outlined in section 6.3. As we will see, the fidelity with which a given superposition state is reconstructed depends on the balance between its contributing eigenstates. As a simple example, we already know from the previous section that the extremal $|0\rangle$ and $|1\rangle$ members of the general superposition class $\alpha_0|0\rangle + \alpha_1|1\rangle$ are reconstructed with vastly different fidelities.

Much of our interest, though, will be in the *average* reconstruction fidelity achievable across an entire class of superposition state, given by integrating across the superposition coefficients as

$$\mathcal{F}_{\text{avg}} = \frac{1}{\mathcal{N}} \int d\boldsymbol{\alpha} \mathcal{F}(\boldsymbol{\alpha}), \tag{283}$$

for some $\mathcal{N} = \int d\boldsymbol{\alpha}$ which normalises the average.

We will here parameterise an N-mode superposition using the unit-radius hyperspherical coordinate system as

$$|\psi\rangle = \cos\theta_1|\psi_1\rangle + \sin\theta_1 \cos\theta_2|\psi_2\rangle + \cdots + \sin\theta_1 \sin\theta_2 \ldots \sin\theta_{N-1}|\psi_N\rangle, \tag{284}$$

to ensure it remains normalised, in which case the average fidelity across the superposition class is given by the integral[3]

$$\mathcal{F}_{\text{avg}} = \frac{\Gamma(\frac{N}{2})}{\pi^{N/2}} \int_0^{\pi} d\theta_1 d\theta_2 \ldots d\theta_{N-1} \Big(\prod_{i=1}^{N-2} \sin^i \theta_{N-1-i} \Big) \mathcal{F}(\boldsymbol{\theta}), \tag{285}$$

for $\Gamma(\cdot)$ the Gamma function.[4] For a two-mode superposition this average can be found simply as

$$\mathcal{F}_{\text{avg}} = \frac{1}{\pi} \int_0^{\pi} d\theta \mathcal{F}(\theta), \tag{287}$$

while for three-mode superpositions it is given by

$$\mathcal{F}_{\text{avg}} = \frac{1}{2\pi} \int_0^{\pi} \int_0^{\pi} d\theta_1 d\theta_2 \sin\theta_1 \mathcal{F}(\theta_1, \theta_2). \tag{288}$$

As the specific superposition within each class is not distinguishable by a single measurement (as the Fock eigenstates are, for example), communication using them is again protected by the no-cloning theorem. The ability to optimally clone arbitrary Fock superpositions has been studied by Bužec & Hillery [58] who prove that the best possible fidelity from two clones of an N-level superposition state is

$$\mathcal{F}_{\text{opt. cloning}} = \frac{N+2}{2N+2}. \tag{289}$$

As we discussed in section 4.3, exceeding this fidelity guarantees that the state reconstruction is secure. For qubit states, this equates to security threshold fidelity of $\mathcal{F} = 2/3$, while for qutrits the security threshold is reduced to $\mathcal{F} = 5/8$.

[3] Ordinarily this integral over the unit hypersphere would consist of $N-2$ integrals over the range $\theta_{1\ldots N-2}: 0 \mapsto \pi$ and one integral over the range $\theta_{N-1}: 0 \mapsto 2\pi$. For a quantum state, though, the lower hemisphere represents only a global phase shift from the upper hemisphere and so can be neglected.

[4] For integer parameters n the Gamma function is given by $\Gamma(n) = (n-1)!$. For half-integer parameters $n + 1/2$, it is given by

$$\Gamma(n + \frac{1}{2}) = \sqrt{\pi} \frac{(2n)!}{4^n n!}. \tag{286}$$

7.2.2 *Qubit states*

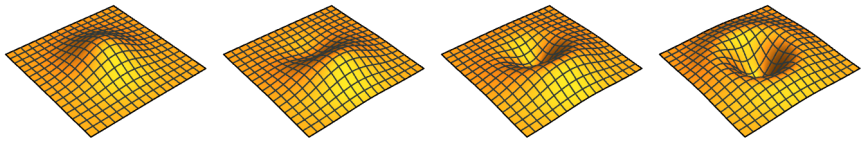
Let us start our discussion of Fock superposition QSS with the lowest-energy qubit state: the superposition of the ground and first excited states given by

$$|\psi\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle. \tag{290}$$

In the Wigner formalism this qubit state is given by

$$W(x,p) = \frac{1}{\pi}\left[\cos^2\theta + \sin^2\theta\,(2x^2 + 2p^2 - 1) + 2\sqrt{2}\sin\theta\cos\theta x\right]\mathrm{e}^{-x^2-p^2}, \tag{291}$$

taking a varying form as the superposition balance changes between the $|0\rangle$ and $|1\rangle$ eigenstates, as shown in figure 28.



(a) $1|0\rangle + 0|1\rangle$ (b) $\frac{\sqrt{2}}{\sqrt{3}}|0\rangle + \frac{1}{\sqrt{3}}|1\rangle$ (c) $\frac{1}{\sqrt{3}}|0\rangle + \frac{\sqrt{2}}{\sqrt{3}}|1\rangle$ (d) $0|0\rangle + 1|1\rangle$

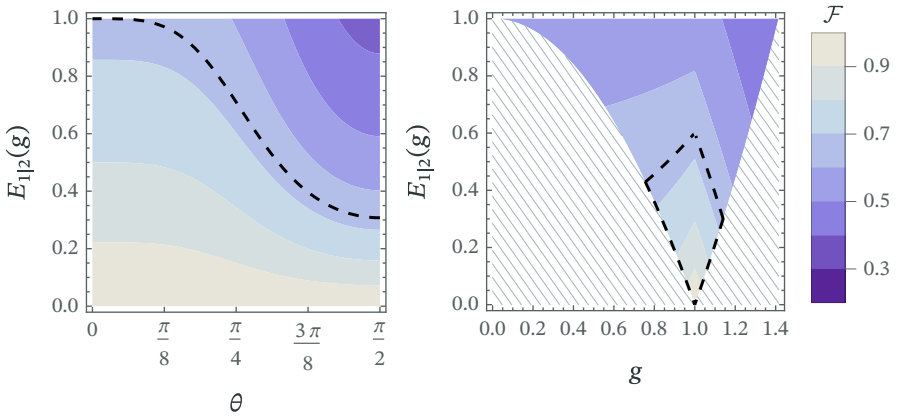
Figure 28: Wigner function representation of indicated qubit superpositions between $|0\rangle$ to $|1\rangle$.

The fidelity after applying the quantum state sharing protocol to this state can be obtained by applying algorithm 1 to the fidelity integral in equation (282) — splitting the resultant product into a sum of individually solvable integrals before summing their solutions. For the standard $g = 1$ case, we find that such a superposition state can be reconstructed with a θ-dependent fidelity given by

$$\mathcal{F}_\theta = \frac{[4\cos(2\theta) - \cos(4\theta) + 2E_{1|2}(g) + 5]E_{1|2}(g) + 8}{(E_{1|2}(g) + 2)^3}, \tag{292}$$

which is shown in figure 29a. Although the full spectrum of superposition states spans from $\theta : 0 \mapsto \pi$, as the achievable fidelity is symmetric about $\pi/2$ we only present the range in which both coefficients are positive.

As might be expected, this superposition is relatively easy to share. Indeed, those superpositions that lie close to the $|0\rangle$ vacuum eigenstate can be shared with $\mathcal{F} = 2/3$ with very small degrees of steering. As these nearly-$|0\rangle$ states are approximately Gaussian in nature, with the vacuum state belonging to the class of coherent states, this result is simply an application



(a) (b)

Figure 29: (a) Single-shot reconstruction fidelity achievable at $g = 1$ for a $\cos\theta|0\rangle + \sin\theta|1\rangle$ qubit secret state for varying steering parameter. (b) Average reconstruction fidelity across a flat θ distribution for the same input state for varying g. Dashed line in each case denotes the $\mathcal{F} > 2/3$ region for which security can be guaranteed.

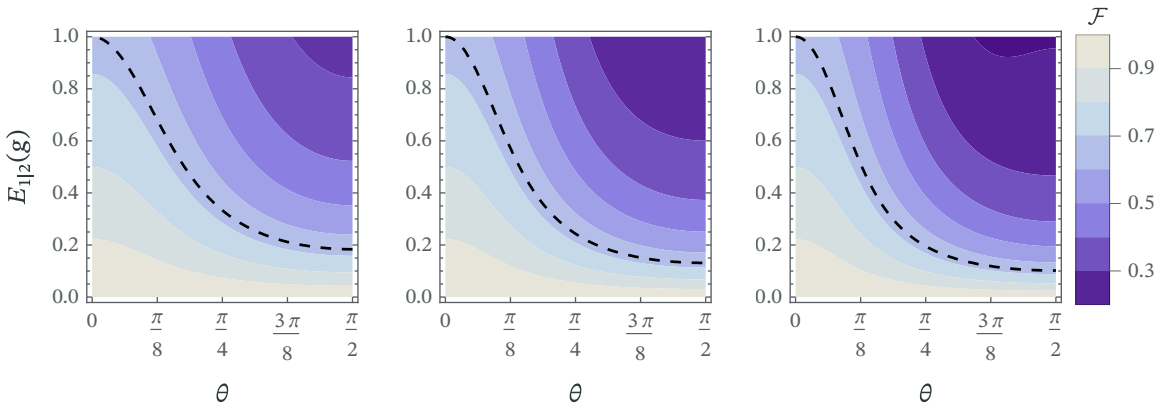
of the Gaussian protocol we have studied extensively in section 5.2. As the superposition ratio moves further towards the $|1\rangle$ eigenstate, though, the reconstruction fidelity achievable for a given level of steering drops monotonically with a steering parameter of $E_{1|2}(g) \approx 0.31$ required to guarantee security at $\theta = \pi/2$. Notably, though, the reconstruction fidelity never drops below that achievable for the $|1\rangle$ eigenstate; any member of this superposition class can be reconstructed with no-worse fidelity than the first excited state.

In analysing the effectiveness of this protocol for the full class of such superposition states, we should consider the expected reconstruction fidelity from an unknown state randomly selected from the class — otherwise known as the average fidelity obtainable across the class from $\theta = 0$ to $\theta = \pi$. In the $g = 1$ special case this averaged fidelity is given by

$$\mathcal{F}_{\text{avg}} = \frac{8 + 5E_{1|2}(g) + 2E_{1|2}(g)^2}{(E_{1|2}(g) + 2)^3}. \tag{293}$$

The average expected fidelity is shown above for varying reconstruction parameter g and steering parameter $E_{1|2}(g)$ in figure 29b. As we found previously for Gaussian and Fock eigenstate secrets, the best average reconstruction fidelity is found around $g = 1$, with decreasing fidelity — and thus greater entanglement requirements for guaranteed security — as we move away from this range.

Although the steering required to achieve security is increased compared to that for wholly-Gaussian quantum state sharing, there remains a sizeable



(a) $\cos\theta|0\rangle + \sin\theta|2\rangle$ (b) $\cos\theta|0\rangle + \sin\theta|3\rangle$ (c) $\cos\theta|0\rangle + \sin\theta|4\rangle$

Figure 30: Single-shot reconstruction fidelity achievable at $g = 1$ for the two-level superpositions of the vacuum and one excited Fock state as labelled. Dashed line in each case denotes the $\mathcal{F} > 2/3$ region for which security can be guaranteed.

region within which secure QSS can be achieved, denoted by the dashed black line in figure 29b . For the standard $g = 1$ special case, for example, the resource state need only exhibit steering of around $E_{1|2}(g) \approx 0.60$, corresponding to marginally more than 5 dB squeezing for a TMSV resource state — well within the routinely achievable range. Although secure QSS for these qubit states is limited to the broadly symmetric region ($0.76 \lesssim g \lesssim 1.14$), this covers a large number of common resource states; greatly asymmetric states are rare in practice as asymmetry severely limits the potential entanglement.

We have found here that secure quantum state sharing of qubit states can be achieved using presently-achievable Gaussian entanglement with no change to the protocol. Let us now consider whether this holds for the more general two-level Fock superposition.

7.2.3 *Other two-level superpositions*

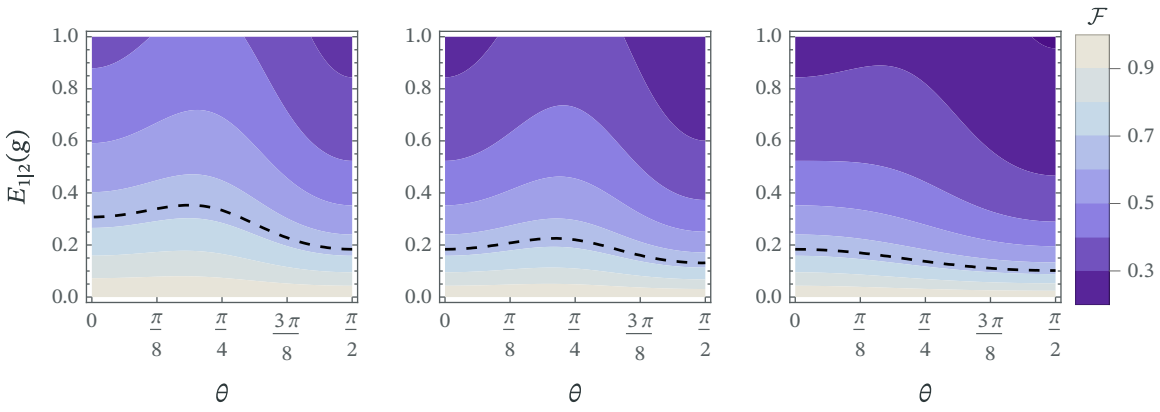
We will here consider the wider class of Fock superposition states of the form

$$|\psi\rangle = \cos\theta|n\rangle + \sin\theta|m\rangle. \tag{294}$$

As discussed in section 6.3, we have not solved this class of state for a general n and m, instead relying on an algorithmic (but fully analytic) approach for specific (n, m) pairings. We will therefore consider the protocol here through a series of example states.

Every example of a two-level superposition that contains the vacuum state acts similarly to the qubit state from the previous subsection. Very high fidelities are achievable while the superposition state remains similar to the $|0\rangle$ eigenstate, while increasingly imperfect copies are reconstructed as the input state moves towards the other eigenstate. A number of examples are shown in figure 30. As might be predicted from the increasing difficulty with which larger Fock eigenstates are reconstructed, as the second contributing eigenstate increases in photon-number the average reconstruction fidelity across the superposition class reduces. Even for relatively high photon numbers, though, the entanglement required to guarantee security remains broadly achievable. The case of the $\cos\theta|0\rangle + \sin\theta|3\rangle$ superposition, for example, is securely sharable using a TMSV resource with ~ 8.8 dB squeezing, which approaches the upper limit of squeezing levels that are readily achievable but remains within the feasible range.

When we consider superpositions of two excited states, though, a more interesting result begins to emerge. Consider, for example, the single-shot fidelity achievable for the $(|1\rangle, |2\rangle)$ and $(|2\rangle, |3\rangle)$ superpositions shown in figures 31a and 31b. Although the achievable reconstruction fidelity at each extreme of the superposition remains low compared to the vacuum state, a peak in fidelity emerges between the eigenstates. It turns out that some Fock superposition states can be shared and reconstructed with a greater fidelity than either of their contributing eigenstates.



(a) $\cos\theta|1\rangle + \sin\theta|2\rangle$ (b) $\cos\theta|2\rangle + \sin\theta|3\rangle$ (c) $\cos\theta|2\rangle + \sin\theta|4\rangle$

Figure 31: Single-shot reconstruction fidelity achievable at $g = 1$ for two-level superpositions of excited Fock states as labelled. Dashed line in each case denotes the $\mathcal{F} > 2/3$ region for which security can be guaranteed.

This initially counter-intuitive result makes sense when one considers the states' similarity to the Gaussian. Recall that the coherent state with complex amplitude α can be defined as the superposition of Fock states

$$|\alpha\rangle = \mathrm{e}^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle. \tag{295}$$

Although the non-vacuum coherent state is a superposition of *all* Fock states, it is likely that a well-selected superposition of two Fock states may be 'more Gaussian' than either eigenstate alone. It should be entirely expected, then, for such superpositions to be more resilient to Gaussian noise than their eigenstate relatives. This is not uniformly the case for all Fock superpositions, though, as shown by the $(|2\rangle, |4\rangle)$ superposition in figure 31c for which the improvement in fidelity is only found at low levels of resource entanglement. Some combinations of eigenstate, such as the $(|1\rangle, |4\rangle)$ superposition, show no improvement at all indicating that not all superpositions have the potential to increase Gaussianess.

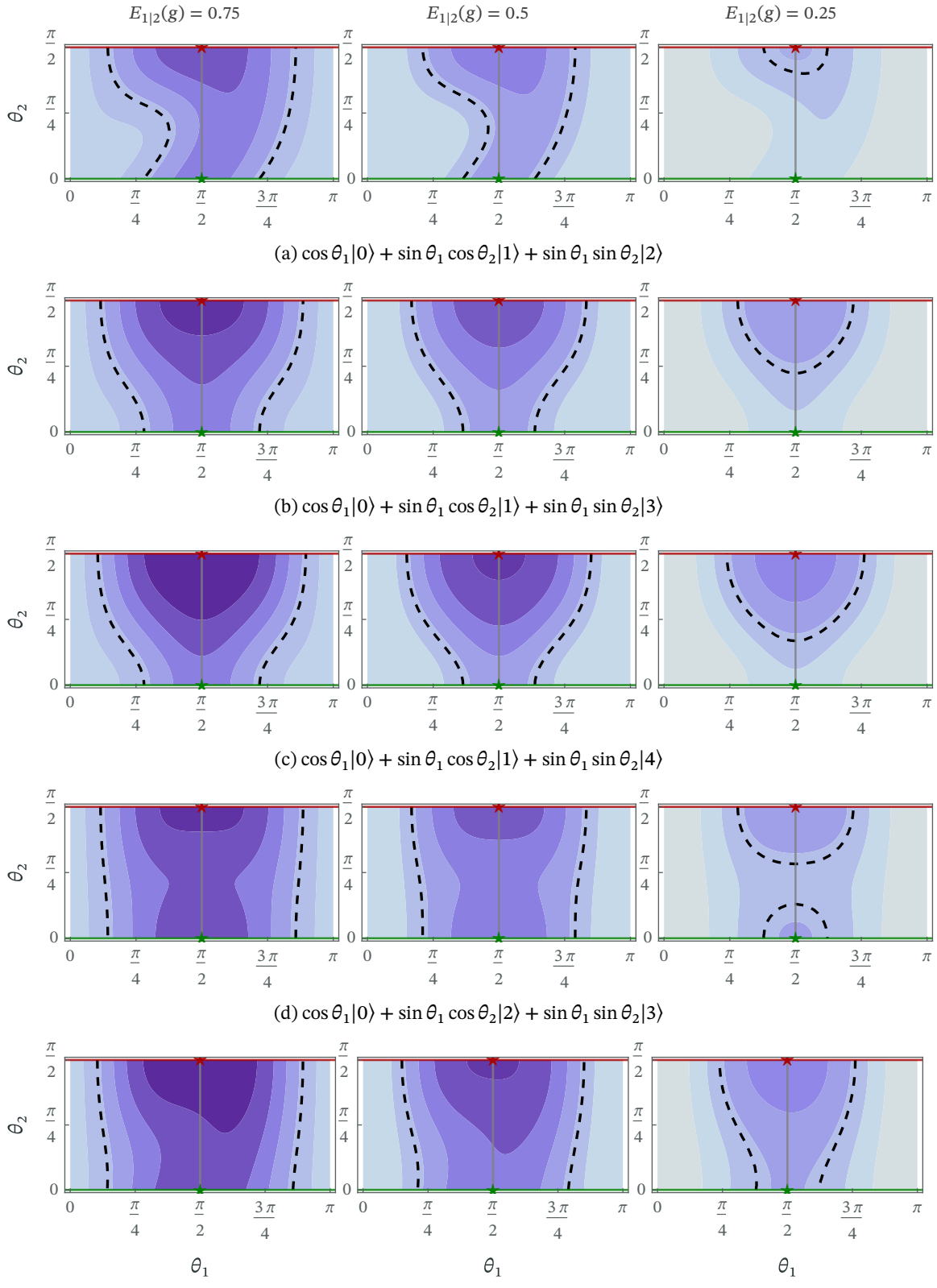
Finally, let us comment on the average reconstruction fidelity for such states and thus their ability to be shared securely. The lowest-energy non-vacuum superposition — the $\alpha_1|1\rangle + \alpha_2|2\rangle$ state — can comfortably be shared securely using presently achievable entanglement, with ~ 8.6 dB squeezing required in a TMSV state. This is an important result as such states are a popular second choice for qubit-style information carriers alongside the $|0\rangle$ and $|1\rangle$ superposition.

Securely sharing higher photon number superpositions may be presently out of reach, however. For example, the $(|1\rangle, |3\rangle)$, $(|2\rangle, |3\rangle)$ and $(|1\rangle, |4\rangle)$ superposition states each require between $10db$ and $11db$ of resource squeezing. Although achievable with existing technology, this would be unlikely to leave enough flexibility to accommodate losses or environment noise in a real-world setup. Gaussian entanglement generation continues to increase at pace, though, so QSS may be routinely usable for such states in the near-to-mid term.

7.2.4 *Multi-level superpositions*

Finally, let us briefly address the question of larger Fock superpositions by considering the three-level superposition of the form

$$|\psi\rangle = \cos\theta_1|n_1\rangle + \sin\theta_1\cos\theta_2|n_2\rangle + \sin\theta_1\sin\theta_2|n_3\rangle, \tag{296}$$



(a) $\cos\theta_1|0\rangle + \sin\theta_1\cos\theta_2|1\rangle + \sin\theta_1\sin\theta_2|2\rangle$

(b) $\cos\theta_1|0\rangle + \sin\theta_1\cos\theta_2|1\rangle + \sin\theta_1\sin\theta_2|3\rangle$

(c) $\cos\theta_1|0\rangle + \sin\theta_1\cos\theta_2|1\rangle + \sin\theta_1\sin\theta_2|4\rangle$

(d) $\cos\theta_1|0\rangle + \sin\theta_1\cos\theta_2|2\rangle + \sin\theta_1\sin\theta_2|3\rangle$

(e) $\cos\theta_1|0\rangle + \sin\theta_1\cos\theta_2|2\rangle + \sin\theta_1\sin\theta_2|4\rangle$

$E_{1|2}(g) = 0.75$ $E_{1|2}(g) = 0.5$ $E_{1|2}(g) = 0.25$

Figure 32: Single-shot reconstruction fidelity at $g = 1$ for indicated superpositions and resource steering parameters. Grey line at $\theta_1 = \pi/2$ indicates the $\cos\theta_2|\psi_2\rangle + \sin\theta_2|\psi_3\rangle$ two-level superposition with no contribution from $|\psi_1\rangle$. Green and red lines at $\theta_2 = 0$, $\pi/2$ similarly indicate two-level superpositions of $|\psi_1\rangle$ & $|\psi_3\rangle$ and $|\psi_1\rangle$ & $|\psi_2\rangle$ respectively. Green and red stars indicate the $|\psi_2\rangle$ & $|\psi_3\rangle$ eigenstates while the $|\psi_1\rangle$ eigenstate exists along the entire axes at $\theta_1 = 0$, π. Black dashed line indicates the secure region for which $\mathcal{F} > 5/8$.

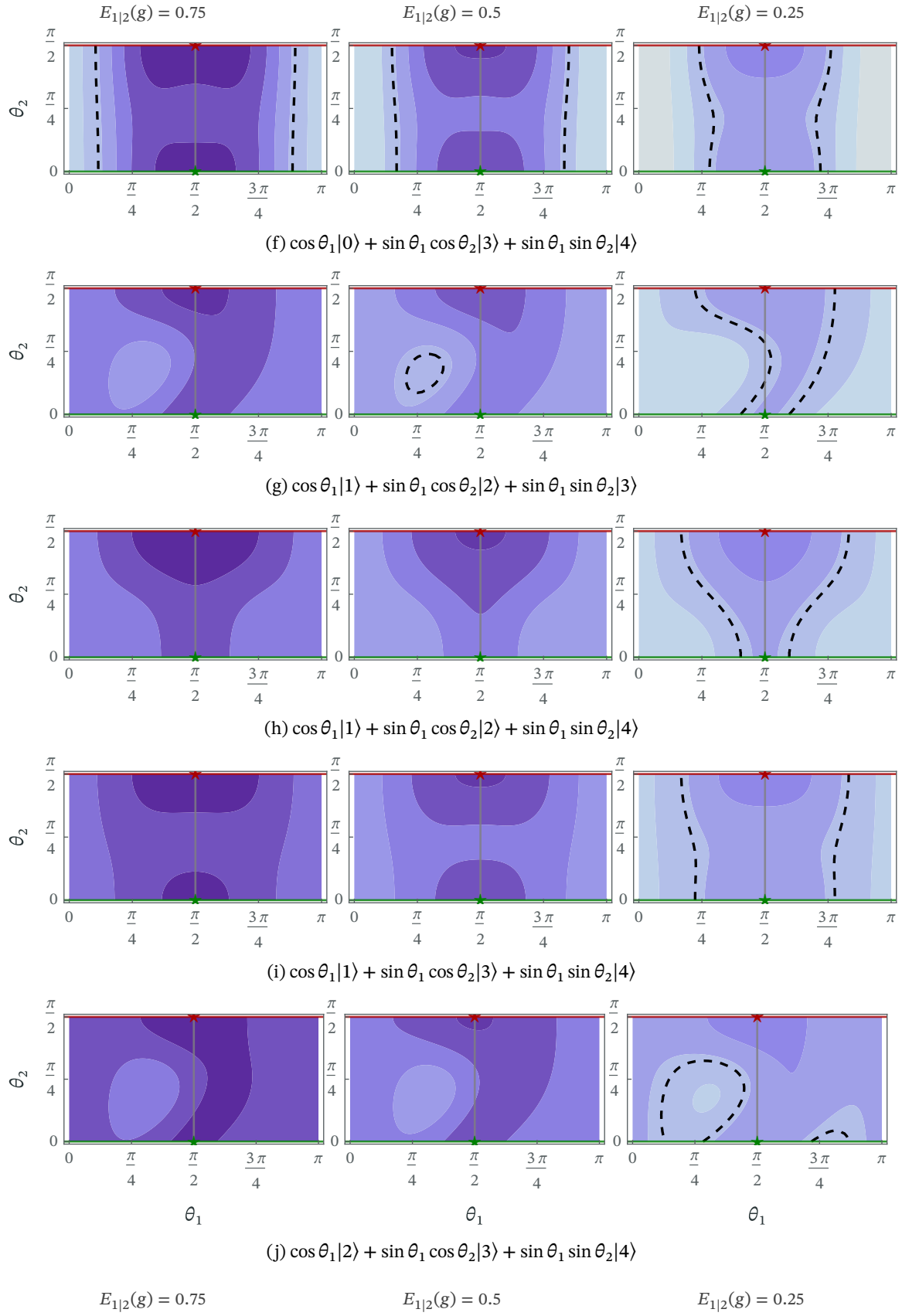
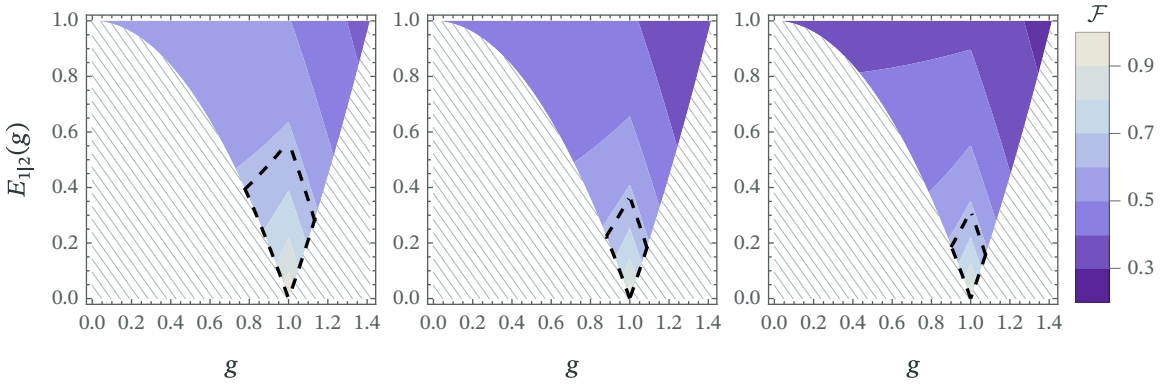


Figure 32 (continued): Single-shot reconstruction fidelity at $g = 1$ for indicated superpositions and resource steering parameters. Grey line at $\theta_1 = \pi/2$ indicates the $\cos\theta_2|\psi_2\rangle + \sin\theta_2|\psi_3\rangle$ two-level superposition with no contribution from $|\psi_1\rangle$. Green and red lines at $\theta_2 = 0,\ \pi/2$ similarly indicate two-level superpositions of $|\psi_1\rangle$ & $|\psi_3\rangle$ and $|\psi_1\rangle$ & $|\psi_2\rangle$ respectively. Green and red stars indicate the $|\psi_2\rangle$ & $|\psi_3\rangle$ eigenstates while the $|\psi_1\rangle$ eigenstate exists along the entire axes at $\theta_1 = 0,\ \pi$. Black dashed line indicates the secure region for which $\mathcal{F} > 5/8$.

for $\theta_1, \theta_2 \in [0, \pi]$. The single-shot fidelity for varying θ_1, θ_2 is shown for every combination of superpositions up to the $|4\rangle$ eigenstate in figure 32. As was the case for two-level superpositions, no improvement can be found over the fidelity achievable for a vacuum input so every superposition containing $|0\rangle$ (figures 32a to 32f) shows maximal fidelity at the $\theta = 0, \pi$ edges of the plot where that eigenstate is found.

For a number of the other superpositions, though, a more complex dependency on the superposition coefficients emerges in which the fidelity is not maximised for any individual eigenstate. Consider, as an example, the ($|1\rangle$, $|2\rangle$, $|3\rangle$) case in figure 32g or the ($|2\rangle$, $|3\rangle$, $|4\rangle$) case in figure 32j. In each case a reasonable fidelity can be achieved along the left and right extremes, indicating the first eigenstate, with a lesser fidelity achievable for the second and third eigenstates indicated by the red and green stars. In both cases, though, an 'egg' of significantly increased fidelity emerges in the bottom left of the plot. In the marginal cases, a fidelity of 5/8 indicating secure reconstruction is only achievable within this 'egg' — and is not obtainable for any eigenstate alone. That this phenomenon appears more pronounced for three-level superpositions than two-level superpositions supports the hypothesis that it derives from the state's increasing similarity to the coherent state as more levels are added to the superposition.

As previously, though, the more important result for our purposes is the expected average fidelity for an unknown member of the superposition class obtained by integrating over the θ_1, θ_2 space. This average reconstruction fidelity is shown for the three lowest-energy superpositions in figure 33. Three-level systems, consisting of more degrees of freedom, are more difficult to clone than two-level ones so the slightly looser threshold of $\mathcal{F} > 5/8$



(a) $|0\rangle$, $|1\rangle$, $|2\rangle$ superposition (b) $|0\rangle$, $|1\rangle$, $|3\rangle$ superposition (c) $|1\rangle$, $|2\rangle$, $|3\rangle$ superposition

Figure 33: Single-shot reconstruction fidelity achievable at $g = 1$ for three-level Fock states superpositions as labelled. Dashed line in each case denotes the $\mathcal{F} > 5/8$ region for which security can be guaranteed.

is sufficient to guarantee security here. This turns out to be only very marginally more difficult to achieve than for two-level superpositions. Using a two-mode squeezed vacuum at $g = 1$ as our usual standard entanglement source, the lowest-energy three-level superposition ($|0\rangle + |1\rangle + |2\rangle$) can be securely shared with only ~ 5.5 dB of resource squeezing ($E_{1|2}(g) \approx 0.57$). Indeed, every three-level superposition consisting of at most the $|4\rangle$ number state can securely shared using no more than 10 dB squeezing, with security achievable for many of them[5] with 8 dB squeezing.

This relative ease with which superposition states can be shared appears to continue as more levels are added to the superposition. Security can be guaranteed for every permutation of 4-level superposition with eigenstates up to $|4\rangle$ for only 9 dB TMSV squeezing, while the 5-level superposition class $\alpha_0|0\rangle + \alpha_1|1\rangle + \alpha_2|2\rangle + \alpha_3|3\rangle + \alpha_4|4\rangle$ needs only ~ 7.75 dB resource squeezing to certify security.

7.3 CONCLUSION

In this chapter we have considered the use of the Gaussian quantum state sharing protocol outlined in chapter 4 for sharing discrete-variable particle number states and their superpositions.

Such states are essential to discrete-variable quantum information processes, but the generation and distribution of non-Gaussian entanglement is a comparatively difficult task. Exploiting Gaussian entanglement for the communication of discrete-variable states then allows one to combine the benefits such states bring to some quantum information tasks with the relative ease of Gaussian entanglement generation.

We have shown that low particle-number states can be shared with high fidelity using readily available Gaussian entanglement. Further, any Fock eigenstate can be shared with arbitrarily good fidelity given a suitably strong entanglement resource. The space of number-state inputs this protocol is useful for will only increase in size as better entanglement sources are developed in the future.

For this class of states, however, security cannot be guaranteed through the no-cloning theorem, which may limit the protocol's usefulness. Other tools may be available for future to analyse the security of Fock-eigenstate protocols; we have considered security based on other considerations to be out of the scope of this thesis.

[5] The $|0\rangle + |1\rangle + |2\rangle$, $|0\rangle + |1\rangle + |3\rangle$, $|0\rangle + |2\rangle + |3\rangle$, and $|1\rangle + |2\rangle + |3\rangle$ superpositions all have a squeezing requirement of around 8 dB.

Considering the Fock superposition states, we have demonstrated that security is possible for a large number of classes of superposition, many with existing, readily-available entanglement sources. Many more are securely sharable with cutting-edge entanglement sources. Beyond two-level superposition states, we have also demonstrated security for a number of three- or four-level superpositions. Although we have only considered a subset of the possible Fock superpositions, we outlined in section 6.3 an algorithm to use for the study of any Fock superposition state.

The results we have found in this chapter indicate that there is significant potential for this protocol to be used for the secure distribution of some of the most commonly used qubit-like and qutrit-like states using only Gaussian entanglement with limited difficulty over sharing Gaussian states. Even modest increases in our ability to generate Gaussian entanglement should expand the set of states for which this protocol can be guaranteed to be secure.

8

CONCLUSION

In this part of the thesis, we have presented a quantum state sharing protocol utilising continuous-variable sources of entanglement for the secure distribution of Gaussian and Fock-like states.

After introducing the protocol in chapter 4, we demonstrated in chapter 5 that it is useful for the secure sharing of coherent states even at very low levels of entanglement. In particular, assuming an input set making up the full range of coherent amplitudes, *any* 2-mode Gaussian state exhibiting *any* degree of EPR-steering can be used for secure QSS. Notably, this is a looser condition than is necessary for teleportation, which requires the existence of EPR-steering in both directions and so excludes the use of weaker single-directional entanglement [45]. Further, our QSS protocol is less state-distorting than teleportation, so for asymmetric steering of any quality a better reproduction of the initial state is obtained from the use of a quantum state sharing protocol. The protocol pays for this, however, by protecting against a slightly weaker threat model. In contrast to quantum teleportation — which is secure against the compromise of any transmission channel — quantum state sharing only defends against attackers unable to gain access to a majority of the communication channels. Assuming the selected state transmission methods are suitably independent — for example, fibre-optical cables taking different routes along the network without overlapping nodes — we believe this represents an acceptable level of risk for many consumer uses. It may be that the lower resource cost and higher-quality reproduction of QSS protocols opens the use of quantum security for more applications which may not justify the increased cost involved in teleportation-based communication.

We have also demonstrated the protocol's effectiveness for the full-range of Gaussian states, including squeezed states and thermal states. Although we are unable to present a security analysis for thermal states as the question of their optimal cloning remains open, we have shown that arbitrarily good reconstructions can be achieved at any level of thermality or squeezing given a suitably-entangled resource. The QSS protocol as presented here is then ready to be immediately-applicable to secure thermal state distribution once their cloning properties are more well understood. Under the assumption that the Gaussian input states remain pure, we have de-

rived stronger security conditions under two different assumptions. First, relying on no unproven conjectures, we show that any squeezed-state QSS protocol that only accepts states up to some defined minimum level of squeezing is provably secure, and derive the level of EPR-steering that is necessary for a set envelope of allowed squeezed states. This sufficient condition for security does not account for the fact that squeezed states are harder to clone than coherent states and assumes the worst-quality reconstruction for all input states, so does not represent a tight bound on security. It has been conjectured [56] that the optimal coherent-state cloning machine is also optimal for squeezed states. Under the assumption that this is true, we have additionally shown that the sharing of squeezed states drawn from a Gaussian distribution is secure under the same conditions as for coherent states — that the resource state exhibit any EPR-steering. As this remains an unproven hypothesis, however, we would caution against relying on this tighter security condition at this time. We present it here in the hope that this conjecture will be proven, in which case our protocol can be immediately trusted for such a class of squeezed states.

In chapter 7 we have considered the use of quantum state sharing as a so-called 'hybrid protocol', utilising Gaussian entanglement for the sharing of discrete-variable Fock-like states. We compared two forms of the protocol, in which the output state is corrected for any de-amplification occurring during state reconstruction and in which it is not, finding that the optimal protocol is dependent on the specific distribution of Fock states the protocol is intended for. An interesting extension to this analysis would be to consider the information obtainable about which Fock state is encoded within one of the shares, to potentially certify security under certain limited circumstances. Despite the input states being orthogonal, the set of potential shares are non-orthogonal and so will not provide perfect knowledge of the input photon number.

We further found that security can be guaranteed for the sharing of Fock superposition states for a large variety of common two- and three-level superpositions. In many of the cases studied, the required level of entanglement is well within what is experimentally achievable with current technology, indicating that QSS is ready as a hybrid protocol for qubit states. Indeed, the $\alpha_0|0\rangle + \alpha_1|1\rangle$ state required only just over 5 dB of squeezing in a two-mode squeezed state resource, a level readily achievable in any quantum optical setup. Although the results presented here represent a full characterisation of the selected superpositions, it would be helpful to have an analytic understanding of the output of an arbitrary superposition, as we do for the Fock eigenstates. Such an understanding could be obtained

by solving the partial-trace integrals for the coherence terms in a similar fashion to the solutions found for the eigenstate terms in appendix B. This would be a fruitful area not only for the study of quantum state sharing, but of hybrid quantum information in general.

8.1 OUTLOOK

As we move towards a quantum-connected future, a wide selection of quantum protocols will be necessary to ensure the security of the quantum internet. Quantum state sharing as a protocol is likely to take a role alongside quantum teleportation and remote state preparation in the secure distribution of quantum information.

Alongside the natural applications QSS has to the secure distribution of quantum information, there has been promising research into the use of QSS schemes for blind computing [34] and transmission loss-recovery [35] applications. Further research is required to fully understand the potential for QSS schemes to be useful for blind computation — including whether such potential exists for the QSS scheme outlined in this thesis and how QSS-blind computing compares to other protocols achieving similar tasks. Quantum loss-recovery is immediately achievable from the quantum state sharing protocol with no further adaptation necessary, however it would be interesting to see whether the redundancy built in to QSS protocols can be used for wider classes of error correction.

A particularly interesting avenue for continued research in this area is the further development of QSS as a hybrid protocol. Despite their usefulness for a large number of quantum information tasks — most notably quantum computation [79, 90], there are a number of fundamental and experimental constraints on the transfer of discrete-variable states [67, 68]. It has been shown for quantum teleportation [91] that under optimal choice of protocol amplification the photon-increasing effects from the thermal noise inherent to Gaussian entanglement can be eliminated. Consequently, it was shown that qubit states can be teleported using CV entanglement without the need to expand the size of their Hilbert space. It would be interesting to apply similar techniques to this protocol to understand whether a similar correction can be made for hybrid quantum state sharing.

Another direct extension of this work may be the investigation of global multi-mode QSS protocols. As discussed in section 5.4, the naïve application of single-mode QSS locally to each mode of a multi-mode secret state has the effect of destroying the state's entanglement. However, the question of whether a global protocol — one acting on all modes of the

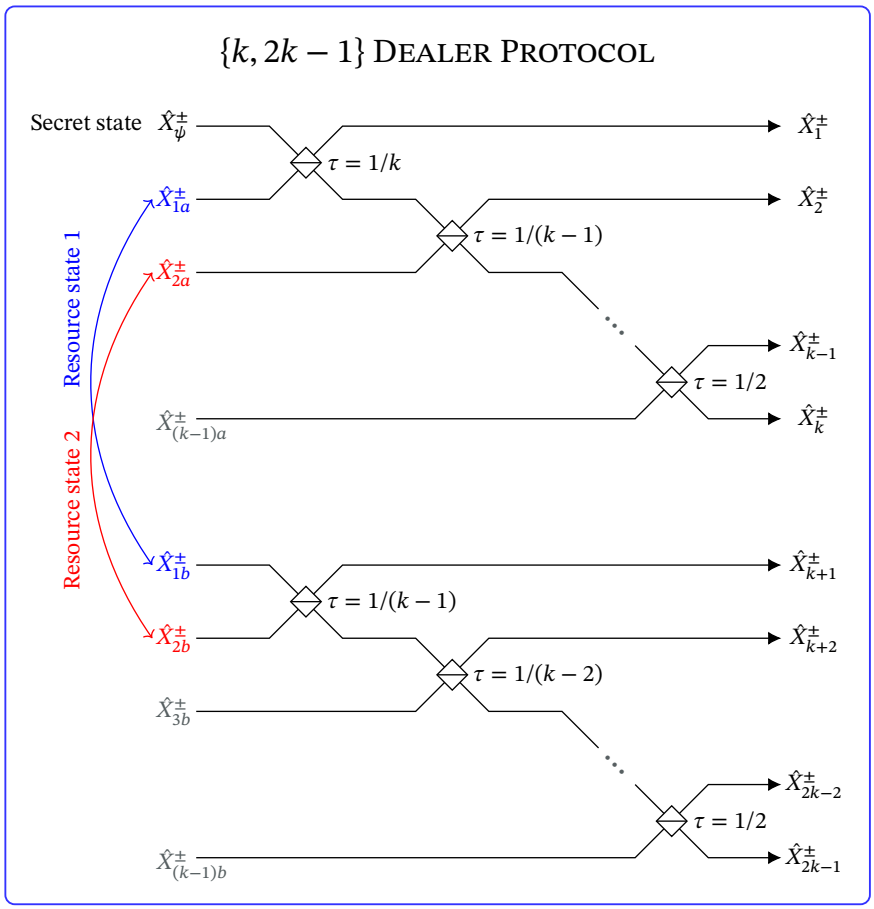


Figure 34: A proposed dealer protocol for general $\{k, 2k-1\}$-threshold QSS. $k-1$ independent two-mode entangled resource states are prepared and passed to the dealer alongside a secret state ψ. The secret state is then mixed with one mode from each resource state in a cascading series of beamsplitters with transmissivity $\tau = 1/(k+1-i)$ for the ith beamsplitter to prepare the first k shares. Each of these shares is left with an equal $1/k$ contribution from the secret state. The remaining $k-1$ modes of the resource states are mixed on a similar setup of cascading beamsplitters with $\tau = 1/(k-i)$ to produce the final $k-1$ shares. Each of these latter shares contain no direct information about the state of ψ but act as keys to disentangle the resource states from the upper set of shares. Any selection of k of these shares is capable of reconstructing the original secret state.

secret collectively — that preserves entanglement exists has been left open. Indeed, in the field of optimal quantum cloning, it has been shown that such global protocols outperform local ones [22]; there is no reason to believe the same would not be true of quantum state sharing.

Perhaps the most interesting avenue for future research, though, is the extension to general $\{k, n\}$-threshold QSS. Recently, work in this area been carried out by two undergraduate students at the University of St Andrews investigating the sharing of coherent-state secrets between five players. Dan Travers [92] investigated the use of dual two-mode squeezed vacuum states to achieve this task, and found that secure QSS between five players was possible with presently-achievable entanglement levels. This work was then extended by Rosie Gittings [93] to allow for the use of two arbitrary, in-general asymmetric and impure resource states of the sort considered in this thesis. She found that, when the two resource states used were identical, they had to exhibit steering of $E_{1|2}(g) \leq 0.5$ to achieve secure QSS; a necessary steering parameter half the value of that required for $\{2, 3\}$ quantum state sharing. It would be interesting to see whether, as conjectured by Gittings, this is the beginning of a trend with $\{k, 2k-1\}$-threshold QSS requiring steering of $E_{1|2}(g) \leq 1/(k-1)$.

A dealer protocol for general $\{k, 2k-1\}$-threshold QSS satisfying the constraints outlined in Ref. [43] (see section 4.1.1) can be trivially constructed through the use of a cascading series of beam-splitters. A proposed setup for such a dealer protocol is shown in figure 34. The observation that the worst reconstruction would be found using the bottom $k-1$ shares in figure 34 (those with no ψ contribution) alongside only one of the top shares may enable a general discussion of security requirements. However, we caution that unlike the $\{2, 3\}$-threshold scheme this dealer protocol has not been proven optimal — it is simply *one* dealer protocol that works.

Finally, let us note that work is ongoing at the Walther-Meißner Institute for Low Temperature Research at the Bavarian Academy of Sciences and Humanities in Munich looking at experimental demonstrations of $\{2, 3\}$-threshold QSS. Led by Karolina Weber under the supervision of Kirill Fedorov, this work is investigating the ability for the protocol outlined in this thesis to be used in the microwave domain — the region of the electromagnetic spectrum used for mobile and wireless communication. This follows similar work from that group demonstrating the successful teleportation of microwave states [53]. Results regarding the quantum state sharing of microwave coherent states at high fidelity are presently being prepared for publication [94, 95], with further research into the demonstration of hybrid QSS sharing Fock states planned.

Part II

QUANTUM METROLOGY FOR FIELD GRADIENT ESTIMATION

9

INTRODUCTION

Quantum sensing — the art of designing quantum states that enable deeper investigation than would otherwise be possible — is perhaps the most immediately-promising quantum technology, with applications already present and significant continued development of quantum measurement technology expected in the next decade [1].

The precise measurement of physical systems is fundamental not only to the academic study of related phenomena, but to the increasingly-powerful technology built on such measurements for healthcare, navigation and research applications. Our ability to measure any physical property is strictly limited, however, with increases in precision requiring quadratic increases in infrastructure or measurement time [96]. Improving the quality of measurements therefore necessitates large increases in equipment size, power usage, and expense. Often, one or the other of these is not feasible and measurement precision is capped by practical considerations. Taking a quantum approach instead, in which a network of measurement probes are entangled prior to being used for a measurement allows one to sidestep these limits and gain more information from the same system size and time. In fact, taking this quantum approach reduces the previously-necessary quadratic increase in system size to a linear increase [96], offering increasing advantage as the size of the available measurement system grows.

In this part of the thesis, we consider the usefulness of quantum metrological approach to the estimation of field gradients. Precise estimations of gradients in the magnetic field are extensively used in a number of real-world tasks from the location of underground mineral deposits [97] and the detection of seaborne mines [98] to the imaging of the human brain [99, 100]. Improving the resolution with which such measurement can be made is therefore of great interest across a number of fields.

Although the measurement of fields is an area in which quantum metrology is anticipated to begin to deliver returns in the short-term [1], there has been little study in to the quantum entanglement networks that are best suited to this task in the abstract. While the generation of the 'designer states' necessary to implement the approach discussed in this thesis remains a longer-term goal of quantum metrology, a greater understanding

of the conditions under which certain quantum states are optimal may help inform the design of future implementations.

We will first consider the estimation of linear field gradients, extending existing results to consider the estimation of multiple gradient directions simultaneously. We will also consider the impact real-world noise fields have on these results, and the adjustments to the entanglement networks necessary to mitigate these noise fields as much as possible. We will show that, while in the ideal case sequential estimation of the gradient directions individually cannot be improved upon, in the presence of noise it becomes optimal to measure multiple directions simultaneously.

9.1 A BRIEF INTRODUCTION TO QUANTUM METROLOGY

Before we go on to consider the estimation of field gradients in detail, in the remainder of this chapter let us first cover the tools from classical and quantum estimation theory we will use. This is by no means a complete overview of the field, nor a formal discussion of these concepts on solid statistical principles;[1] rather its intent is to cover only those core ideas necessary for this thesis. For a more complete introduction to the field, including a full derivation of the (quantum) Fisher information and Cramér-Rao bounds, the reader is directed to the excellent review article by Jasminder Sidhu and Pieter Kok in Ref. [101] or book by Carl Helstrom [102].

At their core, all quantum metrology protocols derive from the same principle: estimating the relative phase between eigenstates of a quantum superposition acts as a proxy for some physical parameter of interest. Consider, for example, the effect of a quantum unitary $\hat{U}$ on a single-mode spin state, $|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$. The two eigenstates will accumulate phase in a different way, depending on the form of the unitary, such that after its operation the state will take the form

$$\hat{U}|\psi\rangle = \alpha\, \mathrm{e}^{\mathrm{i}\,\phi_1}|\uparrow\rangle + \beta\, \mathrm{e}^{\mathrm{i}\,\phi_2}|\downarrow\rangle. \tag{297}$$

Some amount of information about the properties of the unitary is now encoded within the phases of the state. When one knows the structure of the unitary, an estimation of the phase difference can then be directly converted into an estimation of the unitary itself.

[1] In particular, many of these results are only valid in the asymptotic case, but represent a 'good enough' approximation under the assumption that one will be running multiple rounds of measurements.

Now, let us imagine that we have access to a network of N such probes. It is likely that some eigenstates of this wider system will obtain more information than others; for example, perhaps the impact of opposing spins cancel such that the $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$ eigenstates accumulate no useful information. By limiting oneself to the classical case[2] in which no entanglement is permitted between the probes, designing states that maximise the contribution from the useful eigenstates may be impossible. In allowing for an arbitrary entanglement structure within the probe network, on the other hand, the collective superposition that maximises information gain can be constructed regardless of the specific combination of eigenstates this requires. It is in this increased flexibility that quantum metrology finds its advantage over classical approaches.

9.1.1 *Classical parameter estimation*

The aim of any parameter estimation, classical or quantum, is to minimise the space in which we believe the true value of the parameter to lie, termed the 95% confidence interval. This is the region we expect will fail to contain the true value in only 1 in 20 experimental runs.[3] Assuming measurements to be normally distributed[4] around the true value, the size of this confidence region is proportional to the standard deviation of the measurement set [103].

More formally, let us imagine the parameter has a true value denoted by θ, and that in estimating this value we have access to a related observable $\hat{X}$, not necessarily of the parameter itself. As no measurement process is perfect, measurement outcomes x will be drawn from some probability distribution $p(x|\theta)$ centred around the true value of the parameter. Assuming that the form of this probability is known — that is to say, assuming we have a model describing the physical process — an estimator function, $\hat{\theta}(\{x\})$, can be constructed that takes as input a set of measurement outcomes on $\hat{X}$ and returns a best estimate of the value of θ. The accuracy

[2] In the field of quantum metrology the 'classical' case is taken to mean the case in which entanglement is forbidden but superposition is allowed. Strictly speaking, we are discussing the so-called classical-quantum case in which entangled measurements are permitted. It has been shown [96] that this gives no improvement over the fully classical case, however, and so the distinction is not important here.

[3] For reasons discernible only to statisticians, this is *not* the same as there being a 95% likelihood of the true value lying within this interval in any given experiment.

[4] The central value theorem states that when a suitably large set of independent measurements are drawn from the same probability distribution they will tend towards a normal distribution [104]. Each of the problems we consider here consist of measurements drawn in such a way, and so we will assume the estimator to be normally-distributed throughout.

of this estimator can then be assessed by the average expected difference between it and the true value,

$$\sigma^2(\hat{\theta}) = \langle(\hat{\theta}(\{x\}) - \theta)^2\rangle_{p(\{x\}|\theta)} = \int d^n x\, p(\{x\}|\theta)\,(\hat{\theta}(\{x\}) - \theta)^2, \tag{298}$$

for a specific value of the true parameter θ. For an unbiased estimator (one that does not systematically over- or under-estimate the true value) this error coincides with the estimator's variance, [101]

$$V(\hat{\theta}) = \langle\hat{\theta}^2\rangle_{p(\{x\}|\theta)} - \langle\hat{\theta}\rangle^2_{p(\{x\}|\theta)}, \tag{299}$$

and so in 95% of experiments the true value will lie within the range

$$\text{confidence interval} = [\hat{\theta} - 2\sqrt{V}, \hat{\theta} + 2\sqrt{V}]. \tag{300}$$

This quantification of the information we gain from the measurement set is clearly dependent on our choice of estimator function, though; some functions on $\{x\}$ are going to produce a better estimate of θ than others. In assessing an observable $\hat{X}$ more generally, a better question to ask might be: how much information is *fundamentally exposed by the observable*?

Ultimately, the information obtainable from an observable will depend on the extent to which the parameter θ is reflected in its probability distribution, $p(x|\theta)$. The greater the impact a small change in θ has on the distribution of measurement outcomes, the better one will be able to see in them the true value of θ. This tendency of the true value to impact the measurement outcomes can be quantified through the probability distribution's logarithmic derivative, or its score, as [101]

$$L_\theta = \frac{\partial \ln p(x|\theta)}{\partial\theta}. \tag{301}$$

Averaging the variance[5] of this score over the full range of possible measurement outcomes, we obtain an overall measure of the information obtainable from the observable,

$$F_\theta = \langle L^2_\theta\rangle_{p(x|\theta)}, \tag{302}$$

termed the Fisher information (FI) [101]. Notably, this measure (alongside the others introduced earlier in this section) depends on the true value of θ — some areas of the parameter space may be easier to characterise than others.

[5] given only by the expectation of the square, as the expectation value of L itself is 0 due to the symmetry implied by the assumption the estimator is unbiased.

The Fisher information scales linearly with the number of measurements, such that after m measurements — or, equivalently the measurement of a separable system of m identical subsystems — it is given by

$$F_m = m F_1. \tag{303}$$

Similarly, the Fisher information for two different measurements can be combined as

$$F_{\text{total}} = F_1 + F_2. \tag{304}$$

As classical measurements are by-definition separable, this acts to bound the information obtainable from such a network of probes. This limit — allowing at best an N-scaling with probe number in FI and so a $\sqrt{N}$-scaling in confidence interval — is termed the 'shot noise limit' [96].

This Fisher information bounds the variance of any estimator on θ after m measurements of the observable through the Cramér-Rao bound, as

$$V(\hat{\theta}) \geq \frac{1}{m F_\theta}, \tag{305}$$

which is saturated only by the optimal estimators [105]. It has been shown [106] that for a large-enough sample size this bound is always achievable, so the best-possible confidence region size can be further expressed as the equality

$$\text{confidence region size} = \frac{4}{\sqrt{m F_\theta}}, \tag{306}$$

again potentially a function of the true value of θ.

9.1.2 *Quantum parameter estimation*

In assessing quantum metrology protocols, we abstract this concept one step further by considering the quantum state, $\hat{\rho}$, from which these measurements are drawn. Rather than asking how much information could be obtained from a given observable, we instead ask how much information is fundamentally contained within the state irrespective of the measurement used to access it. To answer this question, we turn to the quantum Fisher information (QFI), $F_Q(\hat{\rho})$, which in turn bounds the classical FI as $F(X) \leq F_Q(\hat{\rho})$. It has been shown [101, 107] that under optimal selection of observable this bound is saturated — and thus the QFI and FI coincide —

so in the remainder of this thesis we will drop the distinction and denote the QFI simply as F.

The Cramér-Rao bound is then immediately applicable to the QFI, such that the estimation variance is bounded by

$$V(\hat{\theta}) \geq \frac{1}{m F_Q(\hat{\rho})} \tag{307}$$

after m measurements using identically-prepared states $\hat{\rho}$. Notably, although we remain limited by the m scaling in number of experimental runs, the effect of the size of the measurement system has been moved into the calculation of the quantum Fisher information itself, so the shot-noise limit no longer applies. Instead, it has been shown that a different limit — termed the Heisenberg limit — applies to for entangled states that allows an $1/N^2$-scaling with probe number in QFI and so a $1/N$-scaling in confidence interval [96]. As well as potentially delivering more information from a set number of probes, taking a quantum approach allows for a much faster increase in information from increasing network size than would be possible classically.

When the parameter of interest arises from a quantum unitary of the form $\hat{U} = \exp[-\mathrm{i}\, t\hat{H}(\theta)]$ for Hamiltonian $\hat{H}$, the QFI can be found directly through an intermediate generator matrix [108, 109],

$$\mathcal{H}_\theta = \mathrm{i}\,(\partial_\theta U^\dagger)U. \tag{308}$$

This matrix, independent as yet of the choice of state, contains a complete description of the unitary's phase behaviour. For a pure input state, $|\psi\rangle$, the QFI is then simply the expected variance of this generator matrix with respect to the state,

$$F_\theta = 4\langle \Delta^2 \mathcal{H}_\theta \rangle_\psi, \tag{309}$$

for

$$\langle \Delta^2 \mathcal{H}_\theta \rangle_\psi = \langle \psi | \mathcal{H}_\theta^2 | \psi \rangle - \langle \psi | \mathcal{H}_\theta | \psi \rangle^2. \tag{310}$$

The QFI for a mixed state, $\hat{\rho}$, can be found by first splitting the state into its pure state components, $\hat{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, as

$$F = \sum_i 4 p_i \langle \Delta^2 \mathcal{H}_\theta \rangle_i - \sum_{i \neq j} \frac{8 p_i p_j}{p_i + p_j} |\langle \psi_i | \mathcal{H}_\theta | \psi_j \rangle|^2, \tag{311}$$

where $\langle\cdot\rangle_i = \langle\psi_i|\cdot|\psi_i\rangle$ denotes the expectation over eigenstate ψ_i. Notably, the QFI is a convex function on the set of quantum states [101], so is always maximised by a pure state probe network. In searching for optimal states one can therefore restrict the search space to the set of pure states, vastly reducing the required computational resources. Although this remains true in the presence of noise — the optimal initial setup will continue to be a pure state [110] — by the time one makes a measurement the system will have degraded into a mixed state and the full form of the QFI from equation (311) will remain necessary.

9.1.3 *Multi-parameter estimation*

Our discussion of Fisher information and variance measures thus far has been in the context of the measurement of a single parameter. In the bulk of this part of the thesis, though, we will be interested in the simultaneous measurement of multiple parameters that may or may not interact.

As might be expected, in considering multiple parameters the estimator variance is replaced by a covariance matrix over a set of estimators, $V(\{\hat{\theta}_i\})$, representing the individual variance for each parameter and any correlations between them as

$$V_{i,j} = \langle\hat{\theta}_i\hat{\theta}_j\rangle - \langle\hat{\theta}_i\rangle\langle\hat{\theta}_j\rangle. \tag{312}$$

This covariance matrix defines a 90% confidence region analogous to the single-parameter confidence interval, which has area proportional to the square root of its determinant, $\sqrt{\det V}$ [103]. The quantum Fisher information is similarly replaced by a quantum Fisher information matrix which reproduces the single-parameter QFIs along the diagonal and represents confounding relationships between the parameters (that reduce our ability to distinguish the impact of each parameter in the observable) as off-diagonal elements. The same generator matrices $\mathcal{H}_\theta = \mathrm{i}(\partial_\theta U^\dagger)U$ can be used to find the elements of the QFI matrix as [109]

$$F_{\alpha,\beta} = \sum_{i=1}^{M} 4p_i\mathrm{cov}_i(\mathcal{H}_\alpha,\mathcal{H}_\beta) - \sum_{i\neq j}\frac{8p_ip_j}{p_i+p_j}\mathrm{Re}\Big(\langle\psi_i|\mathcal{H}_\alpha|\psi_j\rangle\langle\psi_i|\mathcal{H}_\beta|\psi_j\rangle\Big), \tag{313}$$

which reduces to the single-parameter QFI from equation (311) when $\alpha = \beta$.

The Cramér-Rao bound is reproduced in matrix form as the positive semi-definiteness condition, [101] [6]

$$V \geq \frac{1}{m} F^{-1}. \tag{314}$$

In contrast to single-parameter estimation, however, this bound is not always achievable. Recall that the Cramér-Rao bound is only saturated for the optimal choice of measurement operator; in the quantum context, though, if the optimal observables for two parameters do not commute they cannot both be measured on a single state. While each parameter can be estimated maximally individually, they may not necessarily all be able to be estimated maximally at the same time. When the generator matrices for two parameters commute, such that

$$[\mathcal{H}_\alpha ; \mathcal{H}_\beta] = 0, \tag{315}$$

though, the parameters represent commuting elements of the Hamiltonian and can always be measured simultaneously [109]; in this case the Cramér-Rao bound is automatically achievable regardless of the form of the state.

Finally, although the diagonal elements of the QFI matrix coincide with the individual QFI of each parameter, we caution the reader that these elements alone do not represent the amount of information obtainable about that parameter. These single-parameter QFIs are only achievable when *all other relevant parameters are perfectly known*; when multiple unknown parameters are present, the full matrix inversion must be performed to find the parameter variances. Indeed, as the QFI matrix is positive definite,[7] the optimal variance of a single parameter,

$$V_{i,i} = (F^{-1})_{i,i} \geq \frac{1}{F_{i,i}}, \tag{316}$$

will be strictly greater than the inverse of the single-parameter QFI unless no covariances with the other parameters exist [111]. The off-diagonal elements of the QFI act to reduce the amount of information available, and represent our inability to separate information obtained about some linear combination of them into information about each individually. As

[6] The positive semi-definiteness condition is equivalent to saying that for every $x \in \mathbb{R}^n$, $x^T(V - \frac{1}{M}F^{-1})x \geq 0$.

[7] Strictly speaking, the QFI matrix need only be positive semi-definite. However, a positive semi-definite matrix with non-zero determinant is by-definition promoted to being positive definite so the only time equation (316) may not apply is when the QFI matrix is singular and the covariance does not exist anyway! Such cases represent no capture of information and so we can assume the QFI matrices we discuss in this thesis to always be positive definite.

the QFI matrix can always be diagonalised through a coordinate rotation, the existence of off-diagonal elements can alternatively be considered indicative of the chosen frame of reference being non-optimal for the given measurement setup. In the tasks we consider in this thesis, the frame of reference — Cartesian coordinates with the measurement device defining the origin — is essentially arbitrary so such a rotation should always be possible. We will consider this point further in section 9.2.2 when we introduce the determinant of the variance as a reference-frame-independent measure.

This distinction between the QFI matrix diagonal elements and the individual QFIs is an important consideration even when one is only interested in estimating a single parameter, as one's ability to estimate it may be impeded by the absence of knowledge of related parameters. Such 'nuisance parameters' must be included in the QFI matrix alongside the desired parameters for the calculation of the covariance matrix before they can be discarded.

9.2 NUMERICAL OPTIMISATION METHODS

The optimal states presented in this part of the thesis have all been found through a numerical optimisation process performed over the full set of n-particle states. Let us outline the optimisation methods we have used for this here.

9.2.1 *Optimisation algorithms*

The core local optimisations are performed using the Scipy [112] implementation of a gradient following optimisation algorithm termed L-BFGS-B [113]. At each step of the optimisation, this algorithm estimates the gradients and (indirectly) the Hessian matrix of the function. Small steps are then repeatedly taken in the direction of steepest descent, until either a local minima is found or the optimisation exceeds the maximum allowed number of rounds.

Optimisations such as this work very well for scenarios where a local minima can be trusted to coincide with the global minima. When this is not the case, however, there is no guarantee that the result found by the optimisation algorithm is indeed the optimal result. A number of pseudo-global optimisation algorithms have been developed that attempt to correct for this deficiency in some way. One such approach to global optimisation is the Basin Hopping algorithm proposed by Wales and Doye in 1997 [114],

Algorithm 2 Population basin hopping (PBH)

▷ *Initialise the population*
randomly select m pops within the state-space, $X' = \{X'_1, X'_2, \ldots X'_m\}$
$X = \{\texttt{minimise}(X'_1), \texttt{minimise}(X'_2), \ldots, \texttt{minimise}(X'_m)\}$
▷ *Run optimisation rounds*
REPEAT
    FOR $k = 1 \ldots m$ DO
        $Y'_k = \texttt{perturb}(X_k)$
        $Y_k = \texttt{minimise}(Y'_k)$
    FOR $k = 1 \ldots m$ DO
        $C =$ nearest X pop to Y_k
        IF $\texttt{dist}(C, Y_k) \geq \texttt{min\_dist}$ THEN
            $C = \texttt{worst\_of}(X)$
        IF $f(Y_k) \geq f(C)$ THEN
            $c =$ index of C in X
            replace $X_c = Y_k$

which involves the repeated operation of the local optimisation process, finding a series of local minima and returning the best of them. After the first local optimisation, its result is stored and a 'hop' is made in a random direction in an attempt to escape the catchment of that local minima, or 'basin'. Ideally, this hop will be such that the next result consists of a neighbouring basin. After the new basin is found, it either replaces the old minimum if it is an improvement or is discarded. In either case, the hop-and-minimise process is repeated to find the next basin for a set number of rounds. This approach is highly effective for the computational chemistry problems for which it was designed — such energy minimisations tend to take the form of a single funnel of minima of increasing depth, leading towards the single lowest energy state [115, 116].

Basin hopping algorithms are less effective, however, when the basins are potentially arranged across multiple of these funnels. While highly capable at escaping local minima to continue moving down a funnel, basin hopping is not designed to be able to escape a so-called local funnel. Take as an example a scenario in which the GHZ state is optimal but the Dicke state also outperforms separable probes. Sampling the entire state-space, one might then find multiple funnels — one leading towards the Dicke state and another towards the GHZ state. The basin hopping algorithm may well get stuck in the Dicke state funnel and never find the GHZ state at all!

To prevent this, we have deployed an adaptation of basin hopping developed by Grosso, Locatelli and Schoen [117, 118] termed population

basin hopping (PBH). We used a custom Python implementation of this algorithm, further augmented with a simple adaptive step-size algorithm.

The core approach taken by population basin hopping is to run multiple basin hopping optimisations (the population, or 'pops') in tandem — finding multiple funnels — before selecting the best result as the presumed global optimum. Coordination between the individual basin hopping optimisations ensures they keep a distance from each other so do not fall into the same funnel. The algorithm is outlined in algorithm 2. The pops begin randomly distributed across the parameter space. At the beginning of each round, every pop independently performs a random hop from their previous position and finds the local minimum around that point, exactly in accordance with the standard basin hopping algorithm. The PBH algorithm then decides whether each of these new minima displaces one of the previous minima according to the following rules.

- If the new minimum is within some set `min_dist` of any old minimum, it is compared against the closest of the old minima. If it beats the nearest old minimum the nearest old minimum is replaced by the new minimum.
- Otherwise, the new minimum is compared to the worst of the old minima. If it beats the worst old minimum, it takes its place.

This process is then repeated a set number of times. In our setup, we have used a population of 10 pops running for 75 rounds for a total of 750 local optimisations per scenario. In every case, the optimum was seen to be stable by the end of these rounds indicating a high likelihood that the global optimum had been found. To establish pop distance we have used the trace distance,[8] [119]

$$\texttt{dist}(\psi, \phi) = \sqrt{1 - |\langle\psi|\phi\rangle|^2}, \tag{317}$$

with a minimum distance of $\texttt{min\_dist} = 0.15$ enforced between the pops.

For our purposes, we have augmented this process with a further adaptive step size algorithm. For an ideal optimisation, the step size should be precisely the distance between two basins while the `min_dist` should be precisely the distance between two funnels. The closer these parameters are to their ideal values the more effective the algorithm will be. As we have no prior estimate of whether multiple funnels exist, or the distance between them, our conservative minimum distance between pops runs the risk of

[8] For pure states this trace distance metric is equivalent to enforcing a maximum fidelity between pops.

multiple pops being trapped in the same funnel. Although this would not prevent the establishment of the bottom of the funnel, with one pop always allowed to move towards the centre, the remaining pops would be stuck along the periphery making small hops that never escape the funnel. To avoid these pops being locked out of the optimisation, at the end of each round the step size for each pop is changed according to algorithm 3. If a

Algorithm 3 Adaptive step

```
FOR k = 1 ... n DO
    IF X_k has not changed in 20 rounds THEN
        increase step_size[k] by 10%
    ELSE
        reset step_size[k] to default_step_size
```

pop has not been able to move for 20 rounds, it starts to be allowed to make larger and larger jumps to attempt to get out of its funnel. Eventually, these jumps will be large enough that the pop is effectively reset to starting at random anywhere within the parameter space. This augmentation allows pops that would otherwise remain useless — those that have already found the bottom of their funnel or are stuck on the periphery of another pop's funnel — to continue to contribute by searching elsewhere. The results of each of these jumps are compared to the old results in the usual way, so the bottom of a funnel will not be lost when its pop jumps unless the pop discovers a better basin.

9.2.2 *Optimisation metrics*

The output of each measurement setup will be a 2×2 quantum Fisher information matrix collectively representing the information obtainable from the setup. To directly compare two states, we must select a metric function,

$$\mathcal{M} : \mathbb{R}^{2\times 2} \to \mathbb{R}, \tag{318}$$

that maps these matrices to a scalar value representing an overall view of the state's usefulness. This allows the optimisation algorithm to compare two QFI matrices as $\mathcal{M}(F_1) \overset{?}{<} \mathcal{M}(F_2)$, to keep only the best of them. In the field of experiment design, there are a huge number of established choices for reducing a Fisher information matrix to a scalar optimisation metric [103]. We considered three options for our optimisation process, as outlined below.

A-OPTIMISATION *Average*-optimisation is the simplest choice: optimising over the trace of the covariance matrix as a proxy for the average variance across the n parameters. This approach naïvely maximises the precision along the selected parameterisation axes. By neglecting the covariances, however, this metric does not account for the potential impact of a poor parameterisation choice. Additionally, by only considering the average variance such optimisations may prioritise one parameter over the other.

E-OPTIMISATION *Eigenvalue*-optimisation or *extreme*-optimisation maximises the information gain of the worst performing parameter by minimising $\max_i V_{i,i}$, for V the measurement covariance matrix from equation (307). Although this may not provide the maximum overall information gain, it ensures all parameters are estimated reasonably well and guards against scenarios in which one parameter is neglected in favour of the others. In most cases,[9] the states returned by this optimisation are optimal under A-optimisation, with the additional constraint that they deliver symmetric estimations.

D-OPTIMISATION *Determinant*-optimisation minimises the determinant of the covariance matrix, otherwise known as the *generalised variance*, equivalent to minimising the area of the likelihood region. This approach fixes the key problem with A-optimisation: that it is dependent on our choice of axes. Unlike the spans of a region, its area is invariant under reparameterisation or axis rotation. Consequently, unlike both other optimisation metrics, this metric fully represents the information content of the measurement. However, by minimising the size of the likelihood region with no other context, this approach is vulnerable to producing measurements that sacrifice precision in one parameter for the other, potentially delivering 'squeezed' covariance matrices.

We have opted here to use a D-optimisation metric. The tasks we will consider in this part of the thesis do not have an objective parameterisation; there is no globally established x and y axis for the gradients to be defined against. Indeed, if there were defined directions one wished to measure, one would still be able to reparameterise the coordinate system simply by rotating the measurement device. An optimisation that maximises inform-

[9] although this is in no way guaranteed by the metric.

ation gain *only* for one essentially-arbitrary coordinate parameterisation — as A- and E-optimisation do — is not a good fit here.[10]

[10]That D-optimisation may produce unbalanced outputs remains a concern here, albeit one that is mitigated through the imposed symmetry in the allowed measurement positions. To ensure that this did not impact our results, the output covariance matrices for each result were manually checked to ensure balance and each was found to be perfectly symmetric between parameters.

10

QUANTUM-ENHANCED FIELD GRADIENT ESTIMATION

The use of physical systems to estimate field gradients has previously been considered in some way by a large number of papers, usually assessing the potential for specific well-defined setups to improve the quality of measurements [120–122]. The task did not receive systematic attention from a quantum metrological point of view, though, until 2017 when Altenburg *et al.* [100] showed that under ideal conditions GHZ states were best suited to the measurement of single-dimensional field gradients. Since then, the analysis of this problem has been generalised to allow for the use of arbitrary spin-j particles [123] and similar problems have been considered for the measurement of a single-dimensional gradient in all three field components [124, 125].

In this thesis, we generalise the problem in a different way by considering a single field component that changes linearly in multiple directions. Initially, in section 10.3, we will consider the ideal case in which no noise is present and the probe network is affected only by the magnetic field. Here, we will find that the optimal approach is simply to measure each gradient individually by devoting the full entanglement network to one gradient direction, before repeating the process for other directions. A significant improvement in resolution can be found over any approach that attempts to measure both gradients simultaneously.

Although we limit our discussion to magnetic fields, these results are immediately applicable to the measurement of spatial gradients in any field that can couple to probes through the Pauli-z operator — that is to say, whenever a basis can be found in which the field affects two eigenstates in opposite ways.

Of particular interest in the analysis of quantum metrological protocols is their performance under real-world conditions with many previously-optimal approaches breaking down in the presence of noise [110, 126, 127]. We will go on to find in section 10.4 that this is true of the ideally-optimal setup in this case, which begins to underperform classical measurements in the presence of mid-strength noise fields. By making changes to the entanglement structures used in the measurement, however, the noise levels under which quantum advantage is obtainable can be extended. Of

particular interest in this section will be the conditions under which the simultaneous estimation of all parameters outperforms sequential single-parameter estimation.

10.1 STATEMENT OF TASK

Consider a magnetic field consisting only of some constant 'offset' field, B_0 and a linear gradient of the form

$$\boldsymbol{B}(\boldsymbol{r}) = B_0\hat{\boldsymbol{z}} + \boldsymbol{G}\cdot(\boldsymbol{r}-\boldsymbol{r}_0)\hat{\boldsymbol{z}}, \tag{319}$$

for $\boldsymbol{G} = (G_x, G_y, G_z)^T = (\partial_x B_z, \partial_y B_z, \partial_z B_z)^T$. We assume for simplicity that the field direction $\hat{z}$ is known and that all higher-order derivatives of the field are 0. Any x and y components to the offset field (but not the gradient component) can be neglected by assuming the z component of the offset field to be sufficiently large[1] [100].

The task we study here is the estimation of the parameters defining the field gradient, G_i, under the assumption that the offset field B_0 is known. In practice, this offset field can be ascertained simply by measuring the field at some defined origin point x_0.

To perform this task, we have at our disposal a set of n probes which we may place any at position with the 'measurement device', defined as a unit box placed such that its bottom-left corner lies at the origin. To minimise the number of measurement rounds — and hence operation time and resource usage — necessary to reach a set precision, we wish to answer the question: what spatial layout and entanglement structure will maximise information gain about the field gradients? Of particular interest will be the ability of entangled probes to beat the information obtainable from a separable set, or show quantum advantage.

The single-dimensional version of this problem — measuring a field that varies only in one direction — was studied extensively by Altenburg *et al.* [100] in 2017. They found that when the offset field is fully known, the best possible strategy is simply to measure the strength of the magnetic field at the furthest allowed point, $x_{\max}$, using a network of n probes all within a single GHZ state,

$$|\psi_{\text{opt}}\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle^{\otimes n} + |\downarrow\rangle^{\otimes n}\right). \tag{320}$$

[1] This can be achieved at will by imposing a suitably large artificial field in the desired direction, as described in Ref. [100].

That GHZ states are particularly well-suited to this task might be intuitively expected, as the magnetic field acts on a spin state proportionally to its spin value — spin up states gain phase while spin down states lose it. The best measurement of a magnetic field is therefore gained from the difference in phase between the all-up and all-down eigenstates.

To make the optimisations used in solving this problem computationally tractable, we will here consider only the 2-dimensional case and assume $G_z = 0$. It is conjectured that the improvements from multi-parameter approaches found here will be applicable to 3-dimensional estimation also.

10.2 METHODS

10.2.1 *Field interaction model*

When a spin state sits in a magnetic field, the spin-up and spin-down eigenstates interact with the field in opposing ways according to the Pauli equation, [100]

$$\hat{H} = \frac{-q\hbar}{2m}\boldsymbol{\sigma} \cdot \boldsymbol{B}, \tag{321}$$

for particle mass m and charge q, and magnetic field strength $\boldsymbol{B}$. Here, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)^T$ is a vector of the Pauli matrices and we have neglected terms such as magnetic potential that affect only global phase.

Substituting the field structure from equation (319), and working in units where $\hbar = -q = m = 1$, we can reduce the effective single-probe Hamiltonian acting on probe i, to

$$\hat{H}^{(i)} = \frac{1}{2}\big[B_0 + G_x(x_i - x_0) + G_y(y_i - y_0)\big]\sigma_z^{(i)}, \tag{322}$$

for (x_i, y_i) the probe position, and where

$$\sigma_z^{(i)} = \underbrace{\mathbb{I} \otimes \ldots \otimes \mathbb{I}}_{i-1} \overbrace{\otimes \sigma_z \otimes}^{\text{mode } i} \underbrace{\mathbb{I} \otimes \ldots \otimes \mathbb{I}}_{n-i} \tag{323}$$

represents the Pauli z operator acting only on mode i.

For an entangled network of probes in a spatially-varying field, each multi-probe eigenstate will accumulate phase depending on the state and positions of its constituent probes. The global field Hamiltonian, describing

the action on the full network, is then given by the sum of each of these local Hamiltonians as

$$\hat{H} = \frac{1}{2}B_0 \sum_i \sigma_z^{(i)} + \frac{1}{2}\sum_i \left[G_x(x_i - x_0) + G_y(y_i - y_0)\right]\sigma_z^{(i)}. \tag{324}$$

We can find the QFI using the two single-parameter generator matrices

$$\mathcal{H}_{x/y} = \mathrm{i}\left[\partial_{x/y} \exp(\mathrm{i}\,\frac{\gamma t}{2}\hat{H})\right] \exp(\mathrm{i}\,\frac{\gamma t}{2}\hat{H}) \tag{325}$$

$$= -\frac{\gamma t}{2}\sum_i (\boldsymbol{r}_i - \boldsymbol{r}_0)_{x/y}\sigma_z^{(i)}, \tag{326}$$

for field interaction time t and coupling strength γ, both of which we here set to 1.[2] Notably, the two generator matrices commute, $[\mathcal{H}_x\,;\mathcal{H}_y] = 0$ so the Cramér-Rao bound can be saturated for this problem and the two parameters can always be simultaneously estimated.

At this point, let us note again that by excluding it as an estimation parameter we have assumed perfect knowledge of B_0. As it acts on all probes in a fully known way, its impact can always be removed in the information post-processing stage and has no impact on the quality of gradient estimation. The case in which B_0 is unknown is discussed in detail for a single-dimensional estimation in Ref. [100] where it was shown that the loss of information about B_0 is equivalent to the action of global dephasing noise.

Recalling the relationship between these generators and the obtainable information from equations (309) and (311), one can immediately see that the greatest information is gained by placing probes at the extreme allowed points, $\boldsymbol{r}_i \in \{(1,0),\,(0,1),\,(1,1)\}$. Additionally assuming an equal interest in each of the gradient components — such that knowledge about one is not prioritised over the other — we have only considered setups in which

- each probe is placed at one of the extreme points of the measurement box: along an axis at $(1,0)$ or $(0,1)$ or at the extremal corner at $(1,1)$, and
- the same number of probes are placed along the x and y axes.

The two 4-probe layouts and three 6-probe layouts satisfying these conditions that we have considered in each case are shown in figure 35. We do not include the case in which all probes are placed at $(1,1)$ despite it meeting these criteria, as trivially a set of field strength measurements from a single location cannot distinguish x and y gradients.

[2] Effectively, this means we now working in temporal units of 'measurement time multiples', although this will be of no consequence as we do not consider measurement time as a factor here.

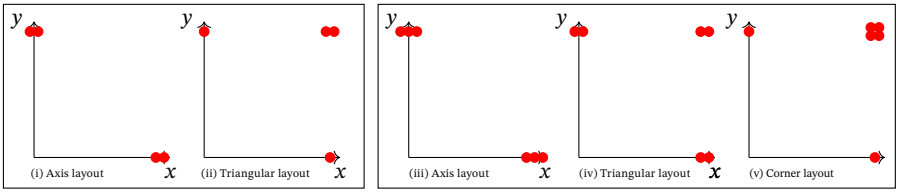


Figure 35: The probe network layouts considered for (left hand box) 4-probe measurement networks, and (right hand box) 6-probe measurement networks. In every case, the *triangular layout* in (ii) and (iv) was found to be optimal for multi-parameter estimation.

The optimal states are found for each of these layouts in turn using the process outlined in section 9.2.1, alongside the optimal sequential approach in which all n probes are allocated to the measurement of a single gradient. The overall optimal approach is then identified as the best of this set. We have found that for multi-parameter estimation, the layout in which two probes are placed in the corner (the *triangular layout*) is optimal in every case. In no case is the 4- or 6-probe *axis layout* or the 6-probe *corner layout* optimal and so the probe layout will not be discussed further.

In some cases, the Schmidt decomposition between subsystems placed at different points appeared to indicate a bi-separable state was optimal. To allow for this, every optimisation was repeated over the space of bi-separable states, with the bi-separable result kept if it was within 1% of the optimum from the full set of states.

A further optimisation was performed for each of these layouts over the space of separable states to establish the classical benchmark.

10.2.2 *Noise modelling*

Much of this chapter will be spent discussing the impact of a variety of sources of noise on our ability to measure field gradients. In each case, the noise acts probabilistically throughout the time the probes are exposed, with the overall impact depending both on the strength of the noise field and on the interaction time. As we have already decided to work in units such that $t = 1$ in section 10.2.1, these implementation details can be neglected here and we will quantify the strength of the noise fields by a single probability $p \in [0, 1]$ that either the state will be affected or unaffected after its time in the noise field. We will use this probability — the 'noise parameter' — to label the strength of the noise fields here.

As the measurement is performed in a single step after the full application of the noise field, the information obtainable from an initially-pure state evolving under the simultaneous application of the noise and mag-

netic field is equivalent to the sequential application of each. The amount of information obtainable from an initially pure probe $|\psi\rangle$ in the presence of a noisy field $\mathcal{N}$ is therefore simply given by the relevant QFI matrix for the noisy state $\hat{\rho} = \mathcal{N}(|\psi\rangle\langle\psi|)$.

This output from a noise channel can be found through the Kraus operators, $\{\hat{K}_i\}$, describing the channel, as outlined in section 2.1.4, as

$$\mathcal{N}(\hat{\rho}) = \sum_i \hat{K}_i|\psi\rangle\langle\psi|\hat{K}_i^\dagger. \tag{327}$$

10.3 NOISELESS GRADIENT DETECTION

Before we go on to discuss the primary question of this chapter (the measurement of field gradients in the presence of noise), let us briefly consider the noiseless case. We already know from Ref. [100] that the optimal approach to measure a single field gradient is to clump probes at the furthest allowed point and entangle them as a single GHZ state of the form

$$|\psi\rangle = |\mathrm{GHZ}\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle^{\otimes n} + |\downarrow\rangle^{\otimes n}\right). \tag{328}$$

In extending this approach to multiple dimensions two answers present themselves.

Within a single measurement round, the greatest information content — that is, the measurement which minimises the size of the resultant confidence region — is obtained by using a single entangled network to measure both gradients simultaneously. In the 4-probe case, this advantage is obtained through the use of the superposition state

$$|\psi\rangle \approx \sin(1.25)|\psi_1\rangle + \cos(1.25)|\psi_2\rangle, \tag{329}$$

for

$$|\psi_1\rangle = \frac{1}{2}\left(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\right) \otimes \left(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle\right), \tag{330}$$

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\uparrow\uparrow\uparrow\rangle + |\downarrow\downarrow\downarrow\downarrow\rangle\right). \tag{331}$$

The single-round information content of $\det F(\psi) = 16.2$ obtainable from this approach marginally outpaces the information obtainable from splitting the probes into two groups, each measuring only a single gradient direction at $\det F(\text{individual}) = 16$. A similar result is found when using 6 probes, where the optimal setup for single-round information gain is to place two probes along each axis and two at the corner, as shown

in figure 35 (iv). The optimal state for this setup is *any* superposition $\sin\theta|\psi_1\rangle + \cos\theta|\psi_2\rangle$ of the states

$$|\psi_1\rangle = \frac{1}{\sqrt{3}}\big(|\uparrow\uparrow\downarrow\downarrow\rangle + |\downarrow\downarrow\uparrow\uparrow\rangle\big)|\downarrow\downarrow\rangle + \frac{1}{\sqrt{3}}|\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\rangle \tag{332}$$

$$|\psi_2\rangle = \frac{1}{\sqrt{3}}\big(|\uparrow\uparrow\downarrow\downarrow\rangle + |\downarrow\downarrow\uparrow\uparrow\rangle\big)|\uparrow\uparrow\rangle + \frac{1}{\sqrt{3}}|\downarrow\downarrow\downarrow\downarrow\downarrow\downarrow\rangle, \tag{333}$$

including the extremal cases of $|\psi_1\rangle$ and $|\psi_2\rangle$ individually. Notably, this class of optimal states includes the 6-probe generalisation of the 4-mode optimal state,

$$\frac{1}{\sqrt{6}}\big(|\uparrow\uparrow\downarrow\downarrow\rangle + |\downarrow\downarrow\uparrow\uparrow\rangle\big)\big(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle\big) + \frac{1}{\sqrt{6}}\big(|\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\rangle + |\downarrow\downarrow\downarrow\downarrow\downarrow\downarrow\rangle\big), \tag{334}$$

in which the tensor product of opposing spins between the axis states with a Bell pair on the extremal corner is combined with the full GHZ state.

This naïve approach to the independent measurement of two parameters does not fully represent the potential information gain from quantum estimation, however. Rather than splitting the N probes into two $N/2$ systems to measure each gradient independently, what if one were to instead devote the entire N-probe network to a single direction, before repeating the process for the opposite direction? Such an approach would enable each single-parameter measurement to take full advantage of the N^2 information scaling allowed from quantum metrological approaches, with the overall information gained after the two measurement rounds given by the sum of the QFI matrices for each individual measurement as

$$F_{\text{total}} = \sum_i F_i. \tag{335}$$

Comparing this result with two rounds of the optimal multi-parameter case[3] — with QFI matrix given by $F_2 = 2F_1$ — therefore represents a more balanced test. In practice, it is never going to be the case that one performs a single measurement and then stops, so this change is representative of real-world uses.

A comparison of the obtainable information after two rounds using each of these approaches is shown in table 1. While the best information possible from a single measurement round is indeed given by the multi-parameter approach — albeit with a coordinate-system reparameterisation necessary

[3] Note that the optimal two-round setup for a multi-parameter measurement is simply to run the optimal single-round setup twice. By testing for the overall information after a single run, we have already optimised for the states that best *simultaneously* measure both field gradients. It is only by fully neglecting one parameter each round that the single-parameter result changes.

	4-PROBE NETWORK	6-PROBE NETWORK
CLASSICAL	$F = \begin{pmatrix} 6 & 4 \\ 4 & 6 \end{pmatrix}$	$F = \begin{pmatrix} 8 & 4 \\ 4 & 8 \end{pmatrix}$
	$V = \begin{pmatrix} 0.3 & -0.2 \\ -0.2 & 0.3 \end{pmatrix}$	$V \approx \begin{pmatrix} 0.167 & -0.0833 \\ -0.0833 & 0.167 \end{pmatrix}$
	$\sqrt{\det V} \approx 0.224$	$\sqrt{\det V} \approx 0.144$
MULTI-PARAMETER	$F = \begin{pmatrix} 10.8 & 7.2 \\ 7.2 & 10.8 \end{pmatrix}$	$F = \begin{pmatrix} 21.3 & 10.7 \\ 10.7 & 21.3 \end{pmatrix}$
	$V \approx \begin{pmatrix} 0.167 & -0.111 \\ -0.111 & 0.167 \end{pmatrix}$	$V \approx \begin{pmatrix} 0.0625 & -0.0313 \\ -0.0313 & 0.0625 \end{pmatrix}$
	$\sqrt{\det V} \approx 0.124$	$\sqrt{\det V} \approx 0.0541$
SINGLE-PARAMETER (SIMULTANEOUS)	$F = \begin{pmatrix} 8 & 0 \\ 0 & 8 \end{pmatrix}$	$F = \begin{pmatrix} 18 & 0 \\ 0 & 18 \end{pmatrix}$
	$V = \begin{pmatrix} 0.125 & 0 \\ 0 & 0.125 \end{pmatrix}$	$V \approx \begin{pmatrix} 0.0556 & 0 \\ 0 & 0.0556 \end{pmatrix}$
	$\sqrt{\det V} = 0.125$	$\sqrt{\det V} \approx 0.0556$
SINGLE-PARAMETER (SEQUENTIAL)	$F = \begin{pmatrix} 16 & 0 \\ 0 & 16 \end{pmatrix}$	$F = \begin{pmatrix} 36 & 0 \\ 0 & 36 \end{pmatrix}$
	$V = \begin{pmatrix} 0.0625 & 0 \\ 0 & 0.0625 \end{pmatrix}$	$V \approx \begin{pmatrix} 0.0278 & 0 \\ 0 & 0.0278 \end{pmatrix}$
	$\sqrt{\det V} = 0.0625$	$\sqrt{\det V} \approx 0.0278$

Table 1: Quantum Fisher information matrices and corresponding covariance matrices (assuming the Cramér-Rao bound to be saturated) after two measurement rounds. Results are for optimal setups in each case in the absence of noise. The size of the confidence region is proportionate to the square root of the determinant of the covariance matrix. Smaller $\sqrt{\det V}$ implies a better overall measurement precision, while smaller diagonal elements of V indicate a better estimation for *this particular parameterisation of the spatial axes.* Consequently, although the parallel single-parameter estimation approach produces a better estimate of the G_x and G_y gradients, this represents a worse estimation of the 2-dimensional field gradient overall as a different parameterisation exists in which the multi-parameter estimation approach provides (marginally) more information. Specific numbers relate to the somewhat convoluted unit system modelled here and should not be taken to have any physical significance beyond comparative.

— any approach that attempts to measure both gradients in one round is vastly outclassed by sequential single-parameter estimation. Indeed, the information advantage from this approach represents a halving of the size of the confidence region compared to a multi-parameter measurement, and so half the number of measurement rounds would achieve the same precision!

The huge increase in precision obtainable from sequential approaches is attributable to their ability to utilise larger entanglement networks for the measurement of a single parameter. As we will see in the following section, however, these large entanglement networks are often no longer optimal in the presence of environment noise. Recalling that aside from their ability to dedicate large networks to each parameter individually single-parameter approaches are (marginally) outperformed by multi-parameter measurements one should expect these networks to become competitive again in the presence of large amounts of background noise.

10.4 NOISY GRADIENT DETECTION

Let us now consider the real-world implementation of this task, in which measurements are performed in the presence of environmental noise. We will consider in this section three of the most common sources of noise affecting discrete-variable systems, which each impact the probe network in different ways. In each case, we assume a flat noise field with no spatial variation in strength, which acts on each probe equivalently.

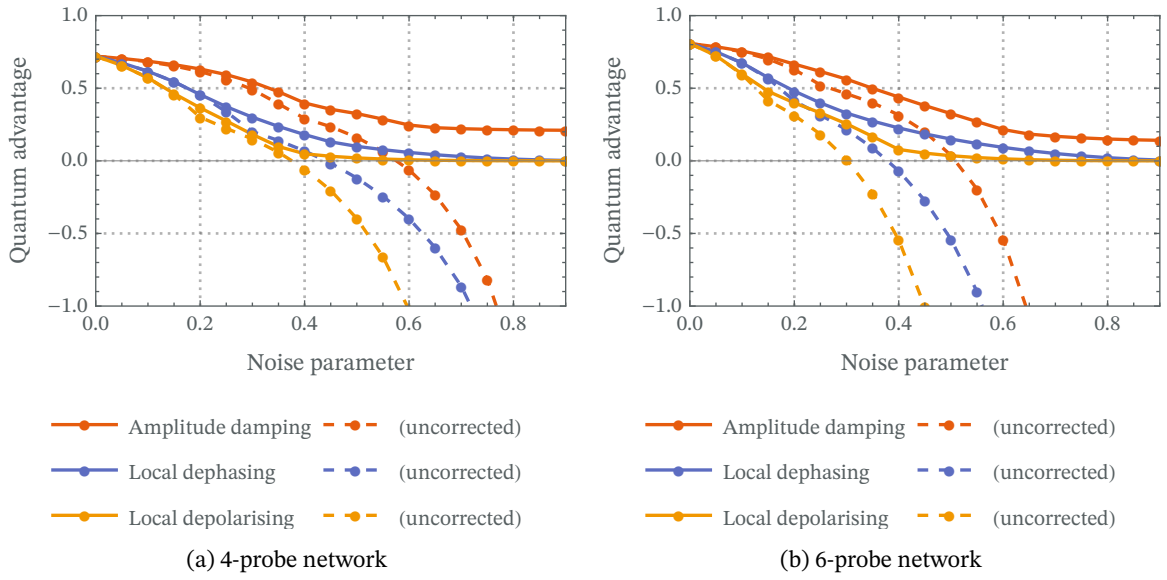
The first of these is *depolarising noise*, representing some probability p that any given probe will be fully depolarised, disconnected from the network and replaced by the maximally mixed state [27]. Such a depolarisation eliminates quantum and classical features alike, so leaves the probe unable to contribute any information to the field measurement, an outcome functionally equivalent to complete probe loss. Consequently, this type of noise additionally models the scenario in which the probes are themselves flawed with a failure rate p.

The *amplitude-damping channel* we will consider next models the decay of spin-up states into spin-down states through random energy loss to the environment; for example, through spontaneous emission of photons [27]. As the mechanism by which probes accumulate information relies on the difference in field action on spin-up and spin-down eigenstates, this form of noise will naturally reduce our ability to measure field strength.

Finally, we will consider local *dephasing noise*, otherwise known as the *phase damping channel*, representing the loss of phase information to de-

coherence effects [27]. By altering the relative phase between eigenstates, this form of noise directly destroys the information encoded within the probes so should be expected to have a significant impact on measurement efficiency. In addition to being a form of noise that may directly impact a physical implementation, the classical uncertainty in phase accumulation resulting from this channel is also representative of experimental uncertainty, such as one's inability to perfectly know the field interaction time or precise probe placement.

We find that the states previously found to be optimal in Ref. [100], discussed in section 10.3, are susceptible to each of the noise channels here. The performance of each compared to a classical approach is shown in the dashed lines in figure 36; in the presence of mid-strength noise fields, these states gain less information than a classical strategy.



(a) 4-probe network

(b) 6-probe network

Figure 36: The advantage obtainable from introducing quantum entanglement to (a) a 4-probe measurement network and (b) a 6-probe measurement network in the presence of indicated noise. Advantage is quantified through the change in the size of the confidence region surrounding the two parameters, with +1 denoting a 100% reduction in size (i.e. no remaining measurement error) and −1 denoting a 100% increase in size. Solid lines correspond to the optimal approaches in each case, while corresponding dashed lines denote the results if the ideal optimum state is not changed to compensate for the noise.

By adjusting the entanglement structures within the probe network to suit the noise field, though, quantum advantage can continue to be found at higher noise levels, shown by the solid lines in figure 36. The degree of quantum advantage obtainable decreases as the noise levels increase, reflecting the loss of useful quantum features in the presence of each of these noise channels. In fact, no significant quantum advantage is possible in the presence of high levels of dephasing and depolarising

noise even with perfect knowledge of the noise field and full control over the entanglement structures. In the presence of amplitude-damping noise, though, a quantum advantage is possible through very high levels of noise ($\approx$ 90% amplitude-damping probability), albeit in a reduced form as we are limited to the use of states found in the codomain of the channel.

Let us now consider each of these noise channels in turn.

10.4.1 *Depolarising noise*

We first consider the probabilistic *depolarising channel*. This type of noise represents a possibility that the state will be completely depolarised (replaced by the maximally mixed state with the loss of all prior classical and quantum features) with some classical likelihood $p \in [0,1]$. For a qubit-like input state, this channel implements the transformation [27]

$$\mathcal{N}_{\text{depolarising}}(\hat{\rho}) = (1-p)\hat{\rho} + p\frac{\mathbb{I}_2}{2}, \tag{336}$$

representative of the $1-p$ probability that the state will be fully unaffected by the noise channel and p probability that it will be fully depolarised. As this is a *classically* probabilistic process, for any $p \neq 0$ the output will be a generally-mixed state representative of our lack of knowledge of the outcome.

The qubit depolarising channel can alternatively be written as the Kraus operators,

$$\mathcal{K} = \left\{\sqrt{1-\frac{3}{4}p}\,\mathbb{I}, \sqrt{p}/2\,\sigma_x, \sqrt{p}/2\,\sigma_y, \sqrt{p}/2\,\sigma_z\right\}, \tag{337}$$

for $\sigma_x, \sigma_y, \sigma_z$ the Pauli matrices, with the output state given by

$$\mathcal{N}_{\text{depolarising}}(\hat{\rho}) = \sum_{\hat{K}\in\mathcal{K}} \hat{K}\hat{\rho}\hat{K}^\dagger. \tag{338}$$

We consider here only the local version of depolarising noise in which each probe may be depolarised individually. Global depolarisation, in which the full system is depolarised or not as a whole, simply represents the failure of some set proportion of measurement rounds. The impact of a global failure rate on Fisher information is well understood, reducing the FI to [101]

$$F(\mathcal{N}(\hat{\rho})) = \frac{(1-p)^2}{1-p+\frac{p}{2^{n+1}}}F(\hat{\rho}), \tag{339}$$

for n the number of probes in the system, and so we do not consider this form of noise further here.

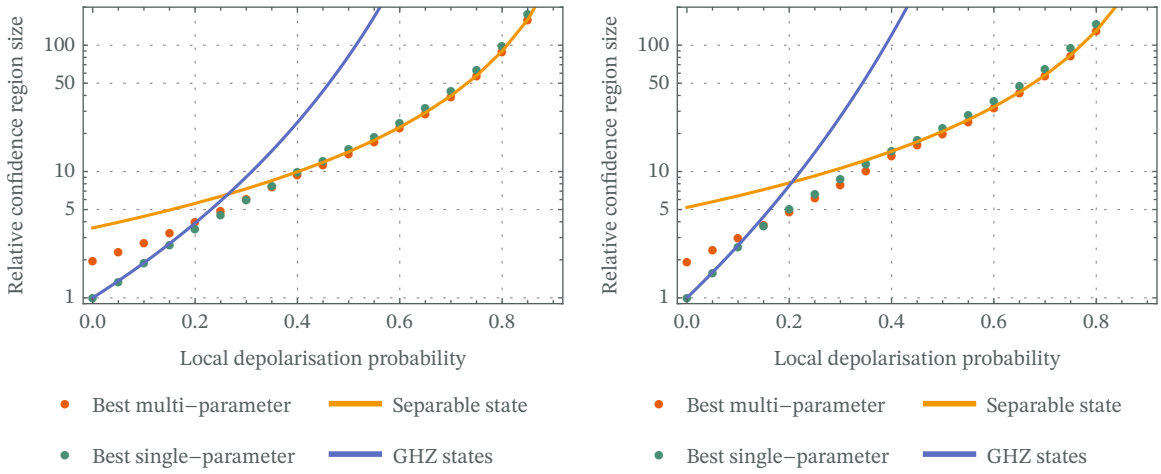
It is widely known that GHZ entanglement is uniquely vulnerable to local depolarisation as its entanglement exists only in multipartite form and cannot be split into smaller bipartite-entangled pairs [25]. The removal of a single particle from a GHZ state leaves the remaining system in a fully mixed state, eliminating its ability to gain useful information about the field. The entire system of our previously-ideal setup would therefore be knocked out by a single depolarisation event! A relatively small local loss rate p will not only be more impactful — destroying the whole system rather than only one probe — but also becomes significantly more likely, spiralling to a system loss probability of $1 - (1 - p)^n$ in a system n probes. It is clear that in the presence of depolarising fields a new strategy will be required.

10.4.1.1 *Results*

Perhaps unsurprisingly, given its tendency to completely disable probes' information collection, we have found depolarising noise to be the most destructive of the three channels we have studied. Its blanket replacement of affected probes with the fully mixed state makes this a variety of noise whose local impact cannot be eliminated simply through judicious choice of initial state. Further, as depolarisation is a highly contagious variety of noise capable of spreading through an entanglement network, our usual information-increasing tool actively works against us in this space. Designing probe networks for depolarisation-resiliance, then, represents a tradeoff between the larger entanglement structures that maximise information gain and smaller ones limiting this cascade effect.

As in every case discussed here, we have optimised over the full space of potential probe states through the mechanism outlined in section 9.2.1, to find the best possible approach at each noise level. In every case, the best multi-parameter setup was that with two probes placed in the corner position with the remaining probes distributed between the axes (the *triangular* layout shown in figure 35 (i) and (iv)).

The amount of information obtainable from the optimal entangled and classical approaches is shown in figure 37, quantified through our ability to use the gained information to narrow the confidence region surrounding the parameters. The presence of a depolarising field very quickly eliminates the previously-optimal setup's ability to collect information. Indeed, using a 4-probe network in the presence of a noise field with a local depolarisation probability of 40%, would require nearly 25 times as many measurements



(a) 4 probe (b) 6 probe

Figure 37: The size of the two-parameter confidence region after 2 measurement rounds in the presence of indicated local depolarisation noise, relative to the ideal confidence region size in the absence of depolarisation, on a logarithmic scale. (a) Results for a 4-probe measurement network; (b) results for a 6-probe measurement network. Solid lines represent the information obtainable from a fully classical measurement (orange) and from a measurement made using the GHZ-state approach, which would be the optimal approach in the absence of noise, with no adaptation for the depolarising field (blue). Red and green dots represent the obtainable information from multi-parameter and alternating single-parameter measurements optimised for the noise field. Smaller values represent a better measurement; larger values indicate an increasing number of measurement rounds would be required to achieve the same precision as possible in the absence of noise.

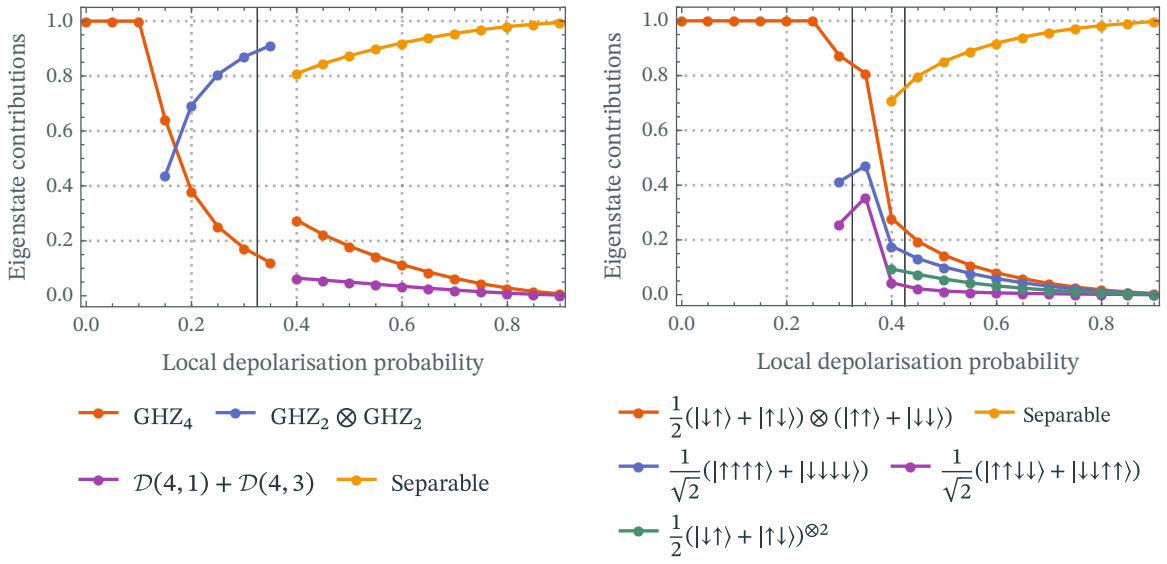
to gain the same confidence as that found in the absence of noise. In the 6-probe case this rises to 120 times as many measurements!

Although there does remain a limit on the noise in which a quantum approach remains useful — indeed in the presence of this 40% depolarising field there remains very little quantum advantage to be found at all — by adapting the states the region in which quantum advantage is possible can be stretched.

OPTIMAL 4-PROBE SETUPS For a measurement network consisting of 4 probes, quantum advantage is almost exclusively obtained through sequential single-parameter measurements. Recall that the best possible measurement in the noiseless case (represented by the blue line in figure 37) is obtained through the use of a single 4-particle GHZ state,

$$|\psi\rangle = \frac{1}{\sqrt{2}}\big(|\uparrow\uparrow\uparrow\uparrow\rangle + |\downarrow\downarrow\downarrow\downarrow\rangle\big). \tag{340}$$

When there is a low risk of depolarisation, this approach remains optimal and no change is necessary to compensate for the field; the huge information advantage obtainable from this approach outweighs the risk of total



(a) Optimal single-parameter measurement states

(b) Optimal multi-parameter measurement states

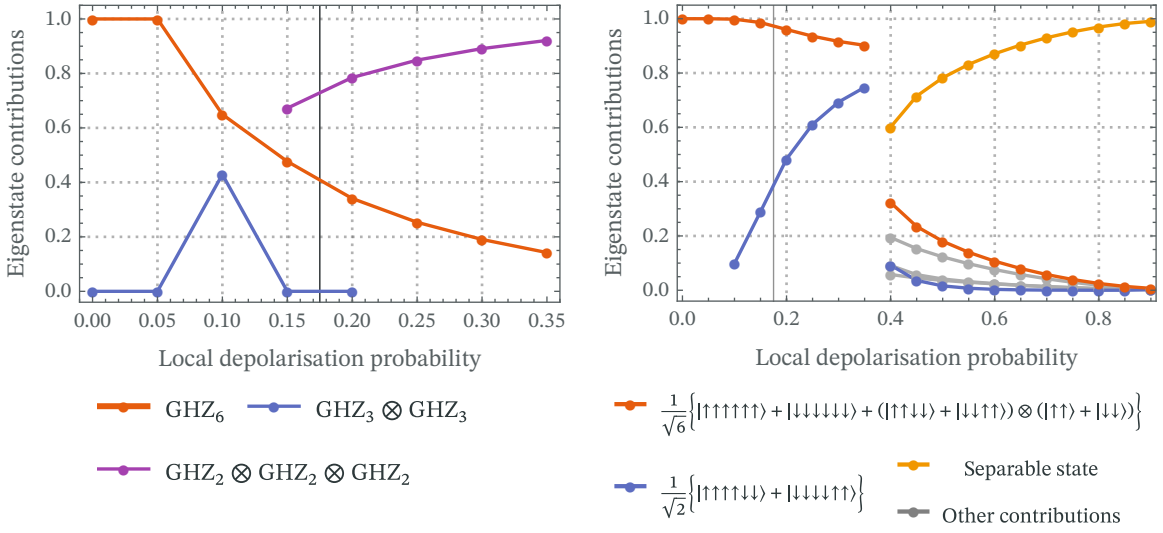
Figure 38: The optimal states for field gradient measurement using a 4-probe network to (a) measure each gradient direction sequentially, (b) measure both gradients simultaneously. The vertical line between $p = 0.30$ and $p = 0.35$ indicates the point at which the optimal approach switches from sequential single-parameter estimation to simultaneous multi-parameter estimation. $\mathcal{D}(n, k)$ represents the (n, k) Dicke state with k excitations distributed across n particles. Note that as GHZ_4 and $\mathrm{GHZ}_2^{\otimes 2}$ are not orthogonal, the values in the early part of (a) do not square and add to 1 despite the states being normalised.

loss from occasional depolarisation. As the depolarisation probability rises, though, the loss of information from failed runs becomes more significant and alterations to the state become necessary to preserve information collection.

Initially, the changes required remain light-touch, as shown on the left of figure 38a: dilute the entanglement by introducing smaller structures to the superposition. These early optimal states consist of a superposition of the (non-orthogonal) 4-mode GHZ state and 2-mode Bell states,

$$\alpha|\mathrm{GHZ}_4\rangle + \beta|\Phi_2\rangle \otimes |\Phi_2\rangle, \tag{341}$$

where $\Phi_2 = \frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle)$ is the Bell state and α, β are dependent on the noise parameter. In the event that any single probe is depolarised, the full 4-mode GHZ component to the superposition will be destroyed but one half of the pair of Bell states will survive. When no depolarisation occurs, though, the significantly larger amount of information accessible from the full GHZ state remains available, albeit in a lesser form. This approach builds redundancy into the system by sacrificing best-case information gain to ensure not all information is lost in the worst-case scenario. Eventually,



(a) Optimal single-parameter states

(b) Optimal multi-parameter states

Figure 39: The optimal states for field gradient measurement using a 6-probe network to (a) measure each gradient direction sequentially, (b) measure two directions simultaneously. The vertical line between $p = 0.15$ and $p = 0.2$ indicates the point at which a sequential single-parameter approach ceases to outperform a multi-parameter estimation. Note that the eigenstates in (a) are not orthogonal; the states presented there are normalised.

though, the probability of depolarisation increases such that even with this correction a GHZ-state approach ceases to be useful.

With a 4-probe network, the region of quantum advantage can be extended marginally by switching to a multi-parameter approach above $p = 0.30$ (this transition point is shown by the vertical line in figure 38). The optimal states for a multi-parameter setup are shown in figure 38b. Although initially (at $p = 0.35$) improvements in information collection can be found by diluting the previously-optimal state with different entanglement structures, the optimal state quickly becomes increasingly indistinguishable from the fully separable state. This similarity between the best possible quantum approach and the separable state is reflected in the extremely minor quantum advantage obtainable in the presence of high depolarising fields shown in figure 37.

OPTIMAL 6-PROBE SETUPS As the number of probes in the network increases, the usefulness of the single-parameter approach breaks down much earlier. Recalling that the probability of full-system collapse in the presence of a locally depolarising field scales with the power of n, it is clear that the advantage in information gain in these cases will more quickly be outweighed. In the case of a 6-probe network, the preservation of quantum advantage requires switching to multi-parameter estimation approaches at depolarisation strengths as low as $p = 0.2$, although ultimately quantum

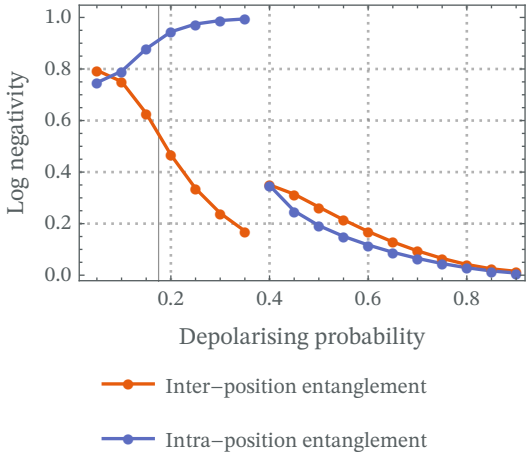


Figure 40: Entanglement structures present within the optimal multi-parameter 6-probe states, quantified by the logarithmic negativity. Red line indicates the entanglement *between* the three different positions probes can take; blue line indicates the entanglement between the two probes at each of the positions. Each of the three positions has the same intra-position entanglement properties. Maximal entanglement is indicated by a log negativity of 1.0 for the intra-position entanglement (blue) and 2.0 for inter-position entanglement (red).

advantage can be found beyond the region possible for a 4-probe network. The optimal measurement state for a 6-probe network in the single- and multi-parameter cases are shown in figure 39. Initially, these optimal setups look very similar to the 4-probe case. In the presence of very small depolarising fields (for 6-probe networks covering only the $p = 0.05$ data point) no adjustment to the previously-optimal GHZ state is needed. After this point, the average information gain can then be improved by again introducing redundancy in the form of smaller systems: initially two 3-mode GHZ states, then three 2-mode Bell states. In the event of a single probe being depolarised the entirety of the 6-mode GHZ component will be destroyed but some subsystems of the smaller-system components will survive.

These smaller systems do not gather as much information as the full 6-probe GHZ state, so quickly find their advantage over the multi-parameter approach diminished. Once the depolarisation probability reaches 20%, it becomes optimal to switch to the multi-parameter states shown in figure 39b. Initially, as expected, these look very similar to the optimal noiseless state (shown in red), mildly adapted with an increasing contribution from the $|\uparrow\rangle^{\otimes 4}|\downarrow\rangle^{\otimes 2} + |\downarrow\rangle^{\otimes 4}|\uparrow\rangle^{\otimes 2}$ state (blue) as the field strength increases. As was the case with the 4-probe networks, though, this wholly quantum approach is still only practical up to a point, after which the optimal state becomes increasingly close to the fully separable state.

Although not immediately clear from the form of the contributions, the shift from full entanglement to subsystem entanglement seen in the single-parameter probes is reflected in the multi-parameter system as well. Consider the entanglement properties (quantified by the logarithmic negativity between subsystems) across the network shown in figure 40. The change in balance between the eigenstates acts to reduce the entanglement between probe positions, while strengthening the quality of the 2-probe entanglement within each of those positions. The wider 6-probe entanglement — highly vulnerable to depolarisation — is progressively swapped out for smaller structures, introducing guards against full depolarisation.

10.4.2 *Amplitude-damping noise*

The *amplitude-damping channel* represents the noise induced by random energy transfer between a quantum state and its environment [27]. Consider a two-level energy system akin to the spin up/down states we are discussing in this chapter. At any given time, there is a small probability of the excited eigenstate decaying — through the spontaneous emission of a photon, for example — back into the ground state.[4]

This has the overall effect of reducing the probability that the state will be found in the spin-up state, represented by the transformation

$$|\downarrow\rangle\langle\downarrow| \mapsto |\downarrow\rangle\langle\downarrow|, \tag{342}$$

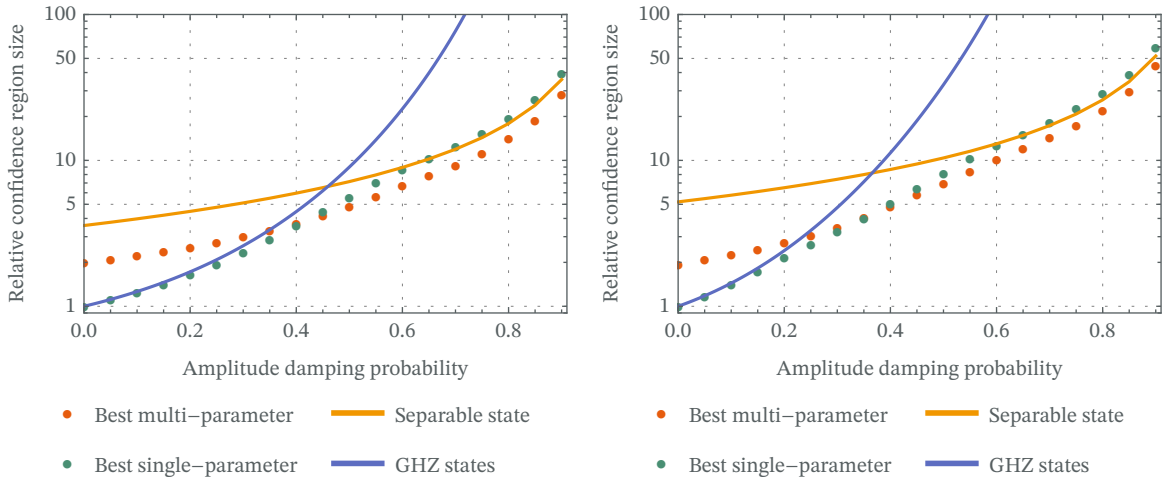
$$|\uparrow\rangle\langle\uparrow| \mapsto p|\downarrow\rangle\langle\downarrow| + (1-p)|\uparrow\rangle\langle\uparrow|, \tag{343}$$

for p the probability of phase transition. As this represents a classically probabilistic process, the density-matrix coherences also degrade, such that $|\downarrow\rangle\langle\uparrow| \mapsto \sqrt{1-p}|\downarrow\rangle\langle\uparrow|$, reducing the purity of the state. The channel can equivalently be represented through the Kraus operators [27]

$$K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}, \tag{344}$$

$$K_1 = \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix}, \tag{345}$$

[4] There is often, of course, additionally a probability that a stray photon will collide with a ground-state particle and promote it into the excited state. Such a two-way energy transfer with the environment can be modelled through the generalised amplitude-damping channel [27] — otherwise known as the thermal channel we saw in part I! To avoid undue complication, we do not consider the energy-gain effects in this part of the thesis.



(a) 4-probe network

(b) 6-probe network

Figure 41: The quantity of information obtainable from (a) a 4-probe measurement network and (b) a 6-probe measurement network in the presence of increasing levels of amplitude-damping noise, relative to the information optimally obtainable in the absence of noise. Solid lines indicate the the information obtainable from a fully-separable probe network (orange) and from the optimum noiseless setup (blue) if the latter was not adapted for the noise. Information is quantified through the size of the confidence region after two measurement rounds. Smaller values indicate a better quality measurement

again acting as

$$\mathcal{N}_{\mathrm{AD}}(\hat{\rho}) = \hat{K}_0\hat{\rho}\hat{K}_0^\dagger + \hat{K}_1\hat{\rho}\hat{K}_1^\dagger. \tag{346}$$

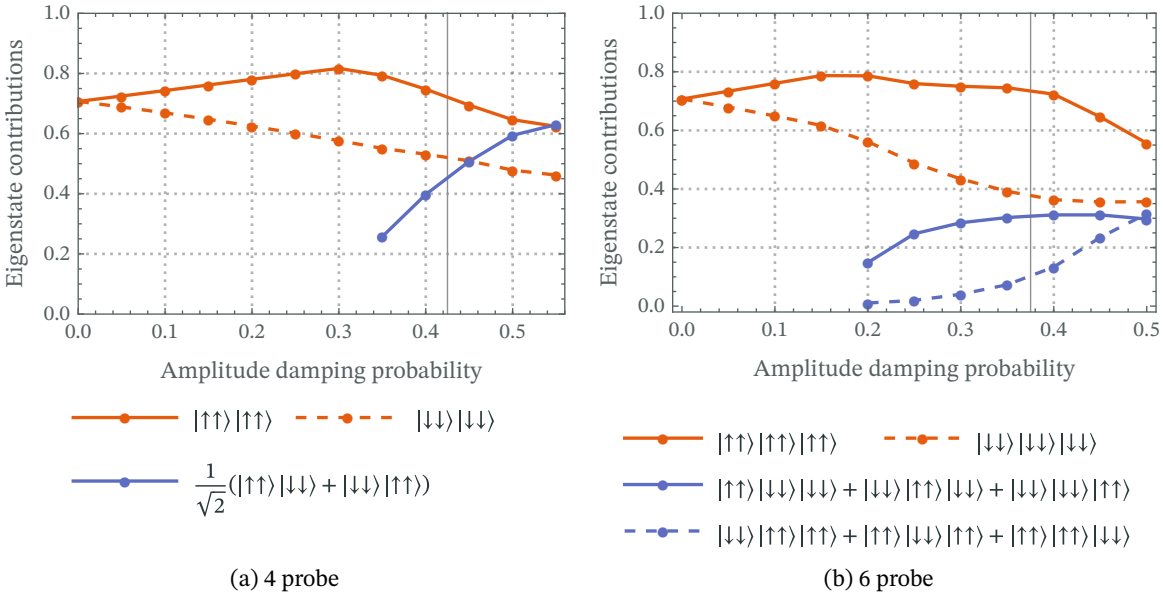
Recalling that our ability to estimate field gradients is rooted in the accumulation of phase between spin-up and spin-down eigenstates, it is anticipated that measurement performance will be significantly degraded by a channel causing random transitions from one to the other. As we will see, though, with prior knowledge of the probability of energy transfer, we are able to compensate for this process to an extent, initially simply by adjusting the balance of eigenstates in the superposition. In the limit of a fully-damping field, though, the entire measurement system will consist of spin-down states with no potential for the gain of information.

10.4.2.1 *Results*

Amplitude damping is the form of noise that gradient detection is most resilient to, with a significant quantum advantage obtainable even through 90% damping fields. As we saw when we considered depolarising fields, the ideally-optimal approach of grouping probes as a single GHZ state breaks down in the presence of noise and eventually ceases to match even the classically available information. The obtainable information from this GHZ setup is shown alongside the optimal quantum and classical approaches in

figure 41. By initially adapting the precise setup used for single-parameter estimation before switching to a multi-parameter approach at higher noise levels, a quantum advantage can be maintained even in the presence of very highly amplitude-damping fields.

OPTIMAL SINGLE-PARAMETER ESTIMATION STRATEGIES The optimal approach to gradient measurements at low levels of noise continues to be the individual measurement of each parameter. The optimal states for this task, found through numerical optimisation, are shown in figure 42. In



(a) 4 probe (b) 6 probe

Figure 42: The optimal setup for measuring a single-dimensional gradient using a measurement network of (a) 4 probes, (b) 6 probes. Vertical line between $p = 0.4$ and $p = 0.45$ in (a) and between $p = 035$ and $p = 0.4$ in (b) indicates the point at which this sequential single-parameter strategy ceases to outperform a multi-parameter approach.

the presence of very low levels of amplitude-damping noise, its effect can be directly compensated for by increasing the pre-noise likelihood of each individual probe being found in the spin-up state. The general structure of the entanglement — GHZ-like superpositions of fully spin-up and fully spin-down eigenstates — is retained, such that

$$|\psi\rangle_{\text{opt}} = \alpha|\uparrow\rangle^{\otimes n} + \sqrt{1-\alpha^2}|\downarrow\rangle^{\otimes n}, \tag{347}$$

with the eigenstate coefficients α varying dependent on the degree of noise present.

This correction is not side-effect free, however. By acting probabilistically on spin-up states only, amplitude damping reduces the purity of a state in proportion to its spin-up coefficient; this effect is shown for a 4-probe state

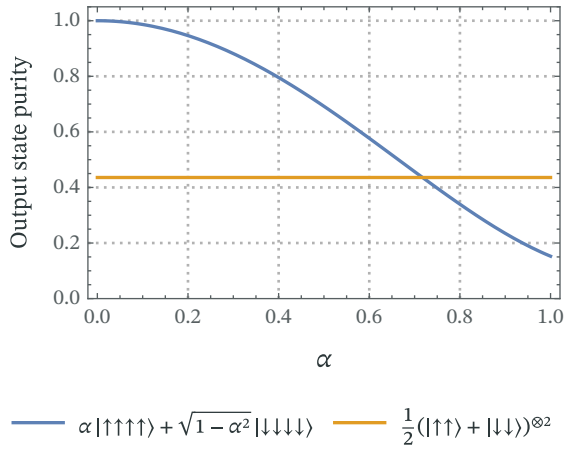


Figure 43: The purity of the denoted GHZ-like state after being passed through an amplitude-damping field with 30% amplitude-damping probability, compared to a dual-Bell-state baseline

in figure 43. By increasing the spin-up contribution to compensate for the energy loss, then, we also inadvertently reduce the purity of the network further.

At a certain point, the loss of purity from this correction becomes so great that it is no longer an effective way to compensate for the noise. At this point, we can take inspiration from the corrections made to counteract depolarising noise: introduce smaller entanglement structures that retain good information-gathering potential but that limit the wider impact of a single amplitude-damping event. The optimal state for the 4-probe system then becomes

$$|\psi_{\text{opt}}\rangle = \alpha|\uparrow\uparrow\uparrow\uparrow\rangle + \beta|\downarrow\downarrow\downarrow\downarrow\rangle + \gamma|\Phi_2\rangle^{\otimes 2}, \tag{348}$$

augmenting the previously optimal GHZ-like states with two-mode bell states of the form

$$|\Phi_2\rangle = \frac{1}{\sqrt{2}}\big(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle\big). \tag{349}$$

A similar effect can be seen in figure 42b for the 6-probe network. After the initial energy-loss-compensation approach breaks down at $p = 0.2$, the GHZ state components are joined by new terms representing smaller entanglement structures of the form

$$\big(\alpha|\uparrow\uparrow\rangle + \sqrt{1-\alpha^2}|\downarrow\downarrow\rangle\big)^{\otimes 3}, \tag{350}$$

for α dependent on p.

OPTIMAL MULTI-PARAMETER ESTIMATION STRATEGIES This performance gain from single-parameter estimation strategies stems from their ability to devote larger numbers of probes simultaneously to each dimension, and thus to establish larger entanglement networks. When the noise becomes such that these larger networks are no longer optimal — and smaller entanglement structures become necessary — this advantage begins to diminish and multi-parameter estimation protocols become competitive. For 4-probe networks the critical point at which simultaneous estimation of both gradients becomes preferable lies at single-probe damping probabilities of $p = 0.45$, while for 6-probe networks the switch becomes necessary sooner at $p = 0.4$. It is likely that as the size of the probe networks increases further, the region in which single-parameter estimation is preferable will continue to shrink and that multi-parameter approaches will become preferable at increasingly-small energy loss probabilities.

Recall from section 10.3 that when one has access to a network of 4 probes, the optimal multi-parameter approach in the absence of noise is to use a superposition of the 4-mode GHZ state $\frac{1}{\sqrt{2}}(|\uparrow\rangle^{\otimes 4} + |\downarrow\rangle^{\otimes 4})$ and the bi-separable state

$$\frac{1}{2}\underbrace{\left(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\right)}_{\text{Axes subsystem}} \otimes \underbrace{\left(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle\right)}_{\text{Corner subsystem}}. \tag{351}$$

In the presence of amplitude-damping noise the optimal GHZ state contribution quickly decays and vanishes by $p = 0.2$, a point at which the multi-parameter approach remains vastly out-classed by single-parameter estimation.

Remarkably, though, the optimal state of the first subsystem — representing the probes spread across the axes — remains constant throughout all noise levels, with the two probes always Bell-entangled as

$$|\psi_A\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle). \tag{352}$$

The second subsystem — the two probes grouped together at the furthest corner — initially retains the general form of the previously-optimal Bell-like state,

$$|\psi_B\rangle = \alpha|\uparrow\uparrow\rangle + \sqrt{1-\alpha^2}|\downarrow\downarrow\rangle, \tag{353}$$

in changing proportions α depending on the strength of the noise field. The state of this second subsystem from the point at which the optimal

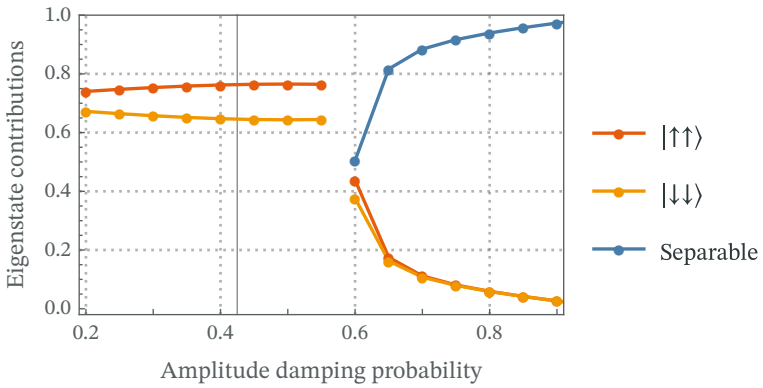


Figure 44: Optimal state of the second subsystem (placed at the extreme corner) of a 4-probe entanglement network for the simultaneous measurement of two-dimensional field gradients in the presence of amplitude-damping noise.The vertical line between $= 0.4$ and $p = 0.45$ indicates the point at which this measurement setup begins to outperform the single-parameter estimation strategy shown in figure 42a.

state becomes bi-separable at $p = 0.2$ is shown in figure 44. The region at which the Bell-like entangled states remain useful for gradient detection is shown by the red ($|\uparrow\uparrow\rangle$) and orange ($|\downarrow\downarrow\rangle$) lines on the left of the plot. Eventually, though, even these very small entanglement structures become an impediment to information gain; as can be seen in figure 44, from $p = 0.6$ the optimal state of this subsystem very quickly becomes dominated by the separable state.

The optimal state of the first subsystem remains entangled throughout all levels of amplitude-damping noise studied. Hence, a quantum advantage remains achievable even in the high-noise regions, despite the second subsystem tending towards a fully classical measurement. This is shown at the right edge of figure 41, in which the optimal quantum protocols (red dots) continue to produce smaller confidence regions than are obtainable from classical approaches.

A similar behaviour is observed for 6-probe multi-parameter gradient estimation. The optimal 6-probe layout is to place the probes in three pairs: along the two axes at positions $(1, 0)$ and $(0, 1)$ and at the extremal corner, $(1, 1)$ as show in figure 35(iv) on page 155. Initially, the optimal approach involves a fully-entangled system consisting of all 6 probes, but in the presence of increasing amplitude-damping probability the optimal setup increasingly has the corner probes disconnected from the 4 probes distributed between the systems. Selected entanglement properties for these states are shown in figure 45, quantified by the remaining logarithmic negativity between two subsystems after the third subsystem has been traced out. The optimal 6-probe entanglement structures echo those for the 4-probe

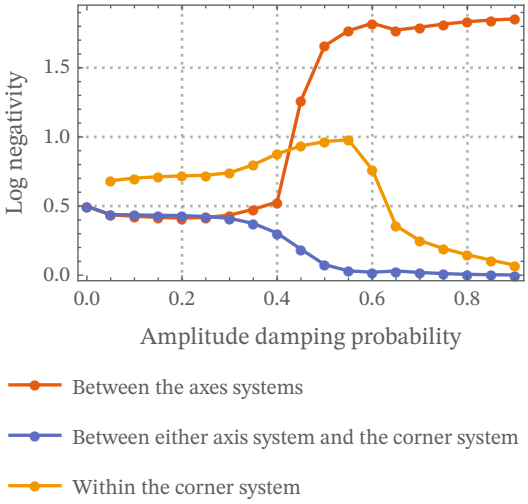


Figure 45: The entanglement structures present within the optimal 6-probe state in the presence of amplitude-damping noise, quantified by the logarithmic negativity between two subsystems. Red line indicates the entanglement between the two subsystems (of two probes each) placed along each axis. Blue line indicates the entanglement present between either of those subsystems and the two probes placed at the extremal corner. Yellow line indicates the entanglement between the two probes placed at the corner. Note that as the subsystems have different sizes, the red and blue inter-position measurement would indicate maximal entanglement at a log negativity of 2.0 while the intra-position measurement would be maximal at a log negativity of 1.0.

network: while cross-axis entanglement (red line) remains useful even in the presence of extremely high levels of noise, there quickly becomes no use for entanglement between those subsystems and the probes placed in the corner (blue line). Indeed, from $p = 0.55$ the optimal state is effectively bi-separable[5] between the two axes and the corner subsystems. Within the corner system, the entanglement between the two probes (orange line) also breaks down in a similar way to the corner probes in the 4-probe network, with the optimal setup in the presence of highly damping fields tending to the fully separable state. As was the case for the 4-probe networks, the existence of this entanglement throughout the spectrum of noise parameters is reflected by the existence of a significant achievable quantum advantage in the presence of any level of amplitude-damping.

10.4.3 *Dephasing noise*

The final form of noise we consider in this chapter is *dephasing noise*, otherwise known as the *phase damping channel*, representing the loss of

[5]The optimal bi-separable state $|\psi\rangle_{\text{axes}} \otimes |\psi\rangle_{\text{corner}}$ produces a confidence region within 1% of the size of the optimal 6-probe-entangled state.

information about relative phase between eigenstates. Traditionally, this is representative of an unknown random rotation in phase space, described by a Gaussian distribution with variance $2p$, such that the output state is described by [27]

$$\mathcal{N}(\hat{\rho}) = \frac{1}{\sqrt{4\pi p}} \int_{-\infty}^{\infty} d\theta \, \mathrm{e}^{-\theta^2/4p} \, \hat{R}(\theta)\hat{\rho}\hat{R}^{\dagger}(\theta), \tag{354}$$

for $\hat{R}(\theta)$ the rotation operator. Under the application of this type of noise, the probability with which the state will be found in any given eigenstate remains unchanged, but it will gain or lose phase in proportion to its eigenvalue. As eigenvalue-proportionate phase accumulation is the conduit through which the probes couple to the field, any loss of relative phase information will have a large impact on our ability to make accurate measurements.

Although dephasing is ordinarily thought of as a source of continuous noise — with the state always impacted to a degree dictated by the noise parameter — it has been shown to be equivalent to the probabilistic phase-flip channel, in which a qubit either remains completely unaffected or accumulates a π phase difference between the eigenstates with probability p. Consequently, dephasing noise can be described through the Kraus operators [27]

$$K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}, \tag{355}$$

$$K_1 = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{p} \end{pmatrix}, \tag{356}$$

with

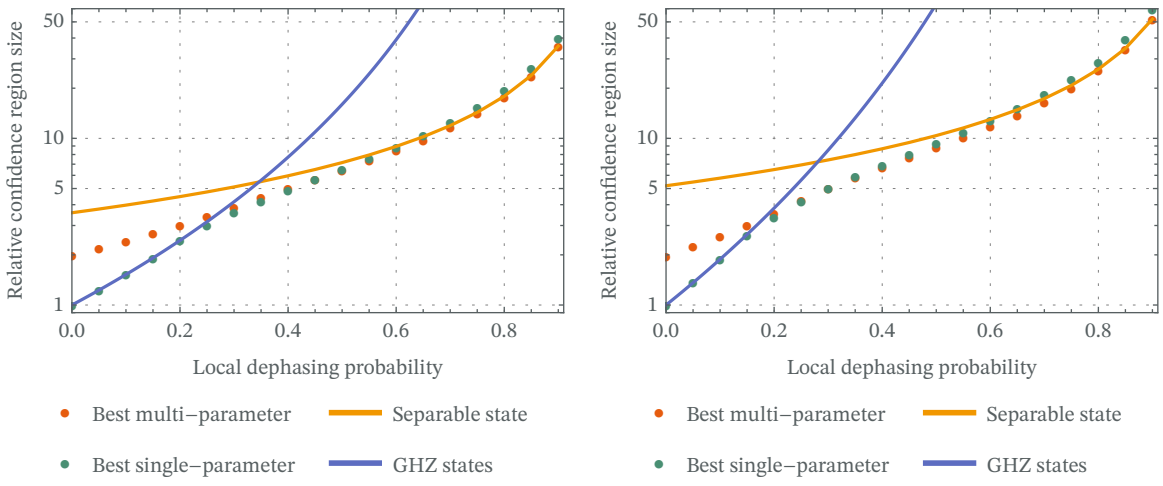
$$\mathcal{N}(\hat{\rho}) = K_0^{\dagger}\hat{\rho}K_0 + K_1^{\dagger}\hat{\rho}K_1. \tag{357}$$

In general, there are two ways for a quantum state to be dephased: locally and globally, distinguished by the degree of coordination between the random phase shifts. In both cases, the random phase shifts on each probe are drawn from the same probability distribution. In a locally-dephasing field, though, a different random phase shift is drawn from this distribution for each probe — meaning, for example, that the eigenstates $|\uparrow\rangle|\downarrow\rangle$ and $|\downarrow\rangle|\uparrow\rangle$ will not necessarily accumulate the same phase. Such random shifts might occur for example from experimental imperfections such as one's inability to decouple all probes from the field simultaneously or perfectly know their spatial position. In a globally-dephasing field, by contrast, a

single random phase shift is applied equally to all probes — for example through an unknown but spatially constant offset field B_0 — with the phase difference accumulated by each eigenstate simply proportional to the number of excited probes it represents. In this case, although an element of randomness is introduced to the eigenstate phases, it remains certain that no phase difference accumulates between eigenstates of equal total energy.

Globally dephasing noise is comparatively simple to combat. As it works equally on all configurations with the same total number of spin-up/spin-down states, there exists a class of state — the so-called decoherence-free states, to which the Dicke states belong — for which global dephasing is reduced to an irrelevant global phase accumulation. By restricting oneself to these probes then, any level of global dephasing can be tolerated. This approach to combating dephasing noise for single-dimensional gradient detection is discussed in detail in Ref. [100], and so in this chapter we will focus exclusively on local dephasing.

10.4.3.1 *Results*



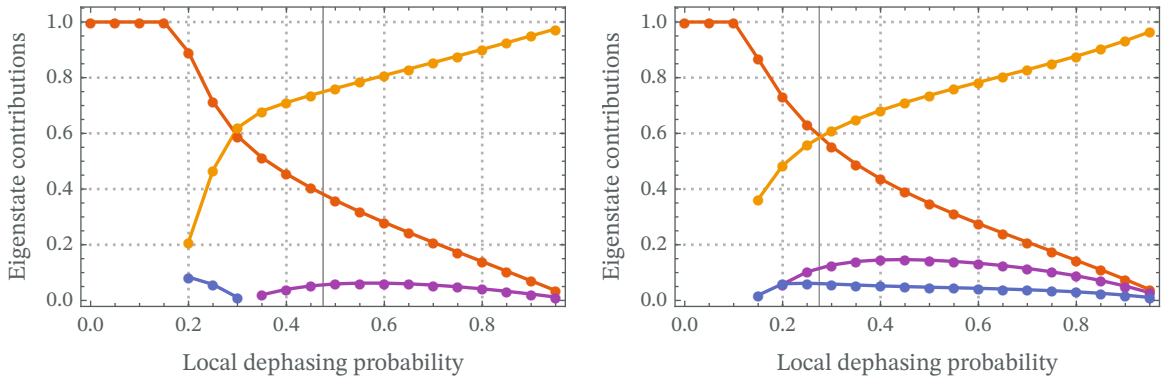
(a) 4 probe (b) 6 probe

Figure 46: The size of the two-parameter confidence region after 2 measurement rounds using (a) a 4-probe measurement network, (b) a 6-probe measurement network in the presence of a locally dephasing field, relative to that achievable in the ideal noiseless case. Solid lines represent the information obtainable from a classical measurement (orange) and a measurement using the ideal-optimum state with no adaptation (blue). Red and green dots represent the obtainable information from multi-parameter and alternating single-parameter measurements respectively that have been optimised for the noise field. Smaller values represent a better measurement.

Considering dephasing noise's direct pollution of any accumulated information, it is unsurprising to see that it is highly destructive and difficult to combat at high levels. Indeed, while quantum advantage can be exten-

ded into the mid-noise region, at high noise levels the optimal quantum approach quickly becomes indistinguishable from a classical measurement. A comparison of the quantum and classical measurements is shown in figure 46. As was the case for depolarising noise, there is very little use for entangled probes in the high-noise limit. Instead, the aim in this section is again to extend the region in which quantum advantage can be found — the area in which the green or red dots outperform the orange separable curve in figure 46 — above that for a naïve GHZ-state approach. Although these previously-optimal states quickly begin to underperform a classical measurement, we have found that the region of quantum advantage can be extended significantly by changing the states used.

OPTIMAL SINGLE-PARAMETER ESTIMATION STRATEGIES As we have found to be the case throughout this chapter, in the first instance the greatest quantum advantage is obtained through alternating single-parameter estimation. In contrast to results for other sources of noise, though, this advantage is not obtained from the introduction of alternative entanglement structures. Rather, the optimal state overwhelmingly consists of a superposition of the original GHZ state and the separable state, with only extremely minor contributions from other Dicke states,[6] as can be seen from the state contributions shown in figure 47.



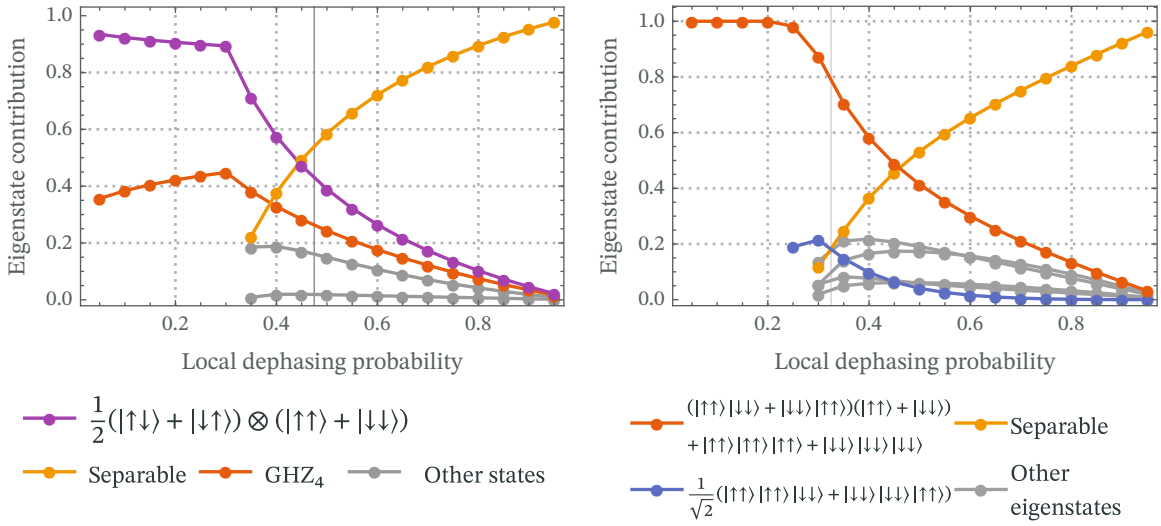
(a) 4 probe (b) 6 probe

Figure 47: The optimal states for estimating a single field gradient in the presence of dephasing noise using (a) a 4-probe network, and (b) a 6-probe network. Vertical line between $p = 0.45$ and $p = 0.5$ in (a) and $p = 0.25$ and $p = 0.3$ in (b) indicates the point at which this approach ceases to be preferable to a multi-parameter estimation strategy.

[6] We note for the avoidance of doubt that although Dicke states *individually* belong to the class of (global) decoherence-free states, combinations of Dicke states do not.

In the case of a 4-probe network, the potential for quantum advantage is almost exclusively provided by this sequential case. By the point at which a multi-parameter approach is preferable (at $p = 0.5$), the size of the confidence region would be reduced by less than 10% by using an entangled approach instead of a classical one. In the 6 probe case, though, the information obtainable from sequential single-parameter estimation breaks down much faster and the multi-parameter estimation approach becomes preferable by $p = 0.3$, where it reduces the size of the confidence region by nearly a third from the classical result.

OPTIMAL MULTI-PARAMETER ESTIMATION STRATEGIES The optimal states for a simultaneous multi-parameter estimation setup are shown in figure 48. Recalling that very little quantum advantage can be



(a) 4 probe

(b) 6 probe

Figure 48: The optimal states for simultaneously estimating two field gradients in presence of dephasing noise using (a) a 4-probe network, and (b) a 6-probe network. Vertical line between $p = 0.45$ and $p = 0.5$ in (a) and $p = 0.25$ and $p = 0.3$ in (b) indicates the point at which this multi-parameter estimation approach becomes preferable to the sequential single-parameter estimation shown in figure 47.

achieved at high noise levels, it is unsurprising to see that the optimal states become increasingly close to the fully separable case. In the case of a 4-probe network, shown in figure 48a, the optimal state is initially primarily made up of Bell pairs and GHZ states — as was optimal for the noiseless case — in changing proportions, before being increasingly dominated by the separable contribution. The optimal setup for a network of 6 probes looks similar. The primary contributions, shown in figure 48b, again come from the optimal noiseless state (red line) and the separable state (orange). However, in this case the entangled state represents the

largest contribution into larger noise levels, resulting in the more persistent quantum advantage seen in figure 46.

That quantum advantage survives longer for larger probe networks is not surprising. Recall that the QFI for quantum-networked measurements can (in the ideal case) scale with N^2, while classical ones only scale with N. Although the noise may impact entangled networks more strongly, the greater potential for quantum advantage from large networks increases their resilience compared to a classical measurement. It should be expected, then, that the region in which quantum advantage will be possible for dephasing noise will only increase as the allowed size of the measurement network increases.

11

CONCLUSION

We have considered in this part of the thesis the usefulness of quantum approaches to the measurement of field gradients, the precise estimation of which is crucial to a wide range of academic, industrial and medical tasks [97–99].

We have found that in the ideal case, in the absence of environment noise and decoherence effects, the optimal approach to the measurement of multi-dimensional field gradients is simply the sequential estimation of each of the gradients in turn. There is no benefit to taking a multi-parameter approach that attempts to use a single probe network to measure multiple gradients simultaneously. The optimal strategy is to entangle the full network as a single n-dimensional GHZ state, placed as a group at the extremal point of one of the axes, as outlined in [100], before repeating this process for the second axis. Practically, this second measurement could be trivially achieved by rotating the measurement device such that the probes now lie along the second axis, so long as that rotation is performed around the defined reference point.

In the presence of any suitably-strong noise field, though, this single-parameter approach ceases to be optimal and better-quality measurements are obtained by switching to an approach that estimates multiple gradients simultaneously. The advantage in information accrual from single-parameter estimation is drawn from their ability to devote the full entanglement network to one parameter at a time, and thus take full advantage of the scaling improvement obtainable from entangled networks. In the presence of noise, however, the advantage from these highly-entangled states either wanes due to the loss of quantum features from the noise or in some cases becomes an active hinderance to the measurement by spreading the impact of local noise beyond a single probe. Once noise levels become such that these large entanglement structures are no longer optimal, this single-parameter advantage vanishes, and the multi-parameter approach becomes necessary.

In the presence of depolarising noise, we have found that the advantage obtainable from a quantum approach very quickly decays, such that in the presence of even mid-level depolarising fields the best possible measurement is obtained from a separable network of probes. That this is the case

reflects the unique way depolarising noise interacts with entanglement networks. By essentially destroying an affected probe, tracing it out of the system, depolarisation can spread through an entanglement network. Often, learning the state of one probe in a system provides significant information about the state of the remaining networks. Tracing a probe out of the system (replacing its quantum superposition with a classical probability) can therefore collapse other parts of the superposition into a mixed state also. By minimising the size of entanglement networks, one limits the effect of any given depolarisation, which in the high-noise region vastly outweighs the potential information gain benefits from entangled approaches. The noise region in which quantum advantage is possible can be extended, though, by making small adjustments to the probe network. By augmenting the initially-optimal N-probe GHZ states with smaller entanglement structures — 2-probe Bell states, for example — one is able to retain some of the information advantage from the larger networks without losing all information in the presence of a single depolarisation event.

The quantum approach does remain robust to even very high levels of amplitude-damping noise, however. Given a suitable adaptation of the entanglement network, a quantum advantage is possible even at a damping probability of up to 90%. Initially, preserving the quantum advantage is a simple matter of adjusting the balance between spin-up and spin-down eigenstates in the prepared state to compensate for the effect of the noise field. This is only an effective approach in the relatively low-noise region, however; to maintain a quantum advantage in the presence of higher levels of amplitude-damping noise, switching to a multi-parameter approach with bespoke entanglement structures that suit the noise level becomes necessary.

A not-insignificant quantum advantage can also be maintained through the mid-noise region in the presence of a dephasing channel, additionally representative of experimental error. Although this advantage is primarily found by increasing the similarity between the state and the wholly separable state, quantum advantage can be found in the presence of higher levels of noise than would be possible for the depolarising channel. In the presence of very high dephasing noise, though, no improvement over classical measurement strategies is possible.

We should note, however, that the ability to find a quantum advantage at these higher noise levels does not necessarily represent the potential for a high-quality measurement. Often many times more measurement rounds will still be necessary to obtain a good estimate of the gradients than would be required in the absence of noise.

11.1 OUTLOOK

A number of open questions remain in this field, most significantly the large-probe-number behaviour. Due to computational constraints, we were only able to investigate networks of four and six probes. Although some trends appear to emerge — such as the optimality of multi-parameter setups emerging at lower noise levels as the probe number increases — it would be desirable to have more datapoints to confirm these. Of particular interest here would be in the distribution of extra probes between the three positions. In both 4- and 6-probe entanglement networks the optimal setup is to place two probes at the corner and distribute the remaining probes between the axes. In particular, this outperforms the case in which there is a single probe along each axis and four probes at the corner. It would be particularly interesting to learn where the additional probes are most useful for 8- and 10- probe networks — whether only two probes are ever necessary on the corner or if there is a more complex allocation algorithm to describe the optimal layout.

The emergence of quantum computers capable of efficiently optimising over large-dimensional systems may make this future study possible [128]. A particularly promising avenue is variational quantum algorithms, which combine quantum operations with classical optimisation processes [129]. The potential for such algorithms to be applied to quantum metrology processes has been studied in Refs [130, 131], where it has been shown to be an effective way to find optimal probe layouts. Once quantum computing infrastructure develops to allow for the implementation of such algorithms, this is likely to be an effective way to continue the research presented in this thesis.

As well as the trivial extension to include the third dimension — excluded here again because of computational constraints — it would be interesting to see results for the estimation of gradients in the presence of partial information about the offset field, B_0. Although the case in which the offset field is fully unknown has been studied in Ref. [100], to the author's knowledge there has as yet been no investigation of the case of partial knowledge. Such an investigation would require taking a Bayesian, rather than Fisher, approach[1] in which the parameters are allocated priors representative of the state of our knowledge of them before the estimation process.

Finally, the estimation of more complex forms of field gradient would pose an interesting task. Estimating the first and second derivatives of the

[1] for example as discussed in Ref. [101]

field simultaneously would be relatively simply to model mathematically, but the more complex form of the estimation problem would require a larger number of probes, making the optimisations more computationally challenging. It is additionally likely that such a task would break the assumption that the optimal layout is trivially one that places the probes at the extremal points, necessitating the inclusion of probe position in the optimisation as well.

APPENDICES

A

RELATIONSHIP BETWEEN THIS WORK AND MY MASTER'S THESIS

A previous iteration of this protocol was presented in a dissertation for the degree of Master of Physics in 2020 [50]. The overwhelming majority of the work presented here is new, with those results that overlap derived from a new framework.

For the avoidance of doubt, let us note any similarities here.

- The *dealer protocol* is the same for both works, but was previously selected only for its simplicity. The principles-based derivation presented in section 4.2.1 is novel.
- The prior work considered {1,3} (and by extension {2,3}) reconstruction by modelling each element in a specific experimental setup. This thesis puts the analysis on a more sound theoretical footing by constructing the quantum channel representing the reconstruction directly. The discussion of the channel presented here is entirely novel, and although some results overlap (see below) those presented here are new in that they are re-derived from this more-sound basis.
- The prior work considered the amplification and attenuation corrections discussed in section 4.2.4 to both occur *after* the reconstruction. In this work we have found improved fidelities by moving the amplification correction to before the dealer protocol. Consequently, all results presented here for $g < 1$ differ from those previously presented and are entirely novel.
- The discussion on mode-swapping in section 5.2.3 is entirely new.
- The {1,2} reconstruction protocol is unchanged, but as the pre-amplification step is new the discussion of the impact this has is novel.
- The discussion of more generalised $\{k, n\}$ QSS schemes is entirely new.
- The previous work only considered coherent states, so all work from section 5.3 onwards is entirely novel, as is the security analysis for a limited codebook presented in section 5.2.5.

Other than some parts of the form of the protocol itself, therefore, the only results produced here that have appeared in previous work are those for coherent states for $g > 1$. Even those, however, have been re-derived based on an information-theoretic approach with a quantum-channel framework that was not used in prior work.

B

GAUSSIAN INTEGRALS

In this appendix we solve a series of Gaussian integrals underpinning the theorems in chapter 6. Ultimately, in theorem B.8, we will solve the integral in equation (242) of the main text representing the potential forms of a coordinate transform of the tensor product of a Fock state and a Gaussian state given as

$$\int_{\mathbb{R}^{2N'}} d^{N'}\boldsymbol{x}_{\mathcal{I}}\, d^{N'}\boldsymbol{p}_{\mathcal{I}}\, L_n\Big([\boldsymbol{\lambda}_x\cdot\boldsymbol{x}]^2+[\boldsymbol{\lambda}_p\cdot\boldsymbol{p}]^2\Big)\exp(-\boldsymbol{q}^T V^{-1}\boldsymbol{q}+\boldsymbol{a}^T\boldsymbol{q}), \tag{B1}$$

for $\boldsymbol{q}=\boldsymbol{x}\oplus\boldsymbol{p}$, where $L_n(\cdot)$ is the Laguerre polynomial.

We will achieve this through a series of stepping stone integrals, starting in theorem B.3 with the simple case of

$$\int_{\mathbb{R}^{N'}} d^{N'}\boldsymbol{x}_{\mathcal{I}}\left(\prod_{i\in\beta} x_i\right)\exp(-\boldsymbol{x}^T V^{-1}\boldsymbol{x}+\boldsymbol{a}^T\boldsymbol{x}), \tag{B2}$$

for β any arbitrary tuple of indices of x.

We will then, in theorems B.5 and B.7 go on to solve the case in which these integrals are drawn from a power of a vector inner product as,

$$\int_{\mathbb{R}^{2N'}} d^{N'}\boldsymbol{x}_{\mathcal{I}}\,(\boldsymbol{\lambda}_x\cdot\boldsymbol{x})^n\exp(-\boldsymbol{x}^T V^{-1}\boldsymbol{x}+\boldsymbol{a}^T\boldsymbol{x}), \tag{B3}$$

and

$$\int_{\mathbb{R}^{2N'}} d^{N'}\boldsymbol{x}_{\mathcal{I}}\, d^{N'}\boldsymbol{p}_{\mathcal{I}}\,[(\boldsymbol{\lambda}_x\cdot\boldsymbol{x})^2+(\boldsymbol{\lambda}_p\cdot\boldsymbol{p})^2]^n\exp(-\boldsymbol{q}^T V^{-1}\boldsymbol{q}+\boldsymbol{a}^T\boldsymbol{q}), \tag{B4}$$

in which we are able to exploit the natural symmetry arising from the multinomial theorem to dramatically simplify the result. As the Laguerre polynomial can be written as a sum of powers of its input, this result then immediately allows us to solve our desired integral.

Finally, having achieved the general result in theorem B.8, we will present a number of special cases of particular interest in the study of Gaussian channels acting on Fock states in corollary B.11 and theorem B.14.

B.1 MATHEMATICAL TOOLS

We will in this appendix make use of some mathematical tools not used in the main body of this thesis. Let us briefly introduce those now.

We will make frequent reference in this appendix to results catalogued in the Digital Library of Mathematical Functions by the National Institute of Standards and Technology. As a large web-based reference for results, we will refer to the

results used directly, with references of the form [132, (15.2.E1)] referring to equation 15.2.E1 within Ref. [132].

MATRIX PROPERTIES Consider an invertible matrix M written in block notation

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}. \tag{B5}$$

The *Schur complement* of M with respect to subblock A is defined as

$$M/A \coloneqq D - CA^{-1}B, \tag{B6}$$

and similarly

$$M/D \coloneqq A - BD^{-1}C. \tag{B7}$$

The matrix inverse can be found using the Schur complement as [133]

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} (A - BD^{-1}C)^{-1} & -(A - BD^{-1}C)^{-1}BD^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & (D - CA^{-1}B)^{-1} \end{pmatrix} \tag{B8}$$

$$= \begin{pmatrix} (M/D)^{-1} & -(M/D)^{-1}BD^{-1} \\ -D^{-1}C(M/D)^{-1} & (M/A)^{-1} \end{pmatrix}, \tag{B9}$$

The Schur complement can also be used to prove the Sherman-Morrisson formula, that [133]

$$(M + \boldsymbol{u}^T\boldsymbol{v})^{-1} = M^{-1} - \frac{M^{-1}\boldsymbol{u}^T\boldsymbol{v}M^{-1}}{1 + \boldsymbol{v}M^{-1}\boldsymbol{u}^T}. \tag{B10}$$

COMBINATORICS We will use $x!$ to denote the factorial of x consisting of the product of x with all smaller positive integers. We will additionally make use of the rising and falling factorials of x, defined as the product of x with the following or preceding $n - 1$ integers, which we denote

$$(x)^{(n)} = \frac{(x + n - 1)!}{(x - 1)!} = x(x + 1)(x + 2)\dots(x + n - 1) \tag{B11}$$

$$(x)_{(n)} = \frac{x!}{(x - n)!} = x(x - 1)(x - 2)\dots(x - n + 1). \tag{B12}$$

Unlike the regular factorial these *do not stop at* 0, and so the falling factorial of $x > 0$ for any $n > x$ is simply 0, and likewise the rising factorial for $x < 0$. The two representations are related two each other by

$$(x)^{(n)} = (-1)^n(-x)_{(n)}, \tag{B13}$$

$$(x)_{(n)} = (-1)^n(-x)^{(n)}. \tag{B14}$$

We will represent as S_σ the set[1] of all permutations of the tuple σ. When σ is a non-repeating tuple this is also the symmetric group over σ. When σ contains multiple identical elements, S_σ as considered here treats those as distinct elements such that

$$S_{(a,a,b)} = \{(a,a,b),(a,b,a),(a,a,b),(a,b,a),(b,a,a),(b,a,a)\}. \tag{B15}$$

To maintain notational consistency, we will denote as S^k_σ the set of all k-permutations of σ: all possible ways to draw k elements from σ. For example,

$$S^2_{(a,a,b)} = \{(a,a),(a,a),(a,b),(a,b),(b,a),(b,a)\}, \tag{B16}$$

again treating repeating elements as distinguishable.

HYPERGEOMETRIC SERIES The set of hypergeometric series are an extension of the geometric series that introduces a fraction consisting of rising factorials to the summand. In this thesis, we will make use only of the ${}_2F_1$ hypergeometric series [134] [135, (15.2.E1)] — also known as the Gaussian hypergeometric series — of the form

$${}_2F_1\left(\begin{matrix}a,b\\c\end{matrix};x\right) = \sum_{n=0}^{\infty} \frac{(a)^{(n)}(b)^{(n)}}{(c)^{(n)}} \frac{x^n}{n!}. \tag{B17}$$

The series terminates only when one of the upper parameters, a or b, is a nonnegative integer, in which case the summation variable n will at some point equal a or b causing the rising factorial to evaluate to 0 for all following elements.

There exist a number of useful transforms on the hypergeometric series, which we will make use of here. These will each be introduced when they are used.

B.2 PRELIMINARY DERIVATIVE

In solving the integral in equation (B2), we shall make repeated use of the derivatives of

$$E = \exp(\frac{1}{2}\boldsymbol{a}V\boldsymbol{t} + \frac{1}{4}\boldsymbol{t}V\boldsymbol{t}), \tag{B18}$$

with respect to some series t_{β_i}. Let us first, then, consider this derivative in isolation.

[1] Although not strictly a mathematical set, as it may contain repeating elements, we will continue to use the notation of sets here as this distinction is not important for our purposes.

The first 4 derivatives in this series are given by

$$\frac{\partial E}{\partial t_{\beta_1}} = \frac{1}{2}(a_{k_1} + t_{k_1})V_{k_1,\beta_1}E \tag{B19}$$

$$\frac{\partial^2 E}{\partial t_{\beta_2}\partial t_{\beta_1}} = \left[\left(\frac{1}{2}\right)^2 (a_{k_1} + t_{k_1})V_{k_1,\beta_1}(a_{k_2} + t_{k_2})V_{k_2,\beta_2} + \frac{1}{2}V_{\beta_1,\beta_2}\right]E \tag{B20}$$

$$\frac{\partial^3 E}{\partial t_{\beta_3}\partial t_{\beta_2}\partial t_{\beta_1}} = \left[\begin{array}{l} \left(\frac{1}{2}\right)^3 (a_{k_1} + t_{k_1})V_{k_1,\beta_1}(a_{k_2} + t_{k_2})V_{k_2,\beta_2}(a_{k_3} + t_{k_3})V_{k_3,\beta_3} \\ +\left(\frac{1}{2}\right)^2 \left\{\begin{array}{l} V_{\beta_1,\beta_2}(a_{k_3} + t_{k_3})V_{k_3,\beta_3} + V_{\beta_1,\beta_3}(a_{k_2} + t_{k_2})V_{k_2,\beta_2} \\ +V_{\beta_2,\beta_3}(a_{k_3} + t_{k_3})V_{k_3,\beta_3} \end{array}\right. \end{array}\right]E \tag{B21}$$

$$\frac{\partial^4 E}{\partial t_{\beta_4}\partial t_{\beta_3}\partial t_{\beta_2}\partial t_{\beta_1}} = \left[\begin{array}{l} \left(\frac{1}{2}\right)^4 \left\{\begin{array}{l} (a_{k_1} + t_{k_1})V_{k_1,\beta_1}(a_{k_2} + t_{k_2})V_{k_2,\beta_2} \\ \times (a_{k_3} + t_{k_3})V_{k_3,\beta_3}(a_{k_4} + t_{k_4})V_{k_4,\beta_4} \end{array}\right. \\ \left(\frac{1}{2}\right)^3 \left\{\begin{array}{l} V_{\beta_1,\beta_2}(a_{k_3} + t_{k_3})V_{k_3,\beta_3}(a_{k_4} + t_{k_4})V_{k_4,\beta_4} \\ +V_{\beta_1,\beta_3}(a_{k_2} + t_{k_2})V_{k_2,\beta_2}(a_{k_4} + t_{k_4})V_{k_4,\beta_4} \\ +V_{\beta_1,\beta_4}(a_{k_2} + t_{k_2})V_{k_2,\beta_2}(a_{k_3} + t_{k_3})V_{k_3,\beta_3} \\ +V_{\beta_2,\beta_3}(a_{k_1} + t_{k_1})V_{k_1,\beta_1}(a_{k_4} + t_{k_4})V_{k_4,\beta_4} \\ +V_{\beta_2,\beta_4}(a_{k_1} + t_{k_1})V_{k_1,\beta_1}(a_{k_3} + t_{k_3})V_{k_3,\beta_3} \\ +V_{\beta_3,\beta_4}(a_{k_1} + t_{k_1})V_{k_1,\beta_1}(a_{k_2} + t_{k_2})V_{k_2,\beta_2} \end{array}\right. \\ \left(\frac{1}{2}\right)^2 \left\{V_{\beta_1,\beta_2}V_{\beta_3,\beta_4} + V_{\beta_1,\beta_3}V_{\beta_2,\beta_4} + V_{\beta_1,\beta_4}V_{\beta_2,\beta_3}\right\} \end{array}\right]E, \tag{B22}$$

where repeated k_i indices indicate implicit summation.

We can see in these first four derivatives a pattern emerging. The nth derivative consists of a weighted sum of every way to distribute n indices $\beta_1, \dots \beta_n$ between $\frac{1}{2}V_{\beta_i,\beta_j}$ and $\frac{1}{2}(a_j+t_j)V_{j,\beta_k}$. The distribution between these two terms can be described by the number of ways to split n indices between i pairs and $n-2i$ single indices for all i. We prove this observation in the following lemma, in which we symmetrise the result as $V_{a,b} = \frac{1}{2}(V_{a,b} + V_{b,a})$,

Lemma B.1. *Given*

$$E = \exp(\frac{1}{2}\boldsymbol{a}V\boldsymbol{t} + \frac{1}{4}\boldsymbol{t}V\boldsymbol{t}), \tag{B23}$$

where $\boldsymbol{a} \in \mathbb{R}^N$, $\boldsymbol{t} = (t_1, \dots, t_N)$, and $V \in \mathbb{R}^{N\times N}$ symmetric, the nth-order derivative with respect to some $\{dt_{\beta_i}\}$ is given by

$$\frac{\partial^n E}{\partial t_{\beta_n} \dots \partial t_{\beta_1}} = \left[\left(\frac{1}{2}\right)^n \sum_{m=0}^{\lfloor \frac{n}{2} \rfloor} \frac{1}{m!(n-2m)!} \sum_{\sigma\in S_\beta} \left(\prod_{i=0}^{m-1} V_{\sigma_{2i},\sigma_{2i+1}}\right)\left(\prod_{i=2m}^{n-1} (a_k + t_k)V_{k,\sigma_i}\right)\right]E, \tag{B24}$$

where S_β is the symmetric group on β (representing the set of all permutations of β), $\lfloor\cdot\rfloor$ is the floor function, and repeated k indices indicate an implicit summation over all $k \in \mathbb{Z}_N$.

Further,

$$\left.\frac{\partial^n E}{\partial t_{\beta_n}\dots\partial t_{\beta_1}}\right|_{\boldsymbol{t}\to 0} = \left(\frac{1}{2}\right)^n \sum_{m=0}^{\lfloor \frac{n}{2}\rfloor} \frac{1}{m!(n-2m)!} \sum_{\sigma\in S_\beta} \left(\prod_{i=0}^{m-1} V_{\sigma_{2i},\sigma_{2i+1}}\right)\left(\prod_{i=2m}^{n-1} (\boldsymbol{a}^T V)_{\sigma_i}\right). \tag{B25}$$

Proof. We will prove this result by induction. It can be seen from equations (B19) to (B22) that the first few derivatives of E fit this pattern. It suffices, then, to show that the formula holds for n if it holds for $n-1$ to prove it true for all n.

Before we begin, let us define for notational convenience the shorthand

$$A_i = (a_k + t_k)V_{i,k} = (a_k + t_k)V_{k,i} = \frac{1}{2}(a_k + t_k)(V_{i,k} + V_{k,i}), \tag{B26}$$

implicitly summed over k, and note that

$$\frac{\partial E}{\partial t_i} = \frac{1}{2}(a_k V_{i,k} + t_k V_{i,k})E = \frac{1}{2}A_i E, \tag{B27}$$

$$\frac{\partial A_i}{\partial t_j} = V_{i,j} = V_{j,i} = \frac{1}{2}(V_{i,j} + V_{j,i}). \tag{B28}$$

Let us now assume that equation (B24) holds upon differentiating over the tuple $\alpha = \beta\backslash\beta_n = (\beta_{n-1},\dots,\beta_1)$ consisting of the first $n-1$ elements of β, such that

$$\frac{\partial^{n-1} E}{\partial t_{\beta_{n-1}}\dots\partial t_{\beta_1}} = \left[\left(\frac{1}{2}\right)^{n-1} \sum_{m=0}^{\lfloor \frac{n-1}{2}\rfloor} \frac{1}{m!(n-2m-1)!} \sum_{\sigma\in S_\alpha} \left(\prod_{i=0}^{m-1} V_{\sigma_{2i},\sigma_{2i+1}}\right)\left(\prod_{i=2m}^{n-2} A_{\sigma_i}\right)\right] E. \tag{B29}$$

We can now find the result of the nth derivative as

$$\frac{\partial^n E}{\partial t_{\beta_n}\dots\partial t_{\beta_1}}$$

$$= \frac{\partial}{\partial t_{\beta_n}}\left[\left(\frac{1}{2}\right)^{n-1} \sum_{m=0}^{\lfloor \frac{n-1}{2}\rfloor} \frac{1}{m!(n-2m-1)!} \sum_{\sigma\in S_\alpha} \left(\prod_{i=0}^{m-1} V_{\sigma_{2i},\sigma_{2i+1}}\right)\left(\prod_{i=2m}^{n-2} A_{\sigma_i}\right)\right] E \tag{B30}$$

$$= \left[\left(\frac{1}{2}\right)^{n-1} \sum_{m=0}^{\lfloor \frac{n-1}{2}\rfloor} \frac{1}{m!(n-2m-1)!} \sum_{\sigma\in S_\alpha} \left(\prod_{i=0}^{m-1} V_{\sigma_{2i},\sigma_{2i+1}}\right)\left(\prod_{i=2m}^{n-2} A_{\sigma_i}\right)\right] \frac{\partial E}{\partial t_{\beta_n}}$$

$$+ \left[\left(\frac{1}{2}\right)^{n-1} \sum_{m=0}^{\lfloor \frac{n-2}{2}\rfloor} \frac{1}{m!(n-2m-1)!} \sum_{\sigma\in S_\alpha} \left(\prod_{i=0}^{m-1} V_{\sigma_{2i},\sigma_{2i+1}}\right)\left(\sum_{j=2m}^{n-2} \frac{\partial A_{\sigma_j}}{\partial t_{\beta_n}} \prod_{\substack{i=2m\\ i\neq j}}^{n-2} A_{\sigma_i}\right)\right] E \tag{B31}$$

$$= \left[\left(\frac{1}{2}\right)^{n-1} \sum_{m=0}^{\lfloor \frac{n-1}{2}\rfloor} \frac{1}{m!(n-2m-1)!} \sum_{\sigma\in S_\alpha} \left(\prod_{i=0}^{m-1} V_{\sigma_{2i},\sigma_{2i+1}}\right)\left(\prod_{i=2m}^{n-2} A_{\sigma_i}\right)\frac{1}{2}A_{\beta_n}\right] E$$

$$+ \left[\left(\frac{1}{2}\right)^{n-1} \sum_{m=0}^{\lfloor \frac{n-2}{2}\rfloor} \frac{1}{m!(n-2m-1)!} \sum_{\sigma\in S_\alpha} \left(\prod_{i=0}^{m-1} V_{\sigma_{2i},\sigma_{2i+1}}\right)\left(\sum_{j=2m}^{n-2} \frac{1}{2}(V_{\sigma_j,\beta_n} + V_{\beta_n,\sigma_j}) \prod_{\substack{i=2m\\ i\neq j}}^{n-2} A_{\sigma_i}\right)\right] E. \tag{B32}$$

We have lowered the upper bound of the second summation here as there are be no A_{σ_i} terms when $m = (n-1)/2$. Let us now consider the two square brackets individually.

FIRST TERM The new A_{β_n} term within the first square bracket can be absorbed into the A_σ product as

$$\left(\frac{1}{2}\right)^{n-1}\sum_{m=0}^{\lfloor\frac{n-1}{2}\rfloor}\frac{1}{m!(n-2m-1)!}\sum_{\sigma\in S_\alpha}\left(\prod_{i=0}^{m-1}V_{\sigma_{2i},\sigma_{2i+1}}\right)\left(\prod_{i=2m}^{n-2}A_{\sigma_i}\right)\frac{1}{2}A_{\beta_n} \tag{B33}$$

$$=\left(\frac{1}{2}\right)^{n}\sum_{m=0}^{\lfloor\frac{n-1}{2}\rfloor}\frac{1}{m!(n-2m-1)!}\sum_{\gamma\in\mathcal{A}'}\left(\prod_{i=0}^{m-1}V_{\gamma_{2i},\gamma_{2i+1}}\right)\left(\prod_{i=2m}^{n-1}A_{\gamma_i}\right), \tag{B34}$$

by defining the set

$$\mathcal{A}' = \Big\{(\sigma,\beta_n) \mid \sigma\in S_\alpha\Big\}, \tag{B35}$$

of all previous $\sigma\in S_\alpha$ tuples with β_n appended. However, this set still imposes that A_{β_n} appear at the end of the product. To symmetrise the product, let us instead define

$$\mathcal{A}_m = \Big\{(\pi,\tau) \mid \pi\in S_\alpha^{2m},\ \tau\in S_{\alpha\setminus\pi\cup\{\beta_n\}}\Big\}, \tag{B36}$$

the set of permutations of β in which the first $2m$ elements are drawn exclusively from α, where we have used S_X^k to denote the set of all k-permutations of X.

This set overcounts $\mathcal{A}'$ by a factor of

$$\frac{|\mathcal{A}_m|}{|\mathcal{A}'|} = \frac{|S_\alpha^{2m}||S_{\alpha\setminus\pi\cup\{\beta_n\}}|}{|S_\alpha|} = \frac{\frac{(n-1)!}{(n-1-2m)!}(n-2m)!}{(n-1)!} = \frac{(n-2m)!}{(n-2m-1)!} = n-2m, \tag{B37}$$

for the $n-2m$ positions A_{β_n} could take in the product. Due to the commutativity of multiplication we can replace $\mathcal{A}'$ by $\mathcal{A}$ simply by correcting for this overcounting as

$$\left(\frac{1}{2}\right)^{n}\sum_{m=0}^{\lfloor\frac{n-1}{2}\rfloor}\frac{1}{m!(n-2m-1)!}\frac{1}{n-2m}\sum_{\gamma\in\mathcal{A}_m}\left(\prod_{i=0}^{m-1}V_{\gamma_{2i},\gamma_{2i+1}}\right)\left(\prod_{i=2m}^{n-1}A_{\gamma_i}\right), \tag{B38}$$

$$=\left(\frac{1}{2}\right)^{n}\sum_{m=0}^{\lfloor\frac{n-1}{2}\rfloor}\frac{1}{m!(n-2m)!}\sum_{\gamma\in\mathcal{A}_m}\left(\prod_{i=0}^{m-1}V_{\gamma_{2i},\gamma_{2i+1}}\right)\left(\prod_{i=2m}^{n-1}A_{\gamma_i}\right), \tag{B39}$$

which resembles our desired result.

SECOND TERM Let us now consider the second square bracket, given by

$$\left(\frac{1}{2}\right)^{n-1}\sum_{m=0}^{\lfloor\frac{n-2}{2}\rfloor}\frac{1}{m!(n-2m-1)!}\sum_{\sigma\in S_\alpha}\left(\prod_{i=0}^{m-1}V_{\sigma_{2i},\sigma_{2i+1}}\right)\left(\sum_{j=2m}^{n-2}\frac{1}{2}(V_{\sigma_j,\beta_n}+V_{\beta_n,\sigma_j})\prod_{\substack{i=2m\\ i\neq j}}^{n-2}A_{\sigma_i}\right) \tag{B40}$$

$$=\left(\frac{1}{2}\right)^{n}\sum_{m=0}^{\lfloor\frac{n-2}{2}\rfloor}\frac{1}{m!(n-2m-1)!}\sum_{\sigma\in S_\alpha}\sum_{j=2m}^{n-2}\left(\prod_{i=0}^{m-1}V_{\sigma_{2i},\sigma_{2i+1}}\right)(V_{\sigma_j,\beta_n}+V_{\beta_n,\sigma_j})\left(\prod_{\substack{i=2m\\ i\neq j}}^{n-2}A_{\sigma_i}\right). \tag{B41}$$

We can similarly absorb the new V terms into the $V_{\sigma_{2i},\sigma_{2i+1}}$ product as

$$\left(\frac{1}{2}\right)^{n}\sum_{m=0}^{\lfloor\frac{n-2}{2}\rfloor}\frac{n-2m-1}{m!(n-2m-1)!}\sum_{\gamma\in\mathcal{B}'_m}\left(\prod_{i=0}^{m}V_{\gamma_{2i},\gamma_{2i+1}}\right)\left(\prod_{i=2(m+1)}^{n-1}A_{\gamma_i}\right), \tag{B42}$$

by noting that the σ and j summations can be equivalently written as the set

$$\begin{aligned}\mathcal{B}'_m=&\left\{(\pi,j,\beta_n,\tau)\mid \pi\in S_\alpha^{2m},\ j\in\alpha\backslash\pi,\ \tau\in S_{\alpha\backslash\pi\backslash\{k\}}\right\}\\ &\cup\left\{(\pi,\beta_n,j,\tau)\mid \pi\in S_\alpha^{2m},\ j\in\alpha\backslash\pi,\ \tau\in S_{\alpha\backslash\pi\backslash\{k\}}\right\},\end{aligned} \tag{B43}$$

which undercounts the previous expression by a factor of $n-2m-1$ for the $n-2m-1$ positions in the A product from which $\frac{\partial A_j}{\partial t_{\beta_n}}$ could have originated.

As we saw when we considered the first square bracket, though, this set still imposes that the new V term exist only at the end of the product. We can symmetrise this set to allow for the new term to take any of the $m+1$ positions within the product by defining the set of permutations of β in which the latter $n-2m-2$ indices are drawn exclusively from α as

$$\mathcal{B}_m=\left\{(\tau,\pi)\mid \pi=S_\alpha^{n-2m-2},\ \tau=S_{\alpha\backslash\pi\cup\{\beta_n\}}\right\}. \tag{B44}$$

This new set, as expected, overcounts $\mathcal{B}'_m$ by a factor of

$$\frac{|\mathcal{B}_m|}{|\mathcal{B}'_m|}=\frac{|S_\alpha^{n-2m-2}||S_{\alpha\backslash\pi\cup\{\beta_n\}}|}{2|S_\alpha^{2m}||\alpha\backslash\pi||S_{\alpha\backslash\pi\backslash\{j\}}|} \tag{B45}$$

$$=\frac{\frac{(n-1)!}{[(n-1)-(n-2m-2)]!}[(n-1)-(n-2m-2)+1]!}{2\frac{(n-1)!}{(n-1-2m)!}(n-1-2m)(n-1-2m-1)!} \tag{B46}$$

$$=m+1, \tag{B47}$$

and so replacing $\mathcal{B}'_m$ by $\mathcal{B}_m$ causes the expression to pick up a $1/(m+1)$ factor and become

$$\left(\frac{1}{2}\right)^{n}\sum_{m=0}^{\lfloor\frac{n-2}{2}\rfloor}\frac{n-2m-1}{m!(n-2m-1)!(m+1)}\sum_{\gamma\in\mathcal{B}_m}\left(\prod_{i=0}^{m}V_{\gamma_{2i},\gamma_{2i+1}}\right)\left(\prod_{i=2(m+1)}^{n-1}A_{\gamma_i}\right) \tag{B48}$$

$$=\left(\frac{1}{2}\right)^{n}\sum_{m=0}^{\lfloor\frac{n-2}{2}\rfloor}\frac{1}{(m+1)!(n-2m-2)!}\sum_{\gamma\in\mathcal{B}_m}\left(\prod_{i=0}^{m}V_{\gamma_{2i},\gamma_{2i+1}}\right)\left(\prod_{i=2(m+1)}^{n-1}A_{\gamma_i}\right). \tag{B49}$$

Reparameterising the summation index as $m \to m-1$ gives

$$\left(\frac{1}{2}\right)^n \sum_{m=1}^{\lfloor \frac{n}{2} \rfloor} \frac{1}{m!(n-2m)!} \sum_{\gamma \in \mathcal{B}_{m-1}} \left(\prod_{i=0}^{m-1} V_{\gamma_{2i},\gamma_{2i+1}}\right)\left(\prod_{i=2m}^{n-1} A_{\gamma_i}\right), \tag{B50}$$

which again resembles the final result.

RECOMBINING THE TERMS Taking together these two terms, equation (B32) can be written

$$\frac{\partial^n E}{\partial t_{\beta_n} \dots \partial t_{\beta_1}} = \left(\frac{1}{2}\right)^n \left[\begin{array}{l} \displaystyle\sum_{m=0}^{\lfloor \frac{n-1}{2} \rfloor} \frac{1}{m!(n-2m)!} \sum_{\gamma \in \mathcal{A}_m} \left(\prod_{i=0}^{m-1} V_{\gamma_{2i},\gamma_{2i+1}}\right)\left(\prod_{i=2m}^{n-1} A_{\gamma_i}\right) \\ + \displaystyle\sum_{m=1}^{\lfloor \frac{n}{2} \rfloor} \frac{1}{m!(n-2m)!} \sum_{\gamma \in \mathcal{B}_{m-1}} \left(\prod_{i=0}^{m-1} V_{\gamma_{2i},\gamma_{2i+1}}\right)\left(\prod_{i=2m}^{n-1} A_{\gamma_i}\right) \end{array} \right] E. \tag{B51}$$

Noting that $\mathcal{B}_{m-1} = \varnothing$ when $m = 0$ and $\mathcal{A}_m = \varnothing$ when $m = n/2$ (because $S^n_\alpha = \varnothing$ when α has fewer than n elements), we can equalise the limits of both summations to $m : 0 \to \lfloor \frac{n}{2} \rfloor$ and so combine the sums as

$$\left(\frac{1}{2}\right)^n \left[\sum_{m=0}^{\lfloor \frac{n}{2} \rfloor} \frac{1}{m!(n-2m)!} \sum_{\gamma \in \mathcal{A}_m \cup \mathcal{B}_{m-1}} \left(\prod_{i=0}^{m-1} V_{\gamma_{2i},\gamma_{2i+1}}\right)\left(\prod_{i=2m}^{n-1} A_{\gamma_i}\right)\right] E. \tag{B52}$$

Let us now consider the combined summation set, $\mathcal{A}_m \cup \mathcal{B}_{m-1}$. The set $\mathcal{A}_m$ consists of the set of all permutations of α with β_n inserted into one of the final $n-2m$ positions. The set $\mathcal{B}_{m-1}$ meanwhile consists of the same set of permutations of α but with β_n inserted to one of the initial $2m$ positions. The union of these two sets is identically S_β, and so we can immediately say that

$$\frac{\partial^n E}{\partial t_{\beta_n} \dots \partial t_{\beta_1}} = \left(\frac{1}{2}\right)^n \left[\sum_{m=0}^{\lfloor \frac{n}{2} \rfloor} \frac{1}{m!(n-2m)!} \sum_{\sigma \in S_\beta} \left(\prod_{i=0}^{m-1} V_{\sigma_{2i},\sigma_{2i+1}}\right)\left(\prod_{i=2m}^{n-1} (a_k + t_k) V_{k,\sigma_i}\right)\right] E, \tag{B53}$$

showing that this formula holds for n derivatives if it holds for $n-1$ and thus completing the proof.

The second result follows trivially by setting $\boldsymbol{t} \to 0$. □

B.3 FOUNDATIONAL INTEGRAL

Let us now consider the foundational integral underpinning every result in this appendix, given in equation (B2). There are two core theorems in this section, theorem B.2 in which we perform the full integration over the entire set of $\boldsymbol{x}$ variables, and theorem B.3 in which we integrate over only a subset of those variables.

Theorem B.2. *The integral*

$$\int_{\mathbb{R}^N} d^N x \left(\prod_{i\in\beta} x_i\right) \exp(-\boldsymbol{x}^T V^{-1}\boldsymbol{x} + \boldsymbol{a}^T\boldsymbol{x}), \tag{B54}$$

for $\boldsymbol{a} \in \mathbb{R}^N$, $V \in \mathbb{R}^{N\times N}$ symmetric positive semi-definite, and $\boldsymbol{\beta}$ a tuple of not necessarily unique indices of $\boldsymbol{x}$, is given by

$$\pi^{\frac{N}{2}}\sqrt{\det V}\left(\frac{1}{2}\right)^n \sum_{m=0}^{\lfloor \frac{n}{2}\rfloor} \frac{1}{m!(n-2m)!} \sum_{\sigma\in S_\beta} \left(\prod_{i=0}^{m-1} V_{\sigma_{2i},\sigma_{2i+1}}\right)\left(\prod_{i=2m}^{n-1} (\boldsymbol{a}^T V)_{\sigma_i}\right) \exp(\frac{1}{4}\boldsymbol{a}^T V \boldsymbol{a}), \tag{B55}$$

where, S_β is the symmetric group of all permutations of β, $\lfloor\cdot\rfloor$ is the floor function, and $x!$ represents the factorial of x.

Notably, in the special case where $\boldsymbol{a} = 0$ only the $m = \frac{n}{2}$ terms remain and this reduces to the previously known result given in equation (244).

Proof. This integral can be solved in the usual Gaussian fashion by first recasting it as the derivative of a known integral as

$$\int_{\mathbb{R}^N} d^N x \left(\prod_{i\in\beta} x_i\right) \exp(-\boldsymbol{x}^T V^{-1}\boldsymbol{x} + \boldsymbol{a}^T\boldsymbol{x}) \tag{B56}$$

$$= \int_{\mathbb{R}^N} d^N x \frac{\partial^n}{\partial t_{\beta_n} \dots \partial t_{\beta_1}} \exp\left[-\boldsymbol{x}^T V^{-1}\boldsymbol{x} + (\boldsymbol{a}+\boldsymbol{t})\boldsymbol{x}\right]\Big|_{\boldsymbol{t}\to 0} \tag{B57}$$

$$= \frac{\partial^n}{\partial t_{\beta_n} \dots \partial t_{\beta_1}} \int_{\mathbb{R}^N} d^N x \exp\left[-\boldsymbol{x}^T V^{-1}\boldsymbol{x} + (\boldsymbol{a}+\boldsymbol{t})\boldsymbol{x}\right]\Big|_{\boldsymbol{t}\to 0}, \tag{B58}$$

for $\boldsymbol{t} = (t_0, \dots, t_{N-1})^T$. Performing the Gaussian integration then reduces the problem to the derivative given by

$$\frac{\pi^{\frac{N}{2}}}{\sqrt{\det V^{-1}}} \frac{\partial^n}{\partial t_{\beta_n} \dots \partial t_{\beta_1}} \exp\left[\frac{1}{4}(\boldsymbol{a}+\boldsymbol{t})^T V (\boldsymbol{a}+\boldsymbol{t})\right]\Big|_{\boldsymbol{t}\to 0} \tag{B59}$$

$$= \pi^{\frac{N}{2}}\sqrt{\det V} \exp(\frac{1}{4}\boldsymbol{a}^T V \boldsymbol{a}) \frac{\partial^n}{\partial t_{\beta_n} \dots \partial t_{\beta_1}} \exp(\frac{1}{2}\boldsymbol{a} V \boldsymbol{t} + \frac{1}{4}\boldsymbol{t} V \boldsymbol{t})\Big|_{\boldsymbol{t}\to 0}, \tag{B60}$$

which we have previously solved in lemma B.1. Substituting in that result completes the proof. □

Let us now consider the case in which the above integral is performed over only some N' of the N elements of $\boldsymbol{x}$.

Theorem B.3. *The integral*

$$\int_{\mathbb{R}^{N'}} d^{N'}x_{\mathcal{I}} \left(\prod_{i\in\beta} x_i\right) \exp(-\boldsymbol{x}^T V^{-1}\boldsymbol{x} + \boldsymbol{a}^T\boldsymbol{x}), \tag{B61}$$

performed over some subset $\boldsymbol{x}_{\mathcal{I}}$ of the variables in $\boldsymbol{x}$, such that (down to an inconsequential reordering of indices)

$$\boldsymbol{x} = \boldsymbol{x}_{\mathcal{J}} \oplus \boldsymbol{x}_{\mathcal{I}}, \qquad \boldsymbol{a} = \boldsymbol{a}_{\mathcal{J}} \oplus \boldsymbol{a}_{\mathcal{I}}, \qquad V = \begin{pmatrix} V_{\mathcal{J}} & V_{\mathcal{J}\mathcal{I}} \\ V_{\mathcal{I}\mathcal{J}} & V_{\mathcal{I}} \end{pmatrix}, \tag{B62}$$

is given by

$$\begin{aligned} &\pi^{\frac{N'}{2}} \sqrt{\frac{\det V}{\det V_{\mathcal{J}}}} \exp(\frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^T B \boldsymbol{a}_{\mathcal{I}}) \left(\prod_{i\in\beta_{\mathcal{J}}} x_i\right) \exp(-\boldsymbol{x}_{\mathcal{J}}^T (V_{\mathcal{J}})^{-1}\boldsymbol{x}_{\mathcal{J}} + \theta^T \boldsymbol{x}_{\mathcal{J}}) \\ &\times \left(\frac{1}{2}\right)^{n'} \sum_{m=0}^{\lfloor \frac{n'}{2} \rfloor} \frac{1}{m!(n'-2m)!} \sum_{\sigma\in S_{\beta_{\mathcal{I}}}} \left(\prod_{i=0}^{m-1} B_{\sigma_{2i},\sigma_{2i+1}}\right) \left(\prod_{i=2m}^{n'-1} (B\boldsymbol{a}_{\mathcal{I}} + 2D\boldsymbol{x}_{\mathcal{J}})_{\sigma_i}\right), \end{aligned} \tag{B63}$$

where $\beta_{\mathcal{I}}$ and $\beta_{\mathcal{J}}$ represent the respective $\mathcal{I}/\mathcal{J}$ elements of β, $n' = |\beta_{\mathcal{I}}|$, and

$$\theta = \boldsymbol{a}_{\mathcal{J}} + D^T \boldsymbol{a}_{\mathcal{I}} \tag{B64}$$

$$B = V/V_{\mathcal{J}} = V_{\mathcal{I}} - V_{\mathcal{I}\mathcal{J}} V_{\mathcal{J}}^{-1} V_{\mathcal{J}\mathcal{I}} \tag{B65}$$

$$D = V_{\mathcal{I}\mathcal{J}} V_{\mathcal{J}}^{-1}. \tag{B66}$$

Although we have presented the 'integrated' and 'non-integrated' indices as a contiguous block in this result, this is for notational convenience only and the result applies immediately to any set $\mathcal{I}$ of indices.

Proof. First, let us note the form of the inverse matrix V^{-1} using blockwise inversion as [133]

$$\begin{pmatrix} V_{\mathcal{I}} & V_{\mathcal{I}\mathcal{J}} \\ V_{\mathcal{J}\mathcal{I}} & V_{\mathcal{J}} \end{pmatrix}^{-1} = \begin{pmatrix} (V_{\mathcal{I}} - V_{\mathcal{I}\mathcal{J}} V_{\mathcal{J}}^{-1} V_{\mathcal{J}\mathcal{I}})^{-1} & -(V_{\mathcal{I}} - V_{\mathcal{I}\mathcal{J}} V_{\mathcal{J}}^{-1} V_{\mathcal{J}\mathcal{I}})^{-1} V_{\mathcal{I}\mathcal{J}} V_{\mathcal{J}}^{-1} \\ -V_{\mathcal{J}}^{-1} V_{\mathcal{J}\mathcal{I}} (V_{\mathcal{I}} - V_{\mathcal{I}\mathcal{J}} V_{\mathcal{J}}^{-1} V_{\mathcal{J}\mathcal{I}})^{-1} & (V_{\mathcal{J}} - V_{\mathcal{J}\mathcal{I}} V_{\mathcal{I}}^{-1} V_{\mathcal{I}\mathcal{J}})^{-1} \end{pmatrix} \tag{B67}$$

$$=: \begin{pmatrix} (V/V_{\mathcal{J}})^{-1} & -(V/V_{\mathcal{J}})^{-1} D \\ -D^T (V/V_{\mathcal{J}})^{-1} & (V/V_{\mathcal{I}})^{-1} \end{pmatrix}, \tag{B68}$$

where $D := V_{\mathcal{I}\mathcal{J}} V_{\mathcal{J}}^{-1}$ and for notational convenience we additionally define $(V/V_{\mathcal{J}}) := B$ and $(V/V_{\mathcal{I}}) := A$.[2]

[2] Each of the inverses in equation (B68) are guaranteed to exist through the non-singularity of V [133].

We can use these definitions to split the integral into 'integrated' $\mathcal{I}$ ($\beta_\mathcal{I}$) and 'non-integrated' $\mathcal{J}$ ($\beta_\mathcal{J}$) components as

$$\begin{aligned}
&\int_{\mathbb{R}^{N'}} d^{N'}x_\mathcal{I} \left(\prod_{i\in\beta} x_i\right) \exp(-\boldsymbol{x}^T V^{-1}\boldsymbol{x} + \boldsymbol{a}^T\boldsymbol{x}) \\
&= \left(\prod_{i\in\beta_\mathcal{J}} x_i\right) \exp(-\boldsymbol{x}_\mathcal{J}^T A^{-1}\boldsymbol{x}_\mathcal{J} + \boldsymbol{a}_\mathcal{J}\boldsymbol{x}_\mathcal{J}) \\
&\quad \times \int_{\mathbb{R}^{N'}} d^{N'}x_\mathcal{I} \left(\prod_{i\in\beta_\mathcal{I}} x_i\right) \exp(-\boldsymbol{x}_\mathcal{I}^T B^{-1}\boldsymbol{x}_\mathcal{I} + (\boldsymbol{a}_\mathcal{I} + 2B^{-1}D\boldsymbol{x}_\mathcal{J})^T\boldsymbol{x}_\mathcal{I}),
\end{aligned} \tag{B69}$$

which is exactly the integral in theorem B.2 for $V \to B$, $\boldsymbol{a} \to \boldsymbol{a}_\mathcal{I} + 2B^{-1}D\boldsymbol{x}_\mathcal{J}$. Substituting the result from that theorem gives

$$\begin{aligned}
&\left(\prod_{i\in\beta_\mathcal{J}} x_i\right) \exp(-\boldsymbol{x}_\mathcal{J}^T A^{-1}\boldsymbol{x}_\mathcal{J} + \boldsymbol{a}_\mathcal{J}\boldsymbol{x}_\mathcal{J}) \pi^{\frac{N'}{2}} \sqrt{\det B} \left(\frac{1}{2}\right)^{n'} \\
&\times \sum_{m=0}^{\lfloor \frac{n'}{2} \rfloor} \frac{1}{m!(n'-2m)!} \sum_{\sigma\in S_{\beta_\mathcal{I}}} \left(\prod_{i=0}^{m-1} B_{\sigma_{2i},\sigma_{2i+1}}\right)\left(\prod_{i=2m}^{n'-1} (\boldsymbol{a}_\mathcal{I} + 2B^{-1}D\boldsymbol{x}_\mathcal{J})_k B_{k,\sigma_i}\right) \\
&\times \exp\left[\frac{1}{4}(\boldsymbol{a}_\mathcal{I} + 2B^{-1}D\boldsymbol{x}_\mathcal{J})^T B (\boldsymbol{a}_\mathcal{I} + 2B^{-1}D\boldsymbol{x}_\mathcal{J})\right]
\end{aligned} \tag{B70}$$

$$\begin{aligned}
&= \pi^{\frac{N'}{2}} \sqrt{\det B} \exp(\frac{1}{4}\boldsymbol{a}_\mathcal{I}^T B \boldsymbol{a}_\mathcal{I}) \left(\prod_{i\in\beta_\mathcal{J}} x_i\right) \exp\left[-\boldsymbol{x}_\mathcal{J}^T (A^{-1} - D^T B^{-1} D)\boldsymbol{x}_\mathcal{J} + (\boldsymbol{a}_\mathcal{J} + D^T\boldsymbol{a}_\mathcal{I})^T\boldsymbol{x}_\mathcal{J}\right] \\
&\times \left(\frac{1}{2}\right)^{n'} \sum_{m=0}^{\lfloor \frac{n'}{2} \rfloor} \frac{1}{m!(n'-2m)!} \sum_{\sigma\in S_{\beta_\mathcal{I}}} \left(\prod_{i=0}^{m-1} B_{\sigma_{2i},\sigma_{2i+1}}\right)\left(\prod_{i=2m}^{n'-1} (\boldsymbol{a}_\mathcal{I} + 2B^{-1}D\boldsymbol{x}_\mathcal{J})_k B_{k,\sigma_i}\right).
\end{aligned} \tag{B71}$$

Applying blockwise inversion again to equation (B68) to find $(V^{-1})^{-1} = V$, we can see that

$$(A^{-1} - D^T B^{-1} D)^{-1} = V_\mathcal{J}. \tag{B72}$$

Substituting in this result and noting that $\det B = \det V \det V_\mathcal{J}^{-1}$ [133] allows us to condense the implicit k sum into a matrix multiplication,[3]

$$\begin{aligned}
&\pi^{\frac{N'}{2}} \sqrt{\frac{\det V}{\det V_\mathcal{J}}} \exp(\frac{1}{4}\boldsymbol{a}_\mathcal{I}^T B \boldsymbol{a}_\mathcal{I}) \left(\prod_{i\in\beta_\mathcal{J}} x_i\right) \exp(-\boldsymbol{x}_\mathcal{J}^T V_\mathcal{J}^{-1}\boldsymbol{x}_\mathcal{J} + (\boldsymbol{a}_\mathcal{J} + D^T\boldsymbol{a}_\mathcal{I})^T\boldsymbol{x}_\mathcal{J}) \\
&\times \left(\frac{1}{2}\right)^{n'} \sum_{m=0}^{\lfloor \frac{n'}{2} \rfloor} \frac{1}{m!(n'-2m)!} \sum_{\sigma\in S_{\beta_\mathcal{I}}} \left(\prod_{i=0}^{m-1} B_{\sigma_{2i},\sigma_{2i+1}}\right)\left(\prod_{i=2m}^{n'-1} (B\boldsymbol{a}_\mathcal{I} + 2D\boldsymbol{x}_\mathcal{J})_{\sigma_i}\right),
\end{aligned} \tag{B73}$$

which completes the proof.

In the $\boldsymbol{x}_\mathcal{I} = \boldsymbol{x}$ case, that is to say integration over the complete $\mathbb{R}^N$ domain, this result reduces to that in theorem B.2. □

[3] Note that B is symmetric so $B_{k,\sigma_i} = B_{\sigma_i,k}$.

B.4 GAUSSIAN INTEGRALS OF POWERS OF VECTOR DOT PRODUCTS

In this section, we consider integrals of Gaussians multiplying vector products. First, let us note a brief lemma on the partitioning of $\mathbb{Z}_N^n$ tuples.

Lemma B.4. *Consider a function acting on a tuple of integers $\sigma \in \mathbb{Z}_N^n$, $f(\sigma) : \mathbb{Z}_N^n \to \mathbb{R}$ for which the order of σ is irrelevant; that is where*

$$f(\sigma) = f(\tau) \; \forall \tau \in S_\sigma. \tag{B74}$$

Then the sum over the set of all tuples can be partitioned into a $0 \ldots N'$ sum and a $N'+1 \ldots N$ sum as

$$\sum_{\sigma \in \mathbb{Z}_N^n} f(\sigma) = \sum_{n'=0}^{n} \sum_{\sigma_1 \in \mathbb{Z}_{N'}^{n'}} \sum_{\sigma_2 \in \mathbb{Z}_{N-N'}^{n-n'}} \frac{n!}{n'!(n-n')!} f(\sigma_1, \sigma_2 + N'), \tag{B75}$$

where $\sigma + N'$ denotes the tuple σ with N' added to each element.

Proof. Consider an n-tuple σ drawn from set of integers Z_N with n' elements drawn from N' and $n-1$ elements drawn from $N \setminus N'$.

The function is not sensitive to reordering within σ, so any element $\sigma \in \mathbb{Z}_N^n$ can be reordered such that all elements drawn from $\mathbb{Z}_{N'}$ precede the remaining elements.

It is trivially true that the set

$$\mathcal{A}_i = \left\{(\sigma, \tau) | \sigma \in \mathbb{Z}_{N'}^{n_i}, \, \tau \in \mathbb{Z}_{N-N'}^{n-n_i}\right\} \tag{B76}$$

contains all such-reordered tuples containing precisely n_i elements from $\mathbb{Z}_{N'}$.

There are $n!$ opportunities to draw such an n-tuple from $\mathbb{Z}_{N'}$ if one does not care about order, but only $n_i!(n-n_i)!$ ways to draw it from $\mathcal{A}_i$, hence a cardinality correction of

$$\frac{n!}{n_i!(n-n_i)!} \tag{B77}$$

is required.

The summation over all such $\mathcal{A}_i$ sets with this correction is then equivalent to a summation over $\mathbb{Z}_N^n$. □

B.4.1 *Gaussian integrals of* $(\lambda \cdot x)^n$

Theorem B.5. *The integral*

$$\int_{\mathbb{R}^{N'}} d^{N'}x_{\mathcal{I}} (\boldsymbol{\lambda}\cdot\boldsymbol{x})^n \exp(-\boldsymbol{x}^T V^{-1}\boldsymbol{x} + \boldsymbol{a}^T\boldsymbol{x}), \tag{B78}$$

performed over only the last N' members of $\boldsymbol{x}$ such that the relevant vectors and matrices split into 'integrated' $\mathcal{I}$ and 'non-integrated' $\mathcal{J}$ parts as

$$\begin{aligned} \boldsymbol{x} &= \boldsymbol{x}_{\mathcal{J}} \oplus \boldsymbol{x}_{\mathcal{I}}, & \boldsymbol{a} &= \boldsymbol{a}_{\mathcal{J}} \oplus \boldsymbol{a}_{\mathcal{I}}, \\ V &= \begin{pmatrix} V_{\mathcal{J}} & V_{\mathcal{J}\mathcal{I}} \\ V_{\mathcal{I}\mathcal{J}} & V_{\mathcal{I}} \end{pmatrix}, & \boldsymbol{\lambda} &= \boldsymbol{\lambda}_{\mathcal{J}} \oplus \boldsymbol{\lambda}_{\mathcal{I}}, \end{aligned}$$

is given by

$$\begin{aligned} &\pi^{\frac{N'}{2}} \sqrt{\frac{\det V}{\det V_{\mathcal{J}}}} \Big(-\frac{\mathrm{i}\sqrt{\boldsymbol{\lambda}_{\mathcal{I}}^T B \boldsymbol{\lambda}_{\mathcal{I}}}}{2}\Big)^n H_n\Big(\mathrm{i}\,\frac{(\boldsymbol{\lambda}_{\mathcal{J}} + D\boldsymbol{\lambda}_{\mathcal{I}})^T \boldsymbol{x}_{\mathcal{J}} + \frac{1}{2}\boldsymbol{\lambda}_{\mathcal{I}}^T B^T \boldsymbol{a}_{\mathcal{I}}}{\sqrt{\boldsymbol{\lambda}_{\mathcal{I}}^T B \boldsymbol{\lambda}_{\mathcal{I}}}}\Big) \\ &\times \exp(-\boldsymbol{x}_{\mathcal{J}}^T V_{\mathcal{J}}^{-1}\boldsymbol{x}_{\mathcal{J}} + \theta^T \boldsymbol{x}_{\mathcal{J}} + \frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^T B \boldsymbol{a}_{\mathcal{I}}), \end{aligned} \tag{B79}$$

where H_n is the nth Hermite polynomial,

$$\begin{aligned} \theta &= \boldsymbol{a}_{\mathcal{J}} + D\boldsymbol{a}_{\mathcal{I}} \\ B &= V/V_{\mathcal{I}} = V_{\mathcal{I}} - V_{\mathcal{I}\mathcal{J}} V_{\mathcal{J}}^{-1} V_{\mathcal{J}\mathcal{I}} \\ D &= V_{\mathcal{J}}^{-1} V_{\mathcal{J}\mathcal{I}} = V_{\mathcal{J}}^{-1} V_{\mathcal{J}\mathcal{I}}, \end{aligned}$$

and $V_{\mathcal{X}}^{-1} = (V_{\mathcal{X}})^{-1}$ is the inverse of the $\mathcal{X}$ block of V.
We have presented this theorem in block form for notational convenience. It equivalently applies to a set of integration and non-integration indices which do not form a contiguous block with appropriately defined $\mathcal{I}$, $\mathcal{J}$ vectors and matrices.

Proof. We can derive this result by first expanding $(\boldsymbol{\lambda}\cdot\boldsymbol{x})^n$ to get the sum of integrals,

$$\int_{\mathbb{R}^{N'}} d^{N'}x_{\mathcal{I}} (\boldsymbol{\lambda}\cdot\boldsymbol{x})^n \exp(-\boldsymbol{x}^T V^{-1}\boldsymbol{x} + \boldsymbol{a}^T\boldsymbol{x}) \tag{B80}$$

$$= \int_{\mathbb{R}^{N'}} d^{N'}x_{\mathcal{I}} \sum_{\sigma\in\mathbb{Z}_N^n} \Big(\prod_{i=0}^{n-1} \lambda_{\sigma_i} x_{\sigma_i}\Big) \exp(-\boldsymbol{x}^T V^{-1}\boldsymbol{x} + \boldsymbol{a}^T\boldsymbol{x}) \tag{B81}$$

$$= \sum_{\sigma\in\mathbb{Z}_N^n} \Big(\prod_{i=0}^{n-1} \lambda_{\sigma_i}\Big) \int_{\mathbb{R}^{N'}} d^{N'}x_{\mathcal{I}} \Big(\prod_{i=0}^{n-1} x_{\sigma_i}\Big) \exp(-\boldsymbol{x}^T V^{-1}\boldsymbol{x} + \boldsymbol{a}^T\boldsymbol{x}). \tag{B82}$$

We can then use lemma B.4 to partition σ between integrated (σ) and non-integrated (τ) indices as

$$\sum_{n'=0}^{n} \sum_{\substack{\sigma\in\mathbb{Z}^{n'}_{N'} \\ \tau\in\mathbb{Z}^{n-n'}_{N-N'}}} \frac{n!}{n'!(n-n')!}\Big[\prod_{i\in\sigma}(\lambda_{\mathcal{I}})_i \prod_{i\in\tau}(\lambda_{\mathcal{J}})_i(x_{\mathcal{J}})_i\Big] \times \int_{\mathbb{R}^{N'}} d^{N'}x_{\mathcal{I}}\Big[\prod_{i\in\sigma}(x_{\mathcal{I}})_i\Big]\exp(-\boldsymbol{x}^T V^{-1}\boldsymbol{x} + \boldsymbol{a}^T\boldsymbol{x}), \tag{B83}$$

where N' denotes the number of integration indices, and evaluate the integrals using theorem B.3 as

$$\begin{aligned}
&\pi^{\frac{N'}{2}}\sqrt{\frac{\det V}{\det V_{\mathcal{J}}}}\exp(\frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^T B\boldsymbol{a}_{\mathcal{I}})\exp(-\boldsymbol{x}_{\mathcal{J}}^T V_{\mathcal{J}}^{-1}\boldsymbol{x}_{\mathcal{J}} + \theta^T\boldsymbol{x}_{\mathcal{J}}) \\
&\times\sum_{n'=0}^{n}(\frac{1}{2})^{n'}\frac{n!}{n'!(n-n')!}\left[\sum_{\tau\in\mathbb{Z}^{n-n'}_{N-N'}}\Big(\prod_{i\in\tau}(\lambda_{\mathcal{J}})_i(x_{\mathcal{J}})_i\Big)\right] \\
&\times\sum_{m=0}^{\lfloor\frac{n'}{2}\rfloor}\frac{1}{m!(n'-2m)!}\left\{\sum_{\sigma\in\mathbb{Z}^{n'}_{N'}}\sum_{\pi\in S_\sigma}\Big[\prod_{i\in\sigma}(\lambda_{\mathcal{I}})_i\Big]\Big[\prod_{i=0}^{m-1}B_{\pi_{2i},\pi_{2i+1}}\Big]\Big[\prod_{i=2m}^{n'-1}(B^T\boldsymbol{a}_{\mathcal{I}})_{\pi_i} + 2(D^T\boldsymbol{x}_{\mathcal{J}})_{\pi_i}\Big]\right\}.
\end{aligned} \tag{B84}$$

The commutativity of multiplication allows us to replace the product over σ with a product over one of its permutations $\pi\in S_\sigma$, removing the σ dependence from the summand. We can then note that for a summation of the form

$$\sum_{\sigma\in\mathbb{Z}^{n'}_{N'}}\sum_{\pi\in S_\sigma} f(\pi), \tag{B85}$$

with no dependence on σ, summing over the symmetric group of elements of a set which already contains every permutation precisely once, as $\mathbb{Z}^n_N$ does, is superfluous and can be replaced by a cardinality multiplication, $|S_\sigma| = n'!$. We can therefore combine the summations into

$$\begin{aligned}
&\pi^{\frac{N'}{2}}\sqrt{\frac{\det V}{\det V_{\mathcal{J}}}}\exp(\frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^T B\boldsymbol{a}_{\mathcal{I}})\exp(-\boldsymbol{x}_{\mathcal{J}}^T V_{\mathcal{J}}^{-1}\boldsymbol{x}_{\mathcal{J}} + \theta^T\boldsymbol{x}_{\mathcal{J}}) \\
&\times\sum_{n'=0}^{n}(\frac{1}{2})^{n'}\frac{n!}{n'!(n-n')!}\left[\sum_{\tau\in\mathbb{Z}^{n-n'}_{N-N'}}\Big(\prod_{i\in\tau}(\lambda_{\mathcal{J}})_i(x_{\mathcal{J}})_i\Big)\right] \\
&\times\sum_{m=0}^{\lfloor\frac{n'}{2}\rfloor}\frac{n'!}{m!(n'-2m)!}\left\{\sum_{\sigma\in\mathbb{Z}^{n'}_{N'}}\Big[\prod_{i=0}^{m-1}(\lambda_{\mathcal{I}})_{\sigma_{2i}}(\lambda_{\mathcal{I}})_{\sigma_{2i+1}}B_{\sigma_{2i},\sigma_{2i+1}}\Big]\Big[\prod_{i=2m}^{n'-1}(\lambda_{\mathcal{I}})_{\sigma_i}(B^T\boldsymbol{a}_{\mathcal{I}} + 2D^T\boldsymbol{x}_{\mathcal{J}})_{\sigma_i}\Big]\right\}.
\end{aligned} \tag{B86}$$

By partitioning $\mathbb{Z}_N^n$ between the two products as

$$\begin{aligned}&\sum_{\sigma\in\mathbb{Z}_{N'}^{n'}}\Big[\prod_{i=0}^{m-1}(\lambda_{\mathcal{I}})_{\sigma_{2i}}(\lambda_{\mathcal{I}})_{\sigma_{2i+1}}B_{\sigma_{2i},\sigma_{2i+1}}\Big]\Big[\prod_{i=2m}^{n'-1}(\lambda_{\mathcal{I}})_{\sigma_i}(B^T\boldsymbol{a}_{\mathcal{I}}+2D^T\boldsymbol{x}_{\mathcal{J}})_{\sigma_i}\Big]\\ =&\Big[\sum_{\sigma\in\mathbb{Z}_{N'}^{2m}}\prod_{i=0}^{m-1}(\lambda_{\mathcal{I}})_{\sigma_{2i}}(\lambda_{\mathcal{I}})_{\sigma_{2i+1}}B_{\sigma_{2i},\sigma_{2i+1}}\Big]\Big[\sum_{\pi\in\mathbb{Z}_{N'}^{n'-2m}}\prod_{i\in\pi}(\lambda_{\mathcal{I}})_i(B^T\boldsymbol{a}_{\mathcal{I}}+2D^T\boldsymbol{x}_{\mathcal{J}})_i\Big],\end{aligned}\tag{B87}$$

and noting that

$$\sum_{\pi\in\mathbb{Z}_N^n}\prod_{i\in\pi}X_i=\Big(\sum_{i=0}^{N}X_i\Big)^n\tag{B88}$$

$$\sum_{\pi\in\mathbb{Z}_N^{2n}}\prod_{i=0}^{n-1}f(\pi_{2i},\pi_{2i+1})=\Big(\sum_{i,j=0}^{N}f(i,j)\Big)^n,\tag{B89}$$

we can factorise this result as

$$\begin{aligned}&\pi^{\frac{N'}{2}}\sqrt{\frac{\det V}{\det V_{\mathcal{J}}}}\exp(\frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^TB\boldsymbol{a}_{\mathcal{I}})\exp(-\boldsymbol{x}_{\mathcal{J}}^TV_{\mathcal{J}}^{-1}\boldsymbol{x}_{\mathcal{J}}+\theta^T\boldsymbol{x}_{\mathcal{J}})\\ &\times\sum_{n'=0}^{n}\sum_{m=0}^{\lfloor\frac{n'}{2}\rfloor}(\frac{1}{2})^{n'}\frac{n!}{m!\,(n-n')!\,(n'-2m)!}\\ &\times\Big[\sum_{i=0}^{N-N'}(\lambda_{\mathcal{J}})_i(x_{\mathcal{J}})_i\Big]^{n-n'}\Big[\sum_{i,j=0}^{N'}(\lambda_{\mathcal{I}})_i(\lambda_{\mathcal{I}})_jB_{i,j}\Big]^m\Big[\sum_{i=0}^{N'}(\lambda_{\mathcal{I}})_i(B^T\boldsymbol{a}_{\mathcal{I}}+2D^T\boldsymbol{x}_{\mathcal{J}})_i\Big]^{n'-2m}\end{aligned}\tag{B90}$$

and identify the summations as vector/matrix products, to get

$$\begin{aligned}&\pi^{\frac{N'}{2}}\sqrt{\frac{\det V}{\det V_{\mathcal{J}}}}\exp(\frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^TB\boldsymbol{a}_{\mathcal{I}})\exp(-\boldsymbol{x}_{\mathcal{J}}^TV_{\mathcal{J}}^{-1}\boldsymbol{x}_{\mathcal{J}}+\theta^T\boldsymbol{x}_{\mathcal{J}})\\ &\times\sum_{n'=0}^{n}\sum_{m=0}^{\lfloor\frac{n'}{2}\rfloor}(\frac{1}{2})^{n'}\frac{n!}{m!\,(n-n')!\,(n'-2m)!}\Big[\boldsymbol{\lambda}_{\mathcal{J}}^T\boldsymbol{x}_{\mathcal{J}}\Big]^{n-n'}\Big[\boldsymbol{\lambda}_{\mathcal{I}}^TB\boldsymbol{\lambda}_{\mathcal{I}}\Big]^m\Big[\boldsymbol{\lambda}_{\mathcal{I}}^TB^T\boldsymbol{a}_{\mathcal{I}}+2\boldsymbol{\lambda}_{\mathcal{I}}^TD^T\boldsymbol{x}_{\mathcal{J}}\Big]^{n'-2m}.\end{aligned}\tag{B91}$$

Let us now swap the order of the summation as

$$\sum_{n'=0}^{n}\sum_{m=0}^{\lfloor\frac{n'}{2}\rfloor}\to\sum_{m=0}^{\lfloor\frac{n}{2}\rfloor}\sum_{n'=2m}^{n}\to\sum_{m=0}^{\lfloor\frac{n}{2}\rfloor}\sum_{n'=0}^{n-2m},\tag{B92}$$

(relabelling $n' \to n - n'$) which leaves the expression in the form

$$\pi^{\frac{N'}{2}} \sqrt{\frac{\det V}{\det V_{\mathcal{J}}}} \exp(\frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^T B \boldsymbol{a}_{\mathcal{I}}) \exp(-\boldsymbol{x}_{\mathcal{J}}^T V_{\mathcal{J}}^{-1} \boldsymbol{x}_{\mathcal{J}} + \theta^T \boldsymbol{x}_{\mathcal{J}}) \tag{B93}$$
$$\times \sum_{m=0}^{\lfloor \frac{n}{2} \rfloor} \sum_{n'=0}^{n-2m} (\frac{1}{2})^{n-n'} \frac{n!}{m!\, n'!\, (n-2m-n')!} \Big[\boldsymbol{\lambda}_{\mathcal{J}}^T \boldsymbol{x}_{\mathcal{J}}\Big]^{n'} \Big[\boldsymbol{\lambda}_{\mathcal{I}}^T B \boldsymbol{\lambda}_{\mathcal{I}}\Big]^{m} \Big[\boldsymbol{\lambda}_{\mathcal{I}}^T B^T \boldsymbol{a}_{\mathcal{I}} + 2\boldsymbol{\lambda}_{\mathcal{I}}^T D^T \boldsymbol{x}_{\mathcal{J}}\Big]^{n-2m-n'}$$
$$= \pi^{\frac{N'}{2}} \sqrt{\frac{\det V}{\det V_{\mathcal{J}}}} \exp(\frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^T B \boldsymbol{a}_{\mathcal{I}}) \exp(-\boldsymbol{x}_{\mathcal{J}}^T V_{\mathcal{J}}^{-1} \boldsymbol{x}_{\mathcal{J}} + \theta^T \boldsymbol{x}_{\mathcal{J}}) \tag{B94}$$
$$\times \sum_{m=0}^{\lfloor \frac{n}{2} \rfloor} (\frac{1}{2})^{n} \Bigg\{ \frac{n!}{m!(n-2m)!} \Big[\boldsymbol{\lambda}_{\mathcal{I}}^T B \boldsymbol{\lambda}_{\mathcal{I}}\Big]^{m} \Bigg\}$$
$$\times \Bigg\{ \sum_{n'=0}^{n-2m} \frac{(n-2m)!}{n'!\,(n-2m-n')!} \Big[2\boldsymbol{\lambda}_{\mathcal{J}}^T \boldsymbol{x}_{\mathcal{J}}\Big]^{n'} \Big[\boldsymbol{\lambda}_{\mathcal{I}}^T B^T \boldsymbol{a}_{\mathcal{I}} + 2\boldsymbol{\lambda}_{\mathcal{I}}^T D^T \boldsymbol{x}_{\mathcal{J}}\Big]^{n-2m-n'} \Bigg\}.$$

Finally, we can use the binomial theorem to factorise the second term as

$$\pi^{\frac{N'}{2}} \sqrt{\frac{\det V}{\det V_{\mathcal{J}}}} \exp(\frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^T B \boldsymbol{a}_{\mathcal{I}}) \exp(-\boldsymbol{x}_{\mathcal{J}}^T V_{\mathcal{J}}^{-1} \boldsymbol{x}_{\mathcal{J}} + \theta^T \boldsymbol{x}_{\mathcal{J}}) \tag{B95}$$
$$\times \sum_{m=0}^{\lfloor \frac{n}{2} \rfloor} (\frac{1}{2})^{n} \frac{n!}{m!(n-2m)!} \Big[\boldsymbol{\lambda}_{\mathcal{I}}^T B \boldsymbol{\lambda}_{\mathcal{I}}\Big]^{m} \Big[2\boldsymbol{\lambda}_{\mathcal{J}}^T \boldsymbol{x}_{\mathcal{J}} + \boldsymbol{\lambda}_{\mathcal{I}}^T B^T \boldsymbol{a}_{\mathcal{I}} + 2\boldsymbol{\lambda}_{\mathcal{I}}^T D^T \boldsymbol{x}_{\mathcal{J}}\Big]^{n-2m}$$
$$= \pi^{\frac{N'}{2}} \sqrt{\frac{\det V}{\det V_{\mathcal{J}}}} \exp(\frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^T B \boldsymbol{a}_{\mathcal{I}}) \exp(-\boldsymbol{x}_{\mathcal{J}}^T V_{\mathcal{J}}^{-1} \boldsymbol{x}_{\mathcal{J}} + \theta^T \boldsymbol{x}_{\mathcal{J}}) \tag{B96}$$
$$\times \sum_{m=0}^{\lfloor \frac{n}{2} \rfloor} (\frac{1}{2})^{n} \frac{n!}{m!(n-2m)!} \Big[\boldsymbol{\lambda}_{\mathcal{I}}^T B \boldsymbol{\lambda}_{\mathcal{I}}\Big]^{m} \Big[2(\boldsymbol{\lambda}_{\mathcal{J}} + D\boldsymbol{\lambda}_{\mathcal{I}})^T \boldsymbol{x}_{\mathcal{J}} + \boldsymbol{\lambda}_{\mathcal{I}}^T B^T \boldsymbol{a}_{\mathcal{I}}\Big]^{n-2m},$$

which we can identify as a Hermite polynomial of the form

$$H_n(x) = n! \sum_{m=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^m}{m!(n-2m)!} (2x)^{n-2m}, \tag{B97}$$

to rewrite the result succinctly as

$$\pi^{\frac{N'}{2}} \sqrt{\frac{\det V}{\det V_{\mathcal{J}}}} \Big(-\mathrm{i}\, \frac{\sqrt{\boldsymbol{\lambda}_{\mathcal{I}}^T B \boldsymbol{\lambda}_{\mathcal{I}}}}{2}\Big)^{n} H_n\Big(\mathrm{i}\, \frac{(\boldsymbol{\lambda}_{\mathcal{J}} + D\boldsymbol{\lambda}_{\mathcal{I}})^T \boldsymbol{x}_{\mathcal{J}} + \frac{1}{2}\boldsymbol{\lambda}_{\mathcal{I}}^T B^T \boldsymbol{a}_{\mathcal{I}}}{\sqrt{\boldsymbol{\lambda}_{\mathcal{I}}^T B \boldsymbol{\lambda}_{\mathcal{I}}}}\Big) \tag{B98}$$
$$\times \exp(-\boldsymbol{x}_{\mathcal{J}}^T V_{\mathcal{J}}^{-1} \boldsymbol{x}_{\mathcal{J}} + \theta^T \boldsymbol{x}_{\mathcal{J}} + \frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^T B \boldsymbol{a}_{\mathcal{I}}),$$

completing the proof. □

Corollary B.6. *The partial integration of theorem B.5 with no linear term, given by*

$$\int_{\mathbb{R}^{N'}} d^{N'}x_{\mathcal{I}}(\boldsymbol{\lambda}\cdot\boldsymbol{x})^n \exp(-\boldsymbol{x}^T V^{-1}\boldsymbol{x}), \tag{B99}$$

over only the first N' members of $\boldsymbol{x}$ so that the relevant vectors and matrices split into 'integrated' $\mathcal{I}$ and 'non-integrated' $\mathcal{J}$ parts as

$$\boldsymbol{x} = \boldsymbol{x}_{\mathcal{I}} \oplus \boldsymbol{x}_{\mathcal{J}}, \qquad \boldsymbol{\lambda} = \boldsymbol{\lambda}_{\mathcal{I}} \oplus \boldsymbol{\lambda}_{\mathcal{J}}, \qquad V = \begin{pmatrix} V_{\mathcal{I}} & V_{\mathcal{I}\mathcal{J}} \\ V_{\mathcal{J}\mathcal{I}} & V_{\mathcal{J}} \end{pmatrix}$$

is given by

$$\pi^{\frac{N'}{2}}\sqrt{\frac{\det V}{\det V_{\mathcal{J}}}}\Big(\mathrm{i}\,\frac{\sqrt{\boldsymbol{\lambda}_{\mathcal{I}}^T B\boldsymbol{\lambda}_{\mathcal{I}}}}{2}\Big)^n H_n\Big(-\mathrm{i}\,\frac{(\boldsymbol{\lambda}_{\mathcal{J}}+D\boldsymbol{\lambda}_{\mathcal{I}})^T\boldsymbol{x}_{\mathcal{J}}}{\sqrt{\boldsymbol{\lambda}_{\mathcal{I}}^T B\boldsymbol{\lambda}_{\mathcal{I}}}}\Big)\exp(-\boldsymbol{x}_{\mathcal{J}}^T V_{\mathcal{J}}^{-1}\boldsymbol{x}_{\mathcal{J}}) \tag{B100}$$

where H_n is the nth Hermite polynomial,

$$\begin{aligned} B &= V/V_{\mathcal{I}} = [(V^{-1})_{\mathcal{I}}]^{-1} = V_{\mathcal{I}} - V_{\mathcal{I}\mathcal{J}}V_{\mathcal{J}}^{-1}V_{\mathcal{J}\mathcal{I}} \\ D &= V_{\mathcal{J}}^{-1}V_{\mathcal{J}\mathcal{I}} = V_{\mathcal{J}}^{-1}V_{\mathcal{J}\mathcal{I}}, \end{aligned}$$

and $V_{\mathcal{X}}^{-1} = (V_{\mathcal{X}})^{-1}$ is the inverse of the $\mathcal{X}$ block of V.
We have presented this theorem in block form for notational convenience. It equivalently applies to a set of integration and non-integration indices which do not form a contiguous block with appropriately defined $\mathcal{I}$, $\mathcal{J}$ vectors and matrices.

Proof. This corollary follows immediately by setting $\boldsymbol{a} = 0$ in theorem B.5 and noting that therefore $\theta = 0$. □

B.4.2 *Gaussian integrals of* $[(\lambda_x\cdot x)^2 + (\lambda_p\cdot p)^2]^n$

Let us now consider integrals over two independent sets of variables, $\boldsymbol{x}$ and $\boldsymbol{p}$ which form a single integrand as

$$\begin{aligned} &[(\lambda_x\cdot x)^2 + (\lambda_p\cdot p)^2]^n \exp(-\boldsymbol{q}^T V^{-1}\boldsymbol{q} + \boldsymbol{a}^T\boldsymbol{q}) \\ =&[(\lambda_x\cdot x)^2 + (\lambda_p\cdot p)^2]^n \exp(-\boldsymbol{x}^T V_x^{-1}\boldsymbol{x} + \boldsymbol{a}_x\boldsymbol{x})\exp(-\boldsymbol{p}^T V_p^{-1}\boldsymbol{p} + \boldsymbol{a}_p\boldsymbol{p}) \end{aligned} \tag{B101}$$

for $\boldsymbol{q} = \boldsymbol{x}\oplus\boldsymbol{p}$.

We will assume in this section that the variables contain some symmetry between x and p such that

$$\boldsymbol{\lambda}_{x\mathcal{I}}^T(V_x/V_{x\mathcal{I}})\boldsymbol{\lambda}_{x\mathcal{I}} = \boldsymbol{\lambda}_{p\mathcal{I}}^T(V_p/V_{p\mathcal{I}})\boldsymbol{\lambda}_{p\mathcal{I}} \tag{B102}$$

for $\lambda_{x\mathcal{I}}, \lambda_{p\mathcal{I}}$ the portions of $\lambda_{x/p}$ respectively corresponding to integrated variables.

Theorem B.7. *The partial integral*

$$\int_{\mathbb{R}^{2N'}} d^{N'}\boldsymbol{x}_{\mathcal{I}}\, d^{N'}\boldsymbol{p}_{\mathcal{I}}\, [(\boldsymbol{\lambda}_x \cdot \boldsymbol{x})^2 + (\boldsymbol{\lambda}_p \cdot \boldsymbol{p})^2]^n \exp(-\boldsymbol{q}^T V^{-1}\boldsymbol{q} + \boldsymbol{a}^T\boldsymbol{q}), \tag{B103}$$

for $\boldsymbol{q} = \boldsymbol{x} \oplus \boldsymbol{p}$ performed over only the final N' members of $\boldsymbol{x}$ and $\boldsymbol{p}$, such that the relevant vectors and matrices split into 'integrated' $\mathcal{I}$ and 'non-integrated' $\mathcal{J}$ parts is given by

$$\begin{aligned}&\pi^{N'}\sqrt{\frac{\det V}{\det V_{\mathcal{J}}}}\exp(\frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^T(V/V_{\mathcal{J}})\boldsymbol{a}_{\mathcal{I}})\exp(-\boldsymbol{q}_{\mathcal{J}}^T V_{\mathcal{J}}^{-1}\boldsymbol{q}_{\mathcal{J}} + \theta^T\boldsymbol{q}_{\mathcal{J}})n!\\ &\times n! G^n L_n\Big(-\frac{(\boldsymbol{A}_x^T\boldsymbol{x}_{\mathcal{J}} + \frac{1}{2}F_x)^2 + (\boldsymbol{A}_p^T\boldsymbol{p}_{\mathcal{J}} + \frac{1}{2}F_p)^2}{G}\Big),\end{aligned} \tag{B104}$$

when the condition

$$\boldsymbol{\lambda}_{x\mathcal{I}}^T(V_x/V_{x\mathcal{J}})\boldsymbol{\lambda}_{x\mathcal{I}} = \boldsymbol{\lambda}_{p\mathcal{I}}^T(V_p/V_{p\mathcal{J}})\boldsymbol{\lambda}_{p\mathcal{I}} \coloneqq G \tag{B105}$$

is satisfied. Here, $L_n(\cdot)$ denotes the nth Laguerre polynomial,

$$\begin{aligned}\theta &= \boldsymbol{a}_{\mathcal{J}} + V_{\mathcal{J}}^{-1}V_{\mathcal{J}\mathcal{I}}\boldsymbol{a}_{\mathcal{I}},\\ \boldsymbol{A}_x &= \boldsymbol{\lambda}_{x\mathcal{J}} + V_{x\mathcal{J}}^{-1}V_{x\mathcal{J}\mathcal{I}}\boldsymbol{\lambda}_{x\mathcal{I}}, & \boldsymbol{A}_p &= \boldsymbol{\lambda}_{p\mathcal{J}} + V_{p\mathcal{J}}^{-1}V_{p\mathcal{J}\mathcal{I}}\boldsymbol{\lambda}_{p\mathcal{I}},\\ F_x &= \boldsymbol{\lambda}_{x\mathcal{I}}^T(V_x/V_{x\mathcal{J}})\boldsymbol{a}_{x\mathcal{I}}, & F_p &= \boldsymbol{\lambda}_{p\mathcal{I}}^T(V_p/V_{p\mathcal{J}})\boldsymbol{a}_{p\mathcal{I}},\end{aligned}$$

and M is assumed to refer to $M_x \oplus M_p$ when presented with no x/p subscript.

Proof. First, let us expand the binomial into two integrals

$$\int_{\mathbb{R}^{2N'}} d^{2N'}\boldsymbol{q}_{\mathcal{I}}[(\boldsymbol{\lambda}_x \cdot \boldsymbol{x})^2 + (\boldsymbol{\lambda}_p \cdot \boldsymbol{p})^2]^n \exp(-\boldsymbol{q}^T V^{-1}\boldsymbol{q} + \boldsymbol{a}^T\boldsymbol{q}) \tag{B106}$$

$$= \sum_{n_x=0}^{n}\binom{n}{n_x}\left[\int_{\mathbb{R}^{N'}} d^{N'}\boldsymbol{x}_{\mathcal{I}}(\boldsymbol{\lambda}_x \cdot \boldsymbol{x})^{2n_x}\exp(-\boldsymbol{x}^T V_x^{-1}\boldsymbol{x} + \boldsymbol{a}_x\boldsymbol{x})\right] \tag{B107}$$

$$\times\left[\int_{\mathbb{R}^{N'}} d^{N'}\boldsymbol{p}_{\mathcal{I}}(\boldsymbol{\lambda}_p \cdot \boldsymbol{p})^{2(n-n_x)}\exp(-\boldsymbol{p}^T V_p^{-1}\boldsymbol{p} + \boldsymbol{a}_p\boldsymbol{p})\right],$$

each of the form solved in theorem B.5. Substituting that result and recombining $\boldsymbol{V} = \boldsymbol{V}_x \oplus \boldsymbol{V}_p$ gives

$$\begin{aligned}&\sum_{n_x=0}^{n}\binom{n}{n_x}\pi^{N'}\sqrt{\frac{\det V}{\det V_{\mathcal{J}}}}\exp(\frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^T B\boldsymbol{a}_{\mathcal{I}})\exp(-\boldsymbol{q}_{\mathcal{J}}^T V_{\mathcal{J}}^{-1}\boldsymbol{q}_{\mathcal{J}} + \theta^T\boldsymbol{q}_{\mathcal{J}})\\ &\times\left\{\Big(-\mathrm{i}\frac{\sqrt{\boldsymbol{\lambda}_{x\mathcal{I}}^T B_x\boldsymbol{\lambda}_{x\mathcal{I}}}}{2}\Big)^{2n_x}H_{2n_x}\Big(\mathrm{i}\frac{(\boldsymbol{\lambda}_{x\mathcal{J}} + D_x\boldsymbol{\lambda}_{x\mathcal{I}})^T\boldsymbol{x}_{\mathcal{J}} + \frac{1}{2}\boldsymbol{\lambda}_{x\mathcal{I}}^T B_x^T\boldsymbol{a}_{x\mathcal{I}}}{\sqrt{\boldsymbol{\lambda}_{x\mathcal{I}}^T B_x\boldsymbol{\lambda}_{x\mathcal{I}}}}\Big)\right\}\\ &\times\left\{\Big(-\mathrm{i}\frac{\sqrt{\boldsymbol{\lambda}_{p\mathcal{I}}^T B_p\boldsymbol{\lambda}_{p\mathcal{I}}}}{2}\Big)^{2(n-n_x)}H_{2(n-n_x)}\Big(\mathrm{i}\frac{(\boldsymbol{\lambda}_{p\mathcal{J}} + D_p\boldsymbol{\lambda}_{p\mathcal{I}})^T\boldsymbol{p}_{\mathcal{J}} + \frac{1}{2}\boldsymbol{\lambda}_{p\mathcal{I}}^T B_p^T\boldsymbol{a}_{p\mathcal{I}}}{\sqrt{\boldsymbol{\lambda}_{p\mathcal{I}}^T B_p\boldsymbol{\lambda}_{p\mathcal{I}}}}\Big)\right\},\end{aligned} \tag{B108}$$

for $\theta^T\boldsymbol{q}_{\mathcal{J}} = \theta_x^T\boldsymbol{x}_{\mathcal{J}} + \theta_p^T\boldsymbol{p}_{\mathcal{J}}$, and $\theta_{x/p}$, $B_{x/p}$, $D_{x/p}$ as defined in theorem B.5.

Let us now, for notational convenience, denote

$$\chi \coloneqq \pi^{N'} \sqrt{\frac{\det V}{\det V_{\mathcal{J}}}} \exp(\frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^T B \boldsymbol{a}_{\mathcal{I}}) \exp(-\boldsymbol{q}_{\mathcal{J}}^T V_{\mathcal{J}}^{-1} \boldsymbol{q}_{\mathcal{J}} + \theta^T \boldsymbol{q}_{\mathcal{J}}) \tag{B109}$$

$$G_x \coloneqq \boldsymbol{\lambda}_{x\mathcal{I}}^T B_x \boldsymbol{\lambda}_{x\mathcal{I}} = \boldsymbol{\lambda}_{x\mathcal{I}}^T (V_{x\mathcal{I}} - V_{x\mathcal{I}\mathcal{J}} V_{x\mathcal{J}}^{-1} V_{x\mathcal{J}\mathcal{I}}) \boldsymbol{\lambda}_{x\mathcal{I}}, \tag{B110}$$

$$\boldsymbol{A}_x \coloneqq \boldsymbol{\lambda}_{x\mathcal{J}} + D_x \boldsymbol{\lambda}_{x\mathcal{I}} = \boldsymbol{\lambda}_{x\mathcal{J}} + V_{x\mathcal{J}}^{-1} V_{x\mathcal{J}\mathcal{I}} \boldsymbol{\lambda}_{x\mathcal{I}} \tag{B111}$$

$$F_x \coloneqq \boldsymbol{\lambda}_{x\mathcal{I}}^T B_x \boldsymbol{a}_{x\mathcal{I}} = \boldsymbol{\lambda}_{x\mathcal{I}}^T (V_{x\mathcal{I}} - V_{x\mathcal{I}\mathcal{J}} V_{x\mathcal{J}}^{-1} V_{x\mathcal{J}\mathcal{I}}) \boldsymbol{a}_{\mathcal{I}}, \tag{B112}$$

and similarly define G_p, $\boldsymbol{A}_p$, and F_p, which allows us to rewrite equation (B108) as

$$\chi \sum_{n_x=0}^{n} \binom{n}{n_x} \Big(-\mathrm{i}\frac{\sqrt{G_x}}{2}\Big)^{2n_x} \Big(-\mathrm{i}\frac{\sqrt{G_p}}{2}\Big)^{2(n-n_x)} H_{2n_x}\Big(\mathrm{i}\frac{\boldsymbol{A}_x^T \boldsymbol{x}_{\mathcal{J}} + \frac{1}{2}F_x}{\sqrt{G_x}}\Big) H_{2(n-n_x)}\Big(\mathrm{i}\frac{\boldsymbol{A}_p^T \boldsymbol{p}_{\mathcal{J}} + \frac{1}{2}F_p}{\sqrt{G_p}}\Big). \tag{B113}$$

We can then use the Hermite-Laguerre polynomial identity [75, (18.7.19)], that

$$H_{2n}(x) = (-4)^n n! L_n^{(-\frac{1}{2})}(x^2), \tag{B114}$$

to rewrite this as a sum of Laguerre polynomials as

$$\chi n! \sum_{n_x=0}^{n} G_x^{n_x} G_p^{n-nx} L_{n_x}^{(-\frac{1}{2})}\Big(-\frac{(\boldsymbol{A}_x^T \boldsymbol{x}_{\mathcal{J}} + \frac{1}{2}F_x)^2}{G_x}\Big) L_{n-n_x}^{(-\frac{1}{2})}\Big(-\frac{(\boldsymbol{A}_p^T \boldsymbol{p}_{\mathcal{J}} + \frac{1}{2}F_p)^2}{G_p}\Big). \tag{B115}$$

In the limited case in which $G_x = G_p \coloneqq G$, we can then use the Laguerre polynomial addition formula [75, (18.18.10)],

$$\sum_{i}^{n} L_i^{(\alpha)}(A) L_{n-i}^{(\beta)}(B) = L_n^{(\alpha+\beta+1)}(A+B), \tag{B116}$$

to rewrite this expression as

$$\chi n! G^n L_n\Big(-\frac{(\boldsymbol{A}_x^T \boldsymbol{x}_{\mathcal{J}} + \frac{1}{2}F_x)^2 + (\boldsymbol{A}_p^T \boldsymbol{p}_{\mathcal{J}} + \frac{1}{2}F_p)^2}{G}\Big), \tag{B117}$$

which completes the proof. □

B.5 GAUSSIAN INTEGRALS OF LAGUERRE POLYNOMIALS

B.5.1 *General case*

In this section we will finally solve the integrals of primary interest to us in this section, that of the product of a Laguerre polynomial with a Gaussian, which closely represents the Fock state.

Theorem B.8. *The partial integral*

$$\int_{\mathbb{R}^{2N'}} d^{N'}x_{\mathcal{I}}\, d^{N'}p_{\mathcal{I}}\, L_n\Big([\boldsymbol{\lambda}_x\cdot\boldsymbol{x}]^2+[\boldsymbol{\lambda}_p\cdot\boldsymbol{p}]^2\Big)\exp(-\boldsymbol{q}^TV^{-1}\boldsymbol{q}+\boldsymbol{a}^T\boldsymbol{q}), \tag{B118}$$

where $L_n(\cdot)$ is the Laguerre polynomial of degree n, and $\boldsymbol{q}=\boldsymbol{x}\oplus\boldsymbol{p}$, performed over only the final N' members of $\boldsymbol{x}$ and $\boldsymbol{p}$ so that the relevant vectors and matrices split into 'integrated' $\mathcal{I}$ and 'non-integrated' $\mathcal{J}$ parts, under the condition that

$$\boldsymbol{\lambda}_{x\mathcal{I}}^T(V_x/V_{x\mathcal{J}})\boldsymbol{\lambda}_{x\mathcal{I}}=\boldsymbol{\lambda}_{p\mathcal{I}}^T(V_p/V_{p\mathcal{J}})\boldsymbol{\lambda}_{p\mathcal{I}} \tag{B119}$$

is given by

$$\begin{aligned}&\pi^{N'}\sqrt{\frac{\det V}{\det V_{\mathcal{J}}}}\exp(\frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^TB\boldsymbol{a}_{\mathcal{I}})\exp(-\boldsymbol{q}_{\mathcal{J}}^TV_{\mathcal{J}}^{-1}\boldsymbol{q}_{\mathcal{J}}+\theta^T\boldsymbol{q}_{\mathcal{J}})\\&\times\sum_{m=0}^{n}\binom{n}{m}(-G)^mL_m\Big(\frac{-1}{4G}\big[(2\boldsymbol{A}_x^T\boldsymbol{x}_{\mathcal{J}}+F_x)^2+(2\boldsymbol{A}_p^T\boldsymbol{p}_{\mathcal{J}}+F_p)^2\big]\Big),\end{aligned} \tag{B120}$$

or by

$$\begin{aligned}=&\pi^{N'}\sqrt{\frac{\det V}{\det V_{\mathcal{J}}}}\exp(\frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^TB\boldsymbol{a}_{\mathcal{I}})\exp(-\boldsymbol{q}_{\mathcal{J}}^TV_{\mathcal{J}}^{-1}\boldsymbol{q}_{\mathcal{J}}+\theta^T\boldsymbol{q}_{\mathcal{J}})\\&\times(1-G)^nL_n\Big(\frac{1}{1-G}\big[(\boldsymbol{A}_x^T\boldsymbol{x}_{\mathcal{J}}+F_x)^2+(\boldsymbol{A}_p^T\boldsymbol{p}_{\mathcal{J}}+F_p)^2\big]\Big)\end{aligned} \tag{B121}$$

where

$$\begin{aligned}\theta&=\boldsymbol{a}_{\mathcal{J}}+V_{\mathcal{J}}^{-1}V_{\mathcal{J}\mathcal{I}}\boldsymbol{a}_{\mathcal{I}},&\qquad G&=\boldsymbol{\lambda}_{x\mathcal{I}}^T(V_x/V_{x\mathcal{J}})\boldsymbol{\lambda}_{x\mathcal{I}}=\boldsymbol{\lambda}_{p\mathcal{I}}^T(V_p/V_{p\mathcal{J}})\boldsymbol{\lambda}_{p\mathcal{I}},\\\boldsymbol{A}_x&=\boldsymbol{\lambda}_{x\mathcal{J}}+V_{x\mathcal{J}}^{-1}V_{x\mathcal{J}\mathcal{I}}\boldsymbol{\lambda}_{x\mathcal{I}},&\qquad\boldsymbol{A}_p&=\boldsymbol{\lambda}_{p\mathcal{J}}+V_{p\mathcal{J}}^{-1}V_{p\mathcal{J}\mathcal{I}}\boldsymbol{\lambda}_{p\mathcal{I}},\\F_x&=\frac{1}{2}\boldsymbol{\lambda}_{x\mathcal{I}}^T(V_x/V_{x\mathcal{J}})\boldsymbol{a}_{x\mathcal{I}},&\qquad F_p&=\frac{1}{2}\boldsymbol{\lambda}_{p\mathcal{I}}^T(V_p/V_{p\mathcal{J}})\boldsymbol{a}_{p\mathcal{I}}.\end{aligned}$$

Proof. The first result follows immediately from theorem B.7. If we first expand the Laguerre polynomial into its closed form and substitute the integral result from that theorem, as

$$\begin{aligned}
&\int_{\mathbb{R}^{2N'}} d^{N'_{\mathcal{I}}}\boldsymbol{x}_{\mathcal{I}}\, d^{N'_{\mathcal{I}}}\boldsymbol{p}_{\mathcal{I}}\, L_n\Big([\boldsymbol{\lambda}_x\cdot\boldsymbol{x}]^2+[\boldsymbol{\lambda}_p\cdot\boldsymbol{p}]^2\Big)\exp(-\boldsymbol{q}^T V^{-1}\boldsymbol{q}+\boldsymbol{a}^T\boldsymbol{q})\\
&=\sum_{m=0}^{n}\binom{n}{m}\frac{(-1)^m}{m!}\int_{\mathbb{R}^{2N'}} d^{N'_{\mathcal{I}}}\boldsymbol{x}_{\mathcal{I}}\, d^{N'_{\mathcal{I}}}\boldsymbol{p}_{\mathcal{I}}\left[(\boldsymbol{\lambda}_x\cdot\boldsymbol{x})^2+(\boldsymbol{\lambda}_p\cdot\boldsymbol{p})^2\right]^m\exp(-\boldsymbol{q}^T V^{-1}\boldsymbol{q}+\boldsymbol{a}^T\boldsymbol{q})
\end{aligned}\tag{B122}$$

$$\begin{aligned}
&=\pi^{N'}\sqrt{\frac{\det V}{\det V_{\mathcal{J}}}}\exp(\frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^T B\boldsymbol{a}_{\mathcal{I}})\exp(-\boldsymbol{q}_{\mathcal{J}}^T V_{\mathcal{J}}^{-1}\boldsymbol{q}_{\mathcal{J}}+\theta^T\boldsymbol{q}_{\mathcal{J}})\\
&\quad\times\sum_{m=0}^{n}\binom{n}{m}\frac{(-1)^m}{m!}m!\,G^m\,L_m\Big(\frac{-1}{4G}\big[(2\boldsymbol{A}_x^T\boldsymbol{x}_{\mathcal{J}}+F_x)^2+(2\boldsymbol{A}_p^T\boldsymbol{p}_{\mathcal{J}}+F_p)^2\big]\Big)
\end{aligned}\tag{B123}$$

$$\begin{aligned}
&=\pi^{N'}\sqrt{\frac{\det V}{\det V_{\mathcal{J}}}}\exp(\frac{1}{4}\boldsymbol{a}_{\mathcal{I}}^T B\boldsymbol{a}_{\mathcal{I}})\exp(-\boldsymbol{q}_{\mathcal{J}}^T V_{\mathcal{J}}^{-1}\boldsymbol{q}_{\mathcal{J}}+\theta^T\boldsymbol{q}_{\mathcal{J}})\\
&\quad\times\sum_{m=0}^{n}\binom{n}{m}(-G)^m L_m\Big(\frac{-1}{4G}\big[(2\boldsymbol{A}_x^T\boldsymbol{x}_{\mathcal{J}}+F_x)^2+(2\boldsymbol{A}_p^T\boldsymbol{p}_{\mathcal{J}}+F_p)^2\big]\Big).
\end{aligned}\tag{B124}$$

To derive the second result we start by expanding the Laguerre polynomials so that

$$\begin{aligned}
&\sum_{m=0}^{n}\binom{n}{m}(-G)^m L_m\Big(\frac{-1}{4G}\big[(2\boldsymbol{A}_x^T\boldsymbol{x}_{\mathcal{J}}+F_x)^2+(2\boldsymbol{A}_p^T\boldsymbol{p}_{\mathcal{J}}+F_p)^2\big]\Big)\\
&=\sum_{m=0}^{n}\binom{n}{m}(-G)^m\sum_{b=0}^{m}\binom{m}{b}\frac{(-1)^b}{b!}\Big(\frac{-1}{4G}\big[(2\boldsymbol{A}_x^T\boldsymbol{x}_{\mathcal{J}}+F_x)^2+(2\boldsymbol{A}_p^T\boldsymbol{p}_{\mathcal{J}}+F_p)^2\big]\Big)^b.
\end{aligned}\tag{B125}$$

Reordering the summations as

$$\sum_{m=0}^{n}\sum_{b=0}^{m}\rightarrow\sum_{b=0}^{n}\sum_{m=b}^{n},\tag{B126}$$

and relabelling the summation index so that $m \to m + b$ then leaves

$$\sum_{b=0}^{n}\sum_{m=0}^{n-b}\binom{n}{m+b}\binom{m+b}{b}(-G)^{m+b}\frac{(-1)^b}{b!}\Big(\frac{-1}{4G}\big[(2\boldsymbol{A}_x^T\boldsymbol{x}_{\mathcal{J}}+F_x)^2+(2\boldsymbol{A}_p^T\boldsymbol{p}_{\mathcal{J}}+F_p)^2\big]\Big)^b \tag{B127}$$

$$=\sum_{b=0}^{n}\sum_{m=0}^{n-b}\frac{n!}{b!b!m![n-m-b]!}(-G)^{m+b}(-1)^b\Big(\frac{-1}{4G}\big[(2\boldsymbol{A}_x^T\boldsymbol{x}_{\mathcal{J}}+F_x)^2+(2\boldsymbol{A}_p^T\boldsymbol{p}_{\mathcal{J}}+F_p)^2\big]\Big)^b \tag{B128}$$

$$=\sum_{b=0}^{n}\frac{n!}{b!b![n-b]!}\Big[\sum_{m=0}^{n-b}\frac{[n-b]!}{m![n-b-m]!}(-G)^m\Big]G^b\Big(\frac{-1}{4G}\big[(2\boldsymbol{A}_x^T\boldsymbol{x}_{\mathcal{J}}+F_x)^2+(2\boldsymbol{A}_p^T\boldsymbol{p}_{\mathcal{J}}+F_p)^2\big]\Big)^b \tag{B129}$$

$$=\sum_{b=0}^{n}\binom{n}{b}\frac{1}{b!}(1-G)^{n-b}\Big(-\frac{1}{4}\big[(2\boldsymbol{A}_x^T\boldsymbol{x}_{\mathcal{J}}+F_x)^2+(2\boldsymbol{A}_p^T\boldsymbol{p}_{\mathcal{J}}+F_p)^2\big]\Big)^b \tag{B130}$$

$$=(1-G)^n\sum_{b=0}^{n}\binom{n}{b}\frac{(-1)^b}{b!}\Big(\frac{1}{4}\frac{1}{1-G}\big[(2\boldsymbol{A}_x^T\boldsymbol{x}_{\mathcal{J}}+F_x)^2+(2\boldsymbol{A}_p^T\boldsymbol{p}_{\mathcal{J}}+F_p)^2\big]\Big)^b \tag{B131}$$

$$=(1-G)^n L_n\Big(\frac{1}{4}\frac{1}{1-G}\big[(2\boldsymbol{A}_x^T\boldsymbol{x}_{\mathcal{J}}+F_x)^2+(2\boldsymbol{A}_p^T\boldsymbol{p}_{\mathcal{J}}+F_p)^2\big]\Big) \tag{B132}$$

Recombining this with the first part from equation (B124) completes the proof. □

Corollary B.9. *The partial integral of theorem B.8 with no linear term,*

$$\int_{\mathbb{R}^{2N'}} d^{N'}\boldsymbol{x}_{\mathcal{I}}\, d^{N'}\boldsymbol{p}_{\mathcal{I}}\, L_n\Big([\boldsymbol{\lambda}_x\cdot\boldsymbol{x}]^2+[\boldsymbol{\lambda}_p\cdot\boldsymbol{p}]^2\Big)\exp(-\boldsymbol{q}^T V^{-1}\boldsymbol{q}), \tag{B133}$$

where $L_n(\cdot)$ is the Laguerre polynomial of degree n, and for $\boldsymbol{q}=\boldsymbol{x}\oplus\boldsymbol{p}$, performed over only the first N' members of $\boldsymbol{x}$ and $\boldsymbol{p}$ so that the relevant vectors and matrices split into 'integrated' $\mathcal{I}$ and 'non-integrated' $\mathcal{J}$ parts as

$$\begin{aligned}
&\boldsymbol{x}=\boldsymbol{x}_{\mathcal{I}}\oplus\boldsymbol{x}_{\mathcal{J}}, && \boldsymbol{p}=\boldsymbol{p}_{\mathcal{I}}\oplus\boldsymbol{p}_{\mathcal{J}}, && \boldsymbol{q}=\boldsymbol{x}\oplus\boldsymbol{p},\\
&V_x=\begin{pmatrix}V_{x\mathcal{I}} & V_{x\mathcal{I}\mathcal{J}}\\ V_{x\mathcal{J}\mathcal{I}} & V_{x\mathcal{J}}\end{pmatrix}, && V_p=\begin{pmatrix}V_{p\mathcal{I}} & V_{p\mathcal{I}\mathcal{J}}\\ V_{p\mathcal{J}\mathcal{I}} & V_{p\mathcal{J}}\end{pmatrix}, && V=V_x\oplus V_p,\\
&\boldsymbol{\lambda}_x=\boldsymbol{\lambda}_{x\mathcal{I}}\oplus\boldsymbol{\lambda}_{x\mathcal{J}}, && \boldsymbol{\lambda}_p=\boldsymbol{\lambda}_{p\mathcal{I}}\oplus\boldsymbol{\lambda}_{p\mathcal{J}}, &&
\end{aligned}$$

under the condition that

$$\boldsymbol{\lambda}_{x\mathcal{I}}^T(V_{x\mathcal{I}}-V_{x\mathcal{I}\mathcal{J}}V_{x\mathcal{J}}^{-1}V_{x\mathcal{J}\mathcal{I}})\boldsymbol{\lambda}_{x\mathcal{I}}=\boldsymbol{\lambda}_{p\mathcal{I}}^T(V_{p\mathcal{I}}-V_{p\mathcal{I}\mathcal{J}}V_{p\mathcal{J}}^{-1}V_{p\mathcal{J}\mathcal{I}})\boldsymbol{\lambda}_{p\mathcal{I}} \tag{B134}$$

is given by

$$\pi^{N'}\sqrt{\frac{\det V}{\det V_{\mathcal{J}}}}(1-G)^n L_n\Big(\frac{1}{1-G}\big[(\boldsymbol{A}_x^T\boldsymbol{x}_{\mathcal{J}})^2+(\boldsymbol{A}_p^T\boldsymbol{p}_{\mathcal{J}})^2\big]\Big)\exp(-\boldsymbol{q}_{\mathcal{J}}^T V_{\mathcal{J}}^{-1}\boldsymbol{q}_{\mathcal{J}}) \tag{B135}$$

where

$$\begin{aligned}
\boldsymbol{A}_x &= \boldsymbol{\lambda}_{x\mathcal{J}} + V_{x\mathcal{J}}^{-1} V_{x\mathcal{J}\mathcal{I}} \boldsymbol{\lambda}_{x\mathcal{I}} \\
\boldsymbol{A}_p &= \boldsymbol{\lambda}_{p\mathcal{J}} + V_{p\mathcal{J}}^{-1} V_{p\mathcal{J}\mathcal{I}} \boldsymbol{\lambda}_{p\mathcal{I}} \\
G &= \boldsymbol{\lambda}_{x\mathcal{I}}^T (V_{x\mathcal{I}} - V_{x\mathcal{I}\mathcal{J}} V_{x\mathcal{J}}^{-1} V_{x\mathcal{J}\mathcal{I}}) \boldsymbol{\lambda}_{x\mathcal{I}} = \boldsymbol{\lambda}_{p\mathcal{I}}^T (V_{p\mathcal{I}} - V_{p\mathcal{I}\mathcal{J}} V_{p\mathcal{J}}^{-1} V_{p\mathcal{J}\mathcal{I}}) \boldsymbol{\lambda}_{p\mathcal{I}}, \\
V_{\mathcal{J}} &= V_{x\mathcal{J}} \oplus V_{p\mathcal{J}},
\end{aligned}$$

and $V_{\mathcal{X}}^{-1} = (V_{\mathcal{X}})^{-1}$ *is the inverse of the* $\mathcal{X}$ *block of* V.
We have presented this theorem in block form for notational convenience. It equivalently applies to a set of integration and non-integration indices which do not form a contiguous block we appropriately defined $\mathcal{I}$, $\mathcal{J}$ *vectors and matrices.*

Proof. This result follows immediately from theorem B.8 by setting $\boldsymbol{\alpha} = 0$ and noting that this implies $\theta = 0$ and $F_{x/p} = 0$. □

B.5.2 *Single-mode output case*

As our output states should be expected to be a mixture of Fock states, it will be instructive to rewrite the output of this integral as a summation over the Fock state Wigner functions. Before we do so, though, we prove a lemma that will come in useful.

B.5.2.1 *Preliminary lemma*

Lemma B.10. *For* $n, m \in \mathbb{Z}_{\geq 0}$, $A, B \in \mathbb{R}$ *and* $|B| < 1$,

$$\begin{aligned}
&\sum_{a=0}^{n} \sum_{b=\max(m-a,0)}^{\infty} \binom{n}{a}\binom{a+b}{m}\binom{a+b}{a} A^a B^b \\
=&\frac{B^m}{(1-B)^{n+m+1}} \sum_{b=0}^{\min(m,n)} \binom{n}{b}\binom{m}{b} \left(\frac{A}{B}\right)^b (A+1-B)^{n-b} \qquad \text{(B136)} \\
=&\frac{1}{(1-B)^{n+m+1}} \sum_{b=0}^{\min(m,n)} \binom{n}{b}\binom{m}{b} A^b B^{m-b} (A+1-B)^{n-b}. \qquad \text{(B137)}
\end{aligned}$$

Proof. Let us start by splitting the summation into $a \leq m$ and $a > m$ components as

$$\begin{aligned}
&\sum_{a=0}^{n} \sum_{b=\max(m-a,0)}^{\infty} \binom{n}{a}\binom{a+b}{m}\binom{a+b}{a} A^a B^b \\
=&\sum_{a=0}^{\min(m,n)} \sum_{b=m-a}^{\infty} \binom{n}{a}\binom{a+b}{m}\binom{a+b}{a} A^a B^b + \sum_{a=m+1}^{n} \sum_{b=0}^{\infty} \binom{n}{a}\binom{a+b}{m}\binom{a+b}{a} A^a B^b,
\end{aligned} \qquad \text{(B138)}$$

which we consider separately.

$a : 0 \mapsto m$ SUMMATION

Let us first consider the $a : 0 \mapsto m$ summation. Noting that the b summation extends to infinity, we can rebase the index as $b \mapsto b + m - a$ without changing the upper limit, to give

$$\sum_{a=0}^{\min(m,n)} \sum_{b=0}^{\infty} \binom{n}{a}\binom{m+b}{m}\binom{m+b}{a} A^a B^{m+b-a} \tag{B139}$$

$$= \sum_{a=0}^{\min(m,n)} \sum_{b=0}^{\infty} \frac{n!}{a!(n-a)!} \frac{(m+b)!}{m!b!} \frac{(m+b)!}{a!(m+b-a)!} A^a B^{m+b-a} \tag{B140}$$

$$= \sum_{a=0}^{\min(m,n)} \frac{m!n!}{a!a!(n-a)!(m-a)!} A^a B^{m-a} \sum_{b=0}^{\infty} \frac{1}{b!} \frac{(m+b)!}{m!} \frac{(m+b)!}{m!} \frac{(m-a)!}{(m-a+b)!} B^b \tag{B141}$$

$$= \sum_{a=0}^{\min(m,n)} \frac{m!n!}{a!a!(n-a)!(m-a)!} A^a B^{m-a} \sum_{b=0}^{\infty} \frac{1}{b!} \frac{(m+1)^{(b)}(m+1)^{(b)}}{(m-a+1)^{(b)}} B^b, \tag{B142}$$

for

$$(x)^{(a)} = \frac{(x+a-1)!}{(x-1)!} \tag{B143}$$

the rising factorial.

We can then identify the b summation as an instance of the ${}_2F_1$ hypergeometric series, defined for $|z| < 1$ as [135, (15.2.E1)]

$${}_2F_1\left(\begin{matrix}a, b\\ c\end{matrix}; z\right) = \sum_{i=0}^{\infty} \frac{1}{i!} \frac{(a)^{(i)}(b)^{(i)}}{(c)^{(i)}} z^i. \tag{B144}$$

Substituting this result in, noting that we have restricted B to $|B| < 1$, gives

$$\sum_{a=0}^{\min(m,n)} \frac{m!n!}{a!a!(n-a)!(m-a)!} A^a B^{m-a} \, {}_2F_1\left(\begin{matrix}m+1, m+1\\ m-a+1\end{matrix}; B\right). \tag{B145}$$

We can now use the Euler transformation [135, (15.8.E1)],

$${}_2F_1\left(\begin{matrix}a, b\\ c\end{matrix}; z\right) = (1-z)^{c-a-b} \, {}_2F_1\left(\begin{matrix}c-a, c-b\\ c\end{matrix}; z\right), \tag{B146}$$

to transform the hypergeometric function into one in which the upper parameters are negative integers,

$$\sum_{a=0}^{\min(m,n)} \frac{m!n!}{a!a!(n-a)!(m-a)!} A^a B^{m-a} (1-B)^{-a-m-1} \, {}_2F_1\left(\begin{matrix}-a, -a\\ m-a+1\end{matrix}; B\right). \tag{B147}$$

Transforming the upper parameters into negative integers means that the previously infinite series now terminates, as $(-a)^{(b)} = 0 \forall\, b > a$ and we can write the hypergeometric series as [135, (15.2.E4)]

$$\begin{aligned}
&\sum_{a=0}^{\min(m,n)} \frac{m!n!}{a!a!(n-a)!(m-a)!} \frac{A^a B^{m-a}}{(1-B)^{m+a+1}} \sum_{b=0}^{a} \frac{1}{b!} \frac{(-a)^{(b)}(-a)^{(b)}}{(m-a+1)^{(b)}} B^b \\
&= \sum_{a=0}^{\min(m,n)} \frac{m!n!}{a!a!(n-a)!(m-a)!} \frac{A^a B^{m-a}}{(1-B)^{m+a+1}} \sum_{b=0}^{a} \frac{1}{b!} \frac{(a)_{(b)}(a)_{(b)}}{(m-a+1)^{(b)}} B^b \qquad \text{(B148)} \\
&= \sum_{a=0}^{\min(m,n)} \frac{m!n!}{a!a!(n-a)!(m-a)!} \frac{A^a B^{m-a}}{(1-B)^{m+a+1}} \sum_{b=0}^{a} \frac{1}{b!} \frac{a!}{(a-b)!} \frac{a!}{(a-b)!} \frac{(m-a)!}{(m-a+b)!} B^b \qquad \text{(B149)} \\
&= \sum_{a=0}^{\min(m,n)} \frac{A^a B^{m-a}}{(1-B)^{m+a+1}} \sum_{b=0}^{a} \frac{m!n!}{b!(n-a)!(a-b)!(a-b)!(m-a+b)!} B^b, \qquad \text{(B150)}
\end{aligned}$$

where we have made use of the relationship between rising and falling factorials that $(-x)^{(a)} = (-1)^a (x)_{(a)}$. Finally, we invert the summation, making the change of index $b \to a - b$, to give

$$\begin{aligned}
&\sum_{a=0}^{\min(m,n)} \sum_{b=0}^{a} \frac{m!n!}{(a-b)!(n-a)!b!b!(m-b)!} \frac{A^a B^{m-b}}{(1-B)^{m+a+1}} \\
&= \sum_{a=0}^{\min(m,n)} \sum_{b=0}^{a} \binom{n}{a}\binom{a}{b}\binom{m}{b} \frac{A^a B^{m-b}}{(1-B)^{m+a+1}}. \qquad \text{(B151)}
\end{aligned}$$

$a : m+1 \mapsto n$ SUMMATION

Let us now return to the $a : m+1 \mapsto n$ summation, given by

$$\sum_{a=m+1}^{n} \sum_{b=0}^{\infty} \binom{n}{a}\binom{a+b}{m}\binom{a+b}{a} A^a B^b. \qquad \text{(B152)}$$

By again expanding the binomials and identifying rising factorials we can write this as a ${}_2F_1$ hypergeometric series,

$$\begin{aligned}
&\sum_{a=m+1}^{n} \sum_{b=0}^{\infty} \binom{n}{a}\binom{a+b}{m}\binom{a+b}{a} A^a B^b. \\
&= \sum_{a=m+1}^{n} \sum_{b=0}^{\infty} \frac{n!}{a!(n-a)!} \frac{(a+b)!}{m!(a+b-m)!} \frac{(a+b)!}{a!b!} A^a B^b \qquad \text{(B153)} \\
&= \sum_{a=m+1}^{n} \frac{n!}{(a-m)!m!} A^a \sum_{b=0}^{\infty} \frac{1}{b!} \frac{(a+b)!}{a!} \frac{(a+b)!}{a!} \frac{(a-m)!}{(a-m+b)!} B^b \qquad \text{(B154)} \\
&= \sum_{a=m+1}^{n} \frac{n!}{(a-m)!m!} A^a \sum_{b=0}^{\infty} \frac{1}{b!} \frac{(a+1)^{(b)}(a+1)^{(b)}}{(a-m+1)^{(b)}} B^b, \qquad \text{(B155)} \\
&= \sum_{a=m+1}^{n} \frac{n!}{(a-m)!m!} A^a \, {}_2F_1\left(\begin{matrix} a+1, a+1 \\ a-m+1 \end{matrix}; B\right). \qquad \text{(B156)}
\end{aligned}$$

Again, applying the Euler transformation allows us to rewrite the upper parameters as negative integers, which ensures the series terminates, as

$$\sum_{a=m+1}^{n} \frac{n!}{(a-m)!m!} A^a (1-B)^{-(m+a+1)} \, {}_2F_1\left(\begin{matrix}-m,-m\\ a-m+1\end{matrix};B\right) \tag{B157}$$

$$= \sum_{a=m+1}^{n} \frac{n!}{(a-m)!m!} \frac{A^a}{(1-B)^{m+a+1}} \sum_{b=0}^{m} \frac{1}{b!} \frac{(m)_{(b)}(m)_{(b)}}{(a-m+1)^{(b)}} B^b \tag{B158}$$

$$= \sum_{a=m+1}^{n} \sum_{b=0}^{m} \frac{n!m!}{b!(m-b)!(m-b)!(a-m+b)!(n-a)!} \frac{A^a B^b}{(1-B)^{m+a+1}}. \tag{B159}$$

Again, reversing the order of the inner summation by taking the index transform $b \to m-b$ allows us to write this as

$$\sum_{a=m+1}^{n} \sum_{b=0}^{m} \binom{n}{a}\binom{a}{b}\binom{m}{b} \frac{A^a B^{m-b}}{(1-B)^{m+a+1}}. \tag{B160}$$

Finally, we note that by definition, $\binom{m}{b} = 0 \, \forall \, b > m$. Hence, for $a > m$ we can set the upper limit of the inner summation to a to match equation (B151), with the additional terms from $b : m+1 \to a$ all vanishing. We finally write this component, then, as

$$\sum_{a=m+1}^{n} \sum_{b=0}^{a} \binom{n}{a}\binom{a}{b}\binom{m}{b} \frac{A^a B^{m-b}}{(1-B)^{m+a+1}}. \tag{B161}$$

COLLECTIVELY Returning to equation (B151), we can recombine the two outer summations as

$$\sum_{a=0}^{\min(m,n)} \sum_{b=m-a}^{\infty} \binom{n}{a}\binom{a+b}{m}\binom{a+b}{a} A^a B^b + \sum_{a=m+1}^{n} \sum_{b=0}^{\infty} \binom{n}{a}\binom{a+b}{m}\binom{a+b}{a} A^a B^b \tag{B162}$$

$$= \sum_{a=0}^{\min(m,n)} \sum_{b=0}^{a} \binom{n}{a}\binom{a}{b}\binom{m}{b} \frac{A^a B^{m-b}}{(1-B)^{m+a+1}} + \sum_{a=m+1}^{n} \sum_{b=0}^{a} \binom{n}{a}\binom{a}{b}\binom{m}{b} \frac{A^a B^{m-b}}{(1-B)^{m+a+1}} \tag{B163}$$

$$= \sum_{a=0}^{n} \sum_{b=0}^{a} \binom{n}{a}\binom{a}{b}\binom{m}{b} \frac{A^a B^{m-b}}{(1-B)^{m+a+1}}. \tag{B164}$$

Let us now swap the order of the two summations, as

$$\sum_{a=0}^{n} \sum_{b=0}^{a} \to \sum_{b=0}^{n} \sum_{a=b}^{n}, \tag{B165}$$

and reindex $a \to a + b$ to get

$$\sum_{b=0}^{n}\sum_{a=0}^{n-b}\binom{n}{a+b}\binom{a+b}{b}\binom{m}{b}\frac{A^{a+b}B^{m-b}}{(1-B)^{m+a+b+1}} \tag{B166}$$

$$=\sum_{b=0}^{n}\binom{n}{b}\binom{m}{b}\frac{A^{b}B^{m-b}}{(1-B)^{m+b+1}}\sum_{a=0}^{n-b}\binom{n-b}{a}\Big(\frac{A}{(1-B)}\Big)^{a} \tag{B167}$$

$$=\sum_{b=0}^{n}\binom{n}{b}\binom{m}{b}\frac{A^{b}B^{m-b}}{(1-B)^{m+b+1}}\Big(1+\frac{A}{(1-B)}\Big)^{n-b} \tag{B168}$$

$$=\frac{B^{m}}{(1-B)^{n+m+1}}\sum_{b=0}^{n}\binom{n}{b}\binom{m}{b}\Big(\frac{A}{B}\Big)^{b}(A+1-B)^{n-b}. \tag{B169}$$

Noticing that $\binom{\min(n,m)}{b} = 0$ for $b > \min(m, n)$, allowing us to short-circuit the summation, completes the proof. □

B.5.2.2 *Theorem*

Corollary B.11. *In the limited case in which the integration is performed over all but one x, p variable termed $x_{\mathcal{J}}$, $p_{\mathcal{J}}$, and in which the x and p components conform to the symmetry conditions*

$$V_{x\mathcal{J}} = V_{p\mathcal{J}} \qquad := V_{\mathcal{J}} \tag{B170}$$

$$\boldsymbol{\lambda}_{x\mathcal{I}}^{T}(V_{x\mathcal{I}}/V_{x\mathcal{J}})\boldsymbol{\lambda}_{x\mathcal{I}} = \boldsymbol{\lambda}_{p\mathcal{I}}^{T}(V_{p\mathcal{I}}/V_{p\mathcal{J}})\boldsymbol{\lambda}_{p\mathcal{I}} \qquad := G \tag{B171}$$

$$(\boldsymbol{\lambda}_{x\mathcal{J}} + V_{x\mathcal{J}}^{-1}V_{x\mathcal{J}\mathcal{I}}\boldsymbol{\lambda}_{x\mathcal{I}})^{2} = (\boldsymbol{\lambda}_{p\mathcal{J}} + V_{p\mathcal{J}}^{-1}V_{p\mathcal{J}\mathcal{I}}\boldsymbol{\lambda}_{p\mathcal{I}})^{2} := A^{2}, \tag{B172}$$

the integral

$$\int_{\mathbb{R}^{2(N-1)}} d^{N-1}x_{\mathcal{I}}\, d^{N-1}p_{\mathcal{I}}\, L_n\Big([\boldsymbol{\lambda}_x\cdot\boldsymbol{x}]^2 + [\boldsymbol{\lambda}_p\cdot\boldsymbol{p}]^2\Big)\exp(-\boldsymbol{q}^{T}V^{-1}\boldsymbol{q}), \tag{B173}$$

is given by

$$\pi^{N-1}\frac{\sqrt{\det V}}{V_{\mathcal{J}}}(-1)^{n}(G-1)^{n}L_n\Big(\frac{A^2}{2(1-G)}\big[2x_{\mathcal{J}}^2 + 2p_{\mathcal{J}}^2\big]\Big)\exp\Big(-\frac{1}{V}[x_{\mathcal{J}}^2 + p_{\mathcal{J}}^2]\Big), \tag{B174}$$

or by the strongly convergent infinite series

$$\pi^{N'}(-1)^{n}\sqrt{\det V}\sum_{m=0}^{\infty}(-1)^{m}C_{nm}\,L_m\Big(2x_{\mathcal{J}}^2 + 2p_{\mathcal{J}}^2\Big)\exp\Big(-x_{\mathcal{J}}^2 - p_{\mathcal{J}}^2\Big) \tag{B175}$$

for

$$C_{nm} = \frac{2}{V_{\mathcal{J}}+1}\sum_{b=0}^{\min(m,n)}\binom{n}{b}\binom{m}{b}\Big(\frac{\sqrt{2}A}{1+1/V_{\mathcal{J}}}\Big)^{2b}\Big(\frac{V_{\mathcal{J}}-1}{V_{\mathcal{J}}+1}\Big)^{m-b}\Big(\frac{A^2}{1+1/V_{\mathcal{J}}}+G-1\Big)^{n-b}. \tag{B176}$$

Proof.

FIRST RESULT The first result is a direct corollary of corollary B.9 where $\boldsymbol{A}_x \mapsto A, \boldsymbol{A}_p \mapsto A, \boldsymbol{q}_{\mathcal{J}}^T V_{\mathcal{J}}^{-1} \boldsymbol{q}_{\mathcal{J}}$ is expanded to x and p terms with equal $V_{\mathcal{J}}$.

SECOND RESULT Let us start with the more general result of corollary B.9 that

$$\int_{\mathbb{R}^{2(N-1)}} d^{N-1}x_{\mathcal{I}}\, d^{N-1}p_{\mathcal{I}}\, L_n\Big([\boldsymbol{\lambda}_x \cdot \boldsymbol{x}]^2 + [\boldsymbol{\lambda}_p \cdot \boldsymbol{p}]^2\Big) \exp(-\boldsymbol{q}^T V^{-1} \boldsymbol{q})$$
$$= \pi^{N-1} \sqrt{\frac{\det V}{\det V_{\mathcal{J}}}} (-1)^n (G-1)^n L_n\Big(\frac{1}{2(1-G)}\big[2(\boldsymbol{A}_x^T \boldsymbol{x}_{\mathcal{J}})^2 + 2(\boldsymbol{A}_p^T \boldsymbol{p}_{\mathcal{J}})^2\big]\Big) \exp(-\boldsymbol{q}_{\mathcal{J}}^T V_{\mathcal{J}}^{-1} \boldsymbol{q}_{\mathcal{J}}), \tag{B177}$$

for $G_x = G_p =: G$ and where $G, \boldsymbol{A}$ are as defined in that result.

We now consider the case where the integration is performed over all but one x, p index. Let us denote the surviving indices $x_{\mathcal{J}}$ and $p_{\mathcal{J}}$. We further impose that $V_{x\mathcal{J}} = V_{p\mathcal{J}} =: V_{\mathcal{J}}$ and $A_x^2 = A_p^2 =: A^2$, noting that $x_{\mathcal{J}}$, $p_{\mathcal{J}}$, and A are now of length one and treated as scalars. This leaves,

$$\pi^{N-1} \frac{\sqrt{\det V}}{V_{\mathcal{J}}} (-1)^n (G-1)^n L_n\Big[\frac{A^2}{2(1-G)}\big(2x_{\mathcal{J}}^2 + 2p_{\mathcal{J}}^2\big)\Big] \exp\Big[-\frac{1}{V_{\mathcal{J}}}(x_{\mathcal{J}}^2 + p_{\mathcal{J}}^2)\Big] \tag{B178}$$
$$= \pi^{N-1} \frac{\sqrt{\det V}}{V_{\mathcal{J}}} (-1)^n (G-1)^n L_n\Big[\frac{A^2}{2(1-G)}\big(2x_{\mathcal{J}}^2 + 2p_{\mathcal{J}}^2\big)\Big]$$
$$\times \exp(-x_{\mathcal{J}}^2 - p_{\mathcal{J}}^2) \exp\Big[\big(\frac{1}{1-V_{\mathcal{J}}}\big)(x_{\mathcal{J}}^2 + p_{\mathcal{J}}^2)\Big] \tag{B179}$$

Let us now rewrite the Laguerre polynomial and the second exponential as a summation of monomials as

$$\pi^{N-1} \frac{\sqrt{\det V}}{V_{\mathcal{J}}} (-1)^n (G-1)^n \left[\sum_{a=0}^{n} \binom{n}{a} \frac{(-1)^a}{a!} \Big(\frac{A^2}{2(1-G)}\Big)^a \big(2x_{\mathcal{J}}^2 + 2p_{\mathcal{J}}^2\big)^a\right]$$
$$\times \left[\sum_{b=0}^{\infty} \frac{1}{2^b b!} \Big(1 - \frac{1}{V_{\mathcal{J}}}\Big)^b \big(2x_{\mathcal{J}}^2 + 2p_{\mathcal{J}}^2\big)^b\right] \exp(-x_{\mathcal{J}}^2 - p_{\mathcal{J}}^2) \tag{B180}$$
$$= \pi^{N-1} \frac{\sqrt{\det V}}{V_{\mathcal{J}}} (-1)^n (G-1)^n$$
$$\times \left[\sum_{a=0}^{n} \sum_{b=0}^{\infty} \binom{n}{a} \frac{1}{a!b!} \Big(\frac{1}{2}\frac{A^2}{G-1}\Big)^a \Big(\frac{1}{2}(1 - 1/V_{\mathcal{J}})\Big)^b \big(2(x_{\mathcal{J}}^2 + p_{\mathcal{J}}^2)\big)^{a+b}\right] \exp(-x_{\mathcal{J}}^2 - p_{\mathcal{J}}^2). \tag{B181}$$

Let us now substitute the polynomial terms for their Laguerre representation using

$$x^n = n! \sum_{m=0}^{n} \binom{n}{m} (-1)^m L_m(x) \tag{B182}$$

to give

$$\pi^{N-1}\frac{\sqrt{\det V}}{V_{\mathcal{J}}}(-1)^n(G-1)^n\exp(-x_{\mathcal{J}}^2-p_{\mathcal{J}}^2)$$
$$\times\left\{\sum_{a=0}^{n}\sum_{b=0}^{\infty}\binom{n}{a}\frac{1}{a!b!}\Big(\frac{1}{2}\frac{A^2}{G-1}\Big)^a\Big(\frac{1-1/V_{\mathcal{J}}}{2}\Big)^b\left[(a+b)!\sum_{m=0}^{a+b}\binom{a+b}{m}(-1)^m L_m\Big(2(x_{\mathcal{J}}^2+p_{\mathcal{J}}^2)\Big)\right]\right\} \tag{B183}$$

$$=\pi^{N-1}\frac{\sqrt{\det V}}{V_{\mathcal{J}}}(-1)^n(G-1)^n\exp(-x_{\mathcal{J}}^2-p_{\mathcal{J}}^2).$$
$$\times\left[\sum_{a=0}^{n}\sum_{b=0}^{\infty}\sum_{m=0}^{a+b}\binom{n}{a}\binom{a+b}{m}\binom{a+b}{a}\Big(\frac{1}{2}\frac{A^2}{G-1}\Big)^a\Big(\frac{1-1/V_{\mathcal{J}}}{2}\Big)^b(-1)^m L_m\Big(2(x_{\mathcal{J}}^2+p_{\mathcal{J}}^2)\Big)\right] \tag{B184}$$

Swapping the order of the summations so that

$$\sum_{a=0}^{n}\sum_{b=0}^{\infty}\sum_{m=0}^{a+b}\to\sum_{m=0}^{\infty}\sum_{a=0}^{n}\sum_{b=\max(m-a,0)}^{\infty} \tag{B185}$$

leaves

$$\pi^{N-1}\frac{\sqrt{\det V}}{V_{\mathcal{J}}}(-1)^n\sum_{m=0}^{\infty}\left[(G-1)^n\sum_{a=0}^{n}\sum_{b=\max(m-a,0)}^{\infty}\binom{n}{a}\binom{a+b}{m}\binom{a+b}{a}\Big(\frac{1}{2}\frac{A^2}{G-1}\Big)^a\Big(\frac{1-1/V_{\mathcal{J}}}{2}\Big)^b\right]$$
$$\times\left[(-1)^m L_m\Big(2(x_{\mathcal{J}}^2+p_{\mathcal{J}}^2)\Big)\exp(-x_{\mathcal{J}}^2-p_{\mathcal{J}}^2)\right] \tag{B186}$$
$$=\pi^{N-1}(-1)^n\sqrt{\det V}\sum_{m=0}^{\infty}C_{nm}\left[(-1)^m L_m\Big(2(x_{\mathcal{J}}^2+p_{\mathcal{J}}^2)\Big)\exp(-x_{\mathcal{J}}^2-p_{\mathcal{J}}^2)\right], \tag{B187}$$

where we define

$$C_{nm}=\frac{(G-1)^n}{V_{\mathcal{J}}}\sum_{a=0}^{n}\sum_{b=\max(m-a,0)}^{\infty}\binom{n}{a}\binom{a+b}{m}\binom{a+b}{a}\Big(\frac{1}{2}\frac{A^2}{G-1}\Big)^a\Big(\frac{1-1/V_{\mathcal{J}}}{2}\Big)^b. \tag{B188}$$

We can here immediately use lemma B.10, that

$$\sum_{a=0}^{n}\sum_{b=\max(m-a,0)}^{\infty}\binom{n}{a}\binom{a+b}{m}\binom{a+b}{a}A^aB^b=\frac{1}{(1-B)^{n+m+1}}\sum_{b=0}^{\min(m,n)}\binom{n}{b}\binom{m}{b}A^bB^{m-b}(A+1-B)^{n-b}, \tag{B189}$$

for

$$A=\frac{1}{2}\frac{A^2}{G-1} \tag{B190}$$
$$B=\frac{1}{2}(1-1/V_{\mathcal{J}}). \tag{B191}$$

The infinite series then reduces to a closed form as

$$C_{nm} = \frac{(G-1)^n}{V_\mathcal{J}} \frac{1}{(\frac{1}{2} + \frac{1}{2V_\mathcal{J}})^{n+m+1}} \tag{B192}$$

$$\times \sum_{b=0}^{\min(m,n)} \binom{n}{b}\binom{m}{b} (\frac{1}{2}\frac{A^2}{G-1})^b (\frac{1}{2}(1-1/V_\mathcal{J}))^{m-b} (\frac{1}{2}\frac{A^2}{G-1} + \frac{1}{2} + \frac{1}{2V_\mathcal{J}})^{n-b},$$

$$= \frac{(G-1)^n}{V_\mathcal{J}} \frac{1}{(1+1/V_\mathcal{J})^{n+m+1}} \tag{B193}$$

$$\times \sum_{b=0}^{\min(m,n)} 2^{b+1} \binom{n}{b}\binom{m}{b} (\frac{A^2}{G-1})^b (1 - \frac{1}{V_\mathcal{J}})^{m-b} (\frac{A^2}{G-1} + 1 + \frac{1}{V_\mathcal{J}})^{n-b}$$

$$= \frac{1}{V_\mathcal{J}(1+1/V_\mathcal{J})^{n+m+1}} \sum_{b=0}^{\min(m,n)} 2^{b+1} \binom{n}{b}\binom{m}{b} A^{2b} (1 - \frac{1}{V_\mathcal{J}})^{m-b} (A^2 + (G-1)(1 + \frac{1}{V_\mathcal{J}}))^{n-b}. \tag{B194}$$

Subsuming the $(1 + 1/V_\mathcal{J})^{n+m}$ term into the relevant terms in the summation by introducing an additional $(1 + 1/V_\mathcal{J})^{2b}$ allows us to simplify the brackets as

$$C_{nm} = \frac{1}{V_\mathcal{J}(1+1/V_\mathcal{J})} \sum_{b=0}^{\min(m,n)} 2^{b+1} \binom{n}{b}\binom{m}{b} A^{2b} (\frac{1-1/V_\mathcal{J}}{1+1/V_\mathcal{J}})^{m-b} (\frac{A^2}{1+1/V_\mathcal{J}} + G - 1)^{n-b} (\frac{1}{1+1/V_\mathcal{J}})^{2b}$$

$$= \frac{2}{V_\mathcal{J}+1} \sum_{b=0}^{\min(m,n)} \binom{n}{b}\binom{m}{b} (\frac{\sqrt{2}A}{1+1/V_\mathcal{J}})^{2b} (\frac{V_\mathcal{J}-1}{V_\mathcal{J}+1})^{m-b} (\frac{A^2}{1+1/V_\mathcal{J}} + G - 1)^{n-b} \tag{B195}$$

which completes the proof. □

B.6 GAUSSIAN INTEGRALS OF A LAGUERRE POLYNOMIAL SUBJECT TO A COORDINATE TRANSFORM

Finally, let us now consider integrals of the form we expect from our output state: a Laguerre component and a Gaussian component with covariance matrix $I_2 \oplus V_r$ that have each been subject to a coordinate transform. Before we go on to solve this integral in theorem B.14, let us first consider two Lemmas that tell us what form the G and A variables will take.

B.6.1 *Lemmas*

Lemma B.12. *Consider an $n+1 \times n+1$ square real matrix Λ, given in block form by*

$$\Lambda = \begin{pmatrix} \Lambda_{\mathcal{J}} & \boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}} \\ \boldsymbol{\Lambda}_{\mathcal{I}\mathcal{J}} & \Lambda_{\mathcal{I}} \end{pmatrix}, \tag{B196}$$

where $\Lambda_{\mathcal{J}} \in \mathbb{R}^{1\times 1}$, $\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}} \in \mathbb{R}^{1\times n}$, $\boldsymbol{\Lambda}_{\mathcal{I}\mathcal{J}} \in \mathbb{R}^{n\times 1}$, and $\boldsymbol{\Lambda}_{\mathcal{I}} \in \mathbb{R}^{n\times n}$.

Consider also a second matrix given by

$$V = \Lambda^{-1}(1 \oplus \boldsymbol{V}_r){\Lambda^{-1}}^T := \begin{pmatrix} V_{\mathcal{J}} & V_{\mathcal{J}\mathcal{I}} \\ V_{\mathcal{I}\mathcal{J}} & V_{\mathcal{I}} \end{pmatrix}. \tag{B197}$$

Then,

$$G := 2\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}} \cdot (V/V_{\mathcal{J}}) \cdot \boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}^T = 2 - \frac{2}{\gamma} \in \mathbb{R}, \tag{B198}$$

$$A := \sqrt{2}(\Lambda_{\mathcal{J}} + \frac{V_{\mathcal{J}\mathcal{I}}\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}^T}{V_{\mathcal{J}}}) = \frac{\sqrt{2}}{\gamma}(\Lambda/\Lambda_{\mathcal{I}}) \in \mathbb{R} \tag{B199}$$

for

$$\gamma = 1 + \boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}(\Lambda_{\mathcal{I}}^T V_r^{-1} \Lambda_{\mathcal{I}})^{-1} \boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}^T = 1 + (\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1}) V_r (\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1})^T \in \mathbb{R}, \tag{B200}$$

where (B/C) denotes the Schur complement of B with respect to block C.

Proof. We will prove here of these results separately. Let us first, though, derive the form of $V/V_{\mathcal{J}}$.

FORM OF $V/V_{\mathcal{J}}$ To derive the form of the Schur complement, let us note the two ways we can write the inverse matrix V^{-1}. First, directly from the form of equation (B197) as

$$V^{-1} = \Lambda^T(1/\upsilon_\psi \oplus \boldsymbol{V}_r^{-1})\Lambda = \begin{pmatrix} \Lambda_{\mathcal{J}}^2 + \boldsymbol{\Lambda}_{\mathcal{I}\mathcal{J}}^T V_R^{-1} \boldsymbol{\Lambda}_{\mathcal{I}\mathcal{J}} & \Lambda_{\mathcal{J}}\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}} + \boldsymbol{\Lambda}_{\mathcal{I}\mathcal{J}}^T V_R^{-1} \Lambda_{\mathcal{I}} \\ \Lambda_{\mathcal{J}}\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}^T + \Lambda_{\mathcal{I}}^T V_R^{-1} \boldsymbol{\Lambda}_{\mathcal{I}\mathcal{J}} & \boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}^T\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}} + \Lambda_{\mathcal{I}}^T V_R^{-1} \Lambda_{\mathcal{I}} \end{pmatrix}, \tag{B201}$$

and second using blockwise inversion as

$$V^{-1} = \begin{pmatrix} (V/V_{\mathcal{I}})^{-1} & -(V_{\mathcal{J}})^{-1}V_{\mathcal{J}\mathcal{I}}(V/V_{\mathcal{J}})^{-1} \\ -(V/V_{\mathcal{J}})^{-1}V_{\mathcal{I}\mathcal{J}}(V_{\mathcal{J}})^{-1} & (V/V_{\mathcal{J}})^{-1} \end{pmatrix}. \tag{B202}$$

These allow us to read off the form of the Schur complement of V with respect to block $V_{\mathcal{J}}$ as

$$V/V_{\mathcal{J}} = [(V^{-1})_{\mathcal{I}}]^{-1} = [\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}^T\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}} + \Lambda_{\mathcal{I}}^T V_R^{-1}\Lambda_{\mathcal{I}}]^{-1}. \tag{B203}$$

The Sherman-Morrison formula for a square invertible matrix M and column vectors $\boldsymbol{u}$, $\boldsymbol{v}$ states that

$$(M + \boldsymbol{u}^T\boldsymbol{v})^{-1} = M^{-1} - \frac{M^{-1}\boldsymbol{u}^T\boldsymbol{v}M^{-1}}{1 + \boldsymbol{v}M^{-1}\boldsymbol{u}^T}. \tag{B204}$$

Applying this formula to equation (B203) we can rewrite the Schur complement as

$$V/V_{\mathcal{J}} = (\Lambda_{\mathcal{I}}^T V_R^{-1}\Lambda_{\mathcal{I}})^{-1} - \frac{(\Lambda_{\mathcal{I}}^T V_R^{-1}\Lambda_{\mathcal{I}})^{-1}(\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}^T\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}})(\Lambda_{\mathcal{I}}^T V_R^{-1}\Lambda_{\mathcal{I}})^{-1}}{1 + \boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}(\Lambda_{\mathcal{I}}^T V_R^{-1}\Lambda_{\mathcal{I}})^{-1}\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}^T} \tag{B205}$$

$$= \Lambda_{\mathcal{I}}^{-1}V_R(\Lambda_{\mathcal{I}}^{-1})^T - \frac{\Lambda_{\mathcal{I}}^{-1}V_R(\Lambda_{\mathcal{I}}^{-1})^T\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}^T\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1}V_R(\Lambda_{\mathcal{I}}^{-1})^T}{1 + \boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1}V_R(\Lambda_{\mathcal{I}}^{-1})^T\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}^T}. \tag{B206}$$

FORM OF G Let us now consider to the first result in this Lemma, shown in equation (B198), that

$$G := 2\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}} \cdot (V/V_{\mathcal{J}}) \cdot \boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}^T. \tag{B207}$$

Substituting in the form of the Schur complement found above, this becomes

$$2\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}} \cdot \left[\Lambda_{\mathcal{I}}^{-1}V_R(\Lambda_{\mathcal{I}}^{-1})^T - \frac{\Lambda_{\mathcal{I}}^{-1}V_R(\Lambda_{\mathcal{I}}^{-1})^T\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}^T\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1}V_R(\Lambda_{\mathcal{I}}^{-1})^T}{1 + \boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1}V_R(\Lambda_{\mathcal{I}}^{-1})^T\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}^T}\right] \cdot \boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}^T \tag{B208}$$

$$= 2(\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1})V_R(\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1})^T - 2\frac{(\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1})V_R(\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1})^T(\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1})V_R(\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1})^T}{1 + (\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1})V_R(\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1})^T} \tag{B209}$$

$$= 2(\gamma - 1) - 2\frac{(\gamma - 1)^2}{\gamma} \tag{B210}$$

$$= 2 - \frac{2}{\gamma}, \tag{B211}$$

for γ defined as

$$\gamma := 1 + (\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1})V_R(\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1})^T \in \mathbb{R}. \tag{B212}$$

FORM OF A Now consider the expression in equation (B199),

$$A := \sqrt{2}(\Lambda_{\mathcal{J}} + \frac{V_{\mathcal{J}\mathcal{I}}\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}^T}{V_{\mathcal{J}}}). \tag{B213}$$

We can use the expression for the inverse of V from equation (B202) to first rewrite $V_{\mathcal{J}\mathcal{I}}$ as

$$V_{\mathcal{J}\mathcal{I}} = -V_{\mathcal{J}}(V^{-1})_{\mathcal{J}\mathcal{I}}(V/V_{\mathcal{J}}). \tag{B214}$$

Consequently, A becomes

$$\sqrt{2}\Big[\Lambda_{\mathcal{J}} - (V^{-1})_{\mathcal{J}\mathcal{I}}(V/V_{\mathcal{J}})\boldsymbol{\Lambda}^T_{\mathcal{J}\mathcal{I}}\Big] \tag{B215}$$

$$=\sqrt{2}\Big[\Lambda_{\mathcal{J}} - (\Lambda_{\mathcal{J}}\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}} + \boldsymbol{\Lambda}^T_{\mathcal{I}\mathcal{J}}V_R^{-1}\Lambda_{\mathcal{I}})(V/V_{\mathcal{J}})\boldsymbol{\Lambda}^T_{\mathcal{J}\mathcal{I}}\Big] \tag{B216}$$

$$=\sqrt{2}\Big[\Lambda_{\mathcal{J}} - \Lambda_{\mathcal{J}}\big[\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}(V/V_{\mathcal{J}})\boldsymbol{\Lambda}^T_{\mathcal{J}\mathcal{I}}\big] - \boldsymbol{\Lambda}^T_{\mathcal{I}\mathcal{J}}V_R^{-1}\Lambda_{\mathcal{I}}(V/V_{\mathcal{J}})\boldsymbol{\Lambda}^T_{\mathcal{J}\mathcal{I}}\Big] \tag{B217}$$

$$=\sqrt{2}\Big[\frac{\Lambda_{\mathcal{J}}}{\gamma} - \boldsymbol{\Lambda}^T_{\mathcal{I}\mathcal{J}}V_R^{-1}\Lambda_{\mathcal{I}}(V/V_{\mathcal{J}})\boldsymbol{\Lambda}^T_{\mathcal{J}\mathcal{I}}\Big], \tag{B218}$$

where we have recognised the term within the inner square bracket as equal to $1/2G$. Let us now focus on the second term. Expanding out the Schur complement using equation (B206), cancelling the $V_R^{-1}\Lambda_{\mathcal{I}}$ matrices, and substituting in γ we can reduce this part to

$$\boldsymbol{\Lambda}^T_{\mathcal{I}\mathcal{J}}V_R^{-1}\Lambda_{\mathcal{I}}(V/V_{\mathcal{J}})\boldsymbol{\Lambda}^T_{\mathcal{J}\mathcal{I}} \tag{B219}$$

$$=\boldsymbol{\Lambda}^T_{\mathcal{I}\mathcal{J}}V_R^{-1}\Lambda_{\mathcal{I}}\left[\Lambda_{\mathcal{I}}^{-1}V_R(\Lambda_{\mathcal{I}}^{-1})^T\left(1 - \boldsymbol{\Lambda}^T_{\mathcal{J}\mathcal{I}}\frac{\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1}V_R(\Lambda_{\mathcal{I}}^{-1})^T}{1+\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1}V_R(\Lambda_{\mathcal{I}}^{-1})^T\boldsymbol{\Lambda}^T_{\mathcal{J}\mathcal{I}}}\right)\right]\boldsymbol{\Lambda}^T_{\mathcal{J}\mathcal{I}} \tag{B220}$$

$$=\boldsymbol{\Lambda}^T_{\mathcal{I}\mathcal{J}}(\Lambda_{\mathcal{I}}^{-1})^T\boldsymbol{\Lambda}^T_{\mathcal{J}\mathcal{I}}\left(1 - \frac{(\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1})V_R(\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1})^T}{1+(\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1})V_R(\boldsymbol{\Lambda}_{\mathcal{J}\mathcal{I}}\Lambda_{\mathcal{I}}^{-1})^T}\right) \tag{B221}$$

$$=\boldsymbol{\Lambda}^T_{\mathcal{I}\mathcal{J}}(\Lambda_{\mathcal{I}}^{-1})^T\boldsymbol{\Lambda}^T_{\mathcal{J}\mathcal{I}}\frac{1}{\gamma}. \tag{B222}$$

Equation (B218) can then be written

$$\sqrt{2}\Big[\frac{\Lambda_{\mathcal{J}}}{\gamma} - \boldsymbol{\Lambda}^T_{\mathcal{I}\mathcal{J}}V_R^{-1}\Lambda_{\mathcal{I}}(V/V_{\mathcal{J}})\boldsymbol{\Lambda}^T_{\mathcal{J}\mathcal{I}}\Big] = \sqrt{2}\frac{\Lambda_{\mathcal{J}} - \boldsymbol{\Lambda}^T_{\mathcal{I}\mathcal{J}}(\Lambda_{\mathcal{I}}^{-1})^T\boldsymbol{\Lambda}^T_{\mathcal{J}\mathcal{I}}}{\gamma}, \tag{B223}$$

which we can recognise as the Schur complement $\Lambda/\Lambda_{\mathcal{I}} = \frac{\det(\Lambda)}{\det(\Lambda_{\mathcal{I}})}$ (noting $\Lambda/\Lambda_{\mathcal{I}}$ is scalar), completing the proof. □

Lemma B.13. *Consider a symplectic matrix Λ given in block form as*

$$\Lambda = \begin{pmatrix} \Lambda_{x\mathcal{J}} & \Lambda_{x\mathcal{J}\mathcal{I}} & & \\ \Lambda_{x\mathcal{I}\mathcal{J}} & \Lambda_{x\mathcal{I}} & & \\ & & \Lambda_{p\mathcal{J}} & \Lambda_{p\mathcal{J}\mathcal{I}} \\ & & \Lambda_{p\mathcal{I}\mathcal{J}} & \Lambda_{p\mathcal{I}} \end{pmatrix}, \tag{B224}$$

with zeros in the empty off-diagonal blocks, and a second matrix given by

$$V = \Lambda^{-1}(1 \oplus \boldsymbol{V}_r)\Lambda^{-1^T} := \begin{pmatrix} V_{\mathcal{J}} & V_{\mathcal{J}\mathcal{I}} \\ V_{\mathcal{I}\mathcal{J}} & V_{\mathcal{I}} \end{pmatrix}. \tag{B225}$$

When the conditions

$$\Lambda_{x\mathcal{J}} = \Lambda_{p\mathcal{J}} := \Lambda_{\mathcal{J}}, \tag{B226}$$

$$V_{x\mathcal{J}} = V_{p\mathcal{J}} := V_{\mathcal{J}} \tag{B227}$$

are met, then $\gamma_x = \gamma_p := \gamma$ *is automatically satisfied for* γ *as given in lemma B.12, which becomes*

$$\gamma = \frac{V_{\mathcal{J}}}{\Lambda_{\mathcal{J}}^2}. \tag{B228}$$

Consequently, the expressions for G, A from lemma B.12 reduce to

$$G_{x/p} = 2\boldsymbol{\Lambda}_{(x/p)JI} \cdot [V_{(x/p)}/V_{(x/p)J}] \cdot \boldsymbol{\Lambda}^T_{(x/p)JI} = 2 - 2\frac{\Lambda_{\mathcal{J}}^2}{V_{\mathcal{J}}}, \tag{B229}$$

$$A_{x/p} = \sqrt{2}\big[\Lambda_{(x/p)J} + \frac{V_{(x/p)JI}\boldsymbol{\Lambda}^T_{(x/p)JI}}{V_{(x/p)J}}\big] = \sqrt{2}\frac{\Lambda_{\mathcal{J}}}{V_{\mathcal{J}}}, \tag{B230}$$

and $G_x = G_p$, $A_x = A_p$ *are automatically satisfied.*

Proof.

FORM OF γ Let us first note the property of symplectic matrices with zero off-diagonal block that

$$(\Lambda_x)^{-1} = (\Lambda_p)^T. \tag{B231}$$

Comparing this to the form of $(\Lambda_x)^{-1}$ found through blockwise matrix inversion,

$$\Lambda_x^{-1} = \begin{pmatrix} (\Lambda_x/\Lambda_{x\mathcal{I}})^{-1} & -(\Lambda_x/\Lambda_{x\mathcal{I}})^{-1}\Lambda_{x\mathcal{J}\mathcal{I}}(\Lambda_{x\mathcal{I}})^{-1} \\ -(\Lambda_{x\mathcal{I}})^{-1}\Lambda_{x\mathcal{J}\mathcal{I}}(\Lambda_x/\Lambda_{x\mathcal{I}})^{-1} & (\Lambda_x/\Lambda_{x\mathcal{J}})^{-1} \end{pmatrix}, \tag{B232}$$

we can read off the form of $\Lambda_{\mathcal{J}}$ as

$$\Lambda_{\mathcal{J}} = (\Lambda_p^T)_{1,1} = (\Lambda_x/\Lambda_{x\mathcal{I}})^{-1}. \tag{B233}$$

Equating these forms as

$$\Lambda_x^{-1} = \begin{pmatrix} \Lambda_{\mathcal{J}} & -\Lambda_{\mathcal{J}}\Lambda_{x\mathcal{J}\mathcal{I}}(\Lambda_{x\mathcal{I}})^{-1} \\ -(\Lambda_{x\mathcal{I}})^{-1}\Lambda_{x\mathcal{I}\mathcal{J}}\Lambda_{\mathcal{J}} & (\Lambda_x/\Lambda_{x\mathcal{J}})^{-1} \end{pmatrix} = \begin{pmatrix} \Lambda_{\mathcal{J}} & \Lambda^T_{p\mathcal{I}\mathcal{J}} \\ \Lambda_{p\mathcal{J}\mathcal{I}}T & \Lambda^T_{p\mathcal{I}} \end{pmatrix} = \Lambda_p^T, \tag{B234}$$

also allows us to read off the result that

$$\boldsymbol{\Lambda}_{x\mathcal{J}\mathcal{I}}\Lambda_{x\mathcal{I}}^{-1} = -\frac{\Lambda^T_{p\mathcal{I}\mathcal{J}}}{\Lambda_{\mathcal{J}}}. \tag{B235}$$

Now let us directly calculate the form of $V_\mathcal{J}$, the upper left element of V_x, as

$$V_x = \Lambda_x^{-1}(1 \oplus V_{rx})(\Lambda_x^{-1})^T \tag{B236}$$

$$= \begin{pmatrix} \Lambda_\mathcal{J} & \Lambda_{p\mathcal{I}\mathcal{J}}^T \\ \Lambda_{p\mathcal{J}\mathcal{I}}^T & \Lambda_{p\mathcal{I}}^T \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & V_{rx} \end{pmatrix} \begin{pmatrix} \Lambda_\mathcal{J} & \Lambda_{p\mathcal{J}\mathcal{I}} \\ \Lambda_{p\mathcal{I}\mathcal{J}} & \Lambda_{p\mathcal{I}} \end{pmatrix} \tag{B237}$$

$$= \begin{pmatrix} (\Lambda_\mathcal{J})^2 + \Lambda_{p\mathcal{I}\mathcal{J}}^T \cdot V_{rx} \cdot \Lambda_{p\mathcal{I}\mathcal{J}} & \cdots \\ \cdots & \cdots \end{pmatrix}, \tag{B238}$$

so

$$V_\mathcal{J} = (\Lambda_\mathcal{J})^2 + \Lambda_{p\mathcal{I}\mathcal{J}}^T \cdot V_{rx} \cdot \Lambda_{p\mathcal{I}\mathcal{J}}. \tag{B239}$$

These results in equations (B235) and (B239) allow us to rewrite the form of γ_x as

$$\gamma_x := 1 + (\boldsymbol{\Lambda}_{x\mathcal{J}\mathcal{I}}\Lambda_{x\mathcal{I}}^{-1})V_{rx}(\boldsymbol{\Lambda}_{x\mathcal{J}\mathcal{I}}\Lambda_{x\mathcal{I}}^{-1})^T \tag{B240}$$

$$= 1 + \frac{1}{\Lambda_\mathcal{J}^2}\Lambda_{p\mathcal{I}\mathcal{J}}^T V_{rx}\Lambda_{p\mathcal{I}\mathcal{J}} \tag{B241}$$

$$= 1 + \frac{V_\mathcal{J} - \Lambda_\mathcal{J}^2}{\Lambda_\mathcal{J}^2} \tag{B242}$$

$$= \frac{V_\mathcal{J}}{\Lambda_\mathcal{J}^2}. \tag{B243}$$

An identical process can be followed to find γ_p, which takes the same final form.

FORMS OF G, A Recall from lemma B.12 that G and A are defined as

$$G_x := 2 - \frac{2}{\gamma_x}, \tag{B244}$$

$$A_x := \sqrt{2}\frac{1}{\gamma_x}(\Lambda_x/\Lambda_{x\mathcal{I}}). \tag{B245}$$

Substituting the form of γ from equation (B243) and $(\Lambda_x/\Lambda_{x\mathcal{I}}) = 1/\Lambda_\mathcal{J}$ from equation (B233), we can immediately write these as

$$G_x = 2 - 2\frac{\Lambda_\mathcal{J}^2}{V_\mathcal{J}}, \tag{B246}$$

$$A_x = \sqrt{2}\frac{\Lambda_\mathcal{J}}{V_\mathcal{J}}. \tag{B247}$$

The form of G_p and A_p can found identically. □

B.6.2 *Theorem*

Theorem B.14. *Consider a symplectic coordinate transform Λ acting on a symmetric positive-definite matrix $I_2 \oplus V_r$ where V_r and Λ are such that the transformed matrix can be written in separate x/p components as*

$$V_x = \Lambda_x^{-1}(1 \oplus V_{rx})(\Lambda_x^{-1})^T \qquad V_p = \Lambda_p^{-1}(1 \oplus V_{rp})(\Lambda_p^{-1})^T.$$

Let us for simplicity denote as $\boldsymbol{\lambda}_x, \boldsymbol{\lambda}_p$ the first row of Λ_x, Λ_p. Then, when the symmetry condition

$$(\Lambda_x)_{1,1} = (\Lambda_p)_{1,1} =: \Lambda_{\mathcal{J}} \qquad (V_x)_{1,1} = (V_p)_{1,1} =: V_{\mathcal{J}} \tag{B248}$$

is satisfied, the integral

$$\int_{\mathbb{R}^{2N-2}} dx_{2,\dots,N} dp_{2,\dots,N} L_n\Big(2[\boldsymbol{\lambda}_x \cdot \boldsymbol{x}]^2 + 2[\boldsymbol{\lambda}_p \cdot \boldsymbol{p}]^2\Big) \exp(-\boldsymbol{x}^T V_x^{-1} \boldsymbol{x} - \boldsymbol{p}^T V_p^{-1} \boldsymbol{p}), \tag{B249}$$

is given by the closed form

$$\pi^{N-1} \frac{\sqrt{\det V}}{V_{\mathcal{J}}^{n+1}} (2\Lambda_{\mathcal{J}}^2 - V_{\mathcal{J}})^n L_n\Big(\frac{\Lambda_{\mathcal{J}}^2}{V_{\mathcal{J}}(2\Lambda_{\mathcal{J}}^2 - V_{\mathcal{J}})}\big[2x_1^2 + 2p_1^2\big]\Big) \exp\Big(-\frac{1}{V_{\mathcal{J}}}(x_1^2 + p_1^2)\Big), \tag{B250}$$

or by the infinite series

$$\pi^{N-1}(-1)^n \sqrt{\det V} \sum_{m=0}^{\infty} C_m (-1)^m L_m\Big(2(x_1^2 + 2p_1^2)\Big) \exp\Big(-x_1^2 - p_1^2\Big) \tag{B251}$$

for

$$C_m = \frac{2}{(V_{\mathcal{J}} + 1)^{n+m+1}} \sum_{b=0}^{\min(m,n)} \binom{n}{b}\binom{m}{b} (4\Lambda_{\mathcal{J}}^2)^b (V_{\mathcal{J}} - 1)^{m-b} (V_{\mathcal{J}} + 1 - 2\Lambda_{\mathcal{J}}^2)^{n-b}. \tag{B252}$$

Proof. This result is a special case of corollary B.11 for

$$V \to \Lambda^{-1}(I_2 \oplus V_r)(\Lambda^{-1})^T \tag{B253}$$

$$\lambda_x \to \sqrt{2}\lambda_x \tag{B254}$$

$$\lambda_p \to \sqrt{2}\lambda_p. \tag{B255}$$

Under the conditions outlined in equation (B248), lemma B.13 shows that the G, A variables in corollary B.11 become

$$G = 2 - 2\frac{\Lambda_{\mathcal{J}}^2}{V_{\mathcal{J}}}, \tag{B256}$$

$$A = \sqrt{2}\frac{\Lambda_{\mathcal{J}}}{V_{\mathcal{J}}}, \tag{B257}$$

and $G_x = G_p$ and $A_x = A_p$ are automatically satisfied.

RESULT 1 Let us consider the first result from corollary B.11, that the solution to the integral can be written

$$\pi^{N-1}\frac{\sqrt{\det V}}{V_\mathcal{J}}(1-G)^n L_n\Big(\frac{A^2}{2(1-G)}\big[2x_\mathcal{J}^2+2p_\mathcal{J}^2\big]\Big)\exp\Big(-\frac{1}{V_\mathcal{J}}(x_\mathcal{J}^2+p_\mathcal{J}^2)\Big), \tag{B258}$$

which in this case becomes

$$\pi^{N-1}\frac{\sqrt{\det V}}{V_\mathcal{J}}(2\frac{\Lambda_\mathcal{J}^2}{V_\mathcal{J}}-1)^n L_n\Big(\frac{1}{2V_\mathcal{J}-\frac{V_\mathcal{J}^2}{\Lambda_\mathcal{J}^2}}\big[2x_\mathcal{J}^2+2p_\mathcal{J}^2\big]\Big)\exp\Big(-\frac{1}{V_\mathcal{J}}(x_\mathcal{J}^2+p_\mathcal{J}^2)\Big) \tag{B259}$$

$$=\pi^{N-1}\frac{\sqrt{\det V}}{V_\mathcal{J}^{n+1}}(2\Lambda_\mathcal{J}^2-V_\mathcal{J})^n L_n\Big(\frac{\Lambda_\mathcal{J}^2}{V_\mathcal{J}(2\Lambda_\mathcal{J}^2-V_\mathcal{J})}\big[2x_\mathcal{J}^2+2p_\mathcal{J}^2\big]\Big)\exp\Big(-\frac{1}{V_\mathcal{J}}(x_\mathcal{J}^2+p_\mathcal{J}^2)\Big). \tag{B260}$$

RESULT 2 Alternatively, the integral output can be written as the infinite series

$$\pi^{N'}(-1)^n\sqrt{\det V}\sum_{m=0}^{\infty}(-1)^m C_{nm}\,L_m\Big(2x_\mathcal{J}^2+2p_\mathcal{J}^2\Big)\exp\Big(-x_\mathcal{J}^2-p_\mathcal{J}^2\Big) \tag{B261}$$

with

$$C_{nm}=\frac{2}{V_\mathcal{J}+1}\sum_{b=0}^{\min(m,n)}\binom{n}{b}\binom{m}{b}(\frac{\sqrt{2}A}{1+1/V_\mathcal{J}})^{2b}(\frac{V_\mathcal{J}-1}{V_\mathcal{J}+1})^{m-b}(\frac{A^2}{1+1/V_\mathcal{J}}+G-1)^{n-b}. \tag{B262}$$

In this case, for G, A as given by lemma B.13, the expression for the coefficients becomes

$$C_{nm}=\frac{2}{V_\mathcal{J}+1}\sum_{b=0}^{\min(m,n)}\binom{n}{b}\binom{m}{b}(\frac{2\Lambda_\mathcal{J}/V_\mathcal{J}}{1+1/V_\mathcal{J}})^{2b}(\frac{V_\mathcal{J}-1}{V_\mathcal{J}+1})^{m-b}(\frac{2\Lambda_\mathcal{J}^2/V_\mathcal{J}^2}{1+1/V_\mathcal{J}}+1-2\frac{\Lambda_\mathcal{J}^2}{V_\mathcal{J}})^{n-b} \tag{B263}$$

$$=\frac{2}{(V_\mathcal{J}+1)^{n+m+1}}\sum_{b=0}^{\min(m,n)}\binom{n}{b}\binom{m}{b}(4\Lambda_\mathcal{J}^2)^b(V_\mathcal{J}-1)^{m-b}(V_\mathcal{J}+1-2\Lambda_\mathcal{J}^2)^{n-b}. \tag{B264}$$

□

BIBLIOGRAPHY

[1] *Strategic research agenda of the quantum flagship*, Policy Paper (European Quantum Flagship, Feb. 2020).

[2] J. Carson, 'The Next Quantum Revolution', Massachusetts Institute of Technology Spectrum.

[3] *Our Prototype Network*, Quantum Internet Alliance, https://quantuminternetalliance.org/our-prototype-network/.

[4] B. Lackey, *Quantum networking: A roadmap to a quantum internet*, Microsoft Azure, (1st Nov. 2023) https://azure.microsoft.com/en-us/blog/quantum/2023/11/01/quantum-networking-a-roadmap-to-a-quantum-internet/.

[5] D. Castelvecchi, 'The quantum internet has arrived (and it hasn't)', Nature 554, 289–292 (2018).

[6] *Plug and play continuous variable quantum key distribution for metropolitan networks*, European Quantum Flagship, (13th June 2025) https://qt.eu/news/2020/plug-and-play-continuous-variable-quantum-key-distribution-for-metropolitan-networks (visited on 13/06/2025).

[7] *Quantum security technologies*, Policy Paper (National Cyber Security Centre, 24th Mar. 2020).

[8] *Underground leak detection in the century of quantum technology*, NSW Smart Sensing Network, (18th Aug. 2020) https://www.nssn.org.au/news/2020/8/18/underground-leak-detection-in-the-century-of-quantum-technology (visited on 29/05/2025).

[9] L. Brooke-Holland, *AUKUS pillar 2: Advanced military capabilities*, Research Briefing (House of Commons Library, 2nd Sept. 2024).

[10] C. Vallance, 'Unjammable navigation tech gets first airborne test', BBC News (2024).

[11] *Quantum science reveals Lisbon's history at EQTC 2024*, European Quantum Flagship, (12th Nov. 2024) https://qt.eu/news/2024/2024-11-12_Quantum-science-reveals-Lisbons-history-at-global-tech-event (visited on 13/06/2025).

[12] *National Quantum Strategy*, Policy Paper (Commonwealth of Australia, Department of Industry, Science and Resources, 3rd May 2023).

[13] M. Shams, J. Choudhari, K. Reyes, S. Prentzas, A. Gapizov, A. Shehryar, M. Affaf, H. Grezenko, R. W. Gasim, S. N. Mohsin, A. Rehman and S. Rehman, 'The Quantum-Medical Nexus: Understanding the Impact of Quantum Technologies on Healthcare', Cureus, 10.7759/cureus.48077 (2023).

[14] *National Quantum Strategy*, Policy Paper (United Kingdom, Department for Science, Innovation and Technology, 15th Mar. 2023).

[15] *Quantum Manifesto*, Policy Paper (European Union, May 2016).

[16] U. Leonhardt, *Essential quantum optics: from quantum measurements to black holes* (Cambridge University Press, Cambridge, 2010), ISBN: 978-0-521-86978-2.

[17] R. W. Robinett, *Quantum mechanics: classical results, modern systems, and visualized examples*, 2nd ed (Oxford University Press, Oxford ; New York, 2006), ISBN: 978-0-19-853097-8.

[18] D. K. Ferry and M. Nedjalkov, *The Wigner function in science and technology*, Version: 20181101, IOP Expanding Physics (IOP Publishing, Bristol, UK, 2018), ISBN: 978-0-7503-1671-2.

[19] A. Einstein, B. Podolsky and N. Rosen, 'Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?', Physical Review 47, 777–780 (1935).

[20] G. Vidal and R. F. Werner, 'Computable measure of entanglement', Physical Review A 65, 032314 (2002).

[21] Q. Y. He, Q. H. Gong and M. D. Reid, 'Classifying Directional Gaussian Entanglement, Einstein-Podolsky-Rosen Steering, and Discord', Physical Review Letters 114, 060402 (2015).

[22] C. Weedbrook, N. B. Grosse, T. Symul, P. K. Lam and T. C. Ralph, 'Quantum cloning of continuous-variable entangled states', Physical Review A 77, 052313 (2008).

[23] A. Serafini, *Quantum continuous variables: a primer of theoretical methods* (CRC Press, Taylor & Francis Group, Boca Raton, 2017), ISBN: 978-1-4822-4634-6.

[24] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro and S. Lloyd, 'Gaussian quantum information', Reviews of Modern Physics 84, 621–669 (2012).

[25] P. P. Rohde, *The Quantum Internet: The Second Quantum Revolution*, 1st ed. (Cambridge University Press, 30th Sept. 2021), ISBN: 978-1-108-86881-5.

[26] P. Kok and B. W. Lovett, *Introduction to optical quantum information processing* (Cambridge University Press, Cambridge, 2010), ISBN: 978-0-511-77618-2.

[27] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*, 10th anniversary edition (Cambridge university press, Cambridge, 2010), ISBN: 978-1-107-00217-3.

[28] A. Shamir, 'How to share a secret', Communications of the ACM 22, 612–613 (1979).

[29] G. R. Blakley, 'Safeguarding cryptographic keys', in 1979 International Workshop on Managing Requirements Knowledge (MARK) (June 1979), pp. 313–318, ISBN: 978-1-5090-3181-8.

[30] A. M. Lance, T. Symul, W. P. Bowen, T. Tyc, B. C. Sanders and P. K. Lam, 'Is quantum secret sharing different to the sharing of a quantum secret?', in Proceedings Volume 5161, Quantum Communications and Quantum Imaging, edited by R. E. Meyers and Y. Shih (3rd Feb. 2004), p. 127.

[31] M. Hillery, V. Bužek and A. Berthiaume, 'Quantum secret sharing', Physical Review A 59, 1829–1834 (1999).

[32] A. Karlsson, M. Koashi and N. Imoto, 'Quantum entanglement for secret sharing and secret splitting', Physical Review A 59, 162–168 (1999).

[33] W. Tittel, H. Zbinden and N. Gisin, 'Experimental demonstration of quantum secret sharing', Physical Review A 63, 042301 (2001).

[34] Y. Ouyang, S.-H. Tan, L. Zhao and J. F. Fitzsimons, 'Computing on quantum shared secrets', Physical Review A 96, 052333 (2017).

[35] T. Symmul, A. Lance, W. Bowen, P. Lam, B. Sanders and T. C. Ralph, 'Quantum State Sharing', in *Quantum communications and cryptography*, edited by A. V. Sergienko (CRC Press, Taylor & Francis Group, Boca Raton London New York, 2019), ISBN: 978-0-367-39174-4.

[36] C. Crépeau, D. Gottesman and A. Smith, 'Secure multi-party quantum computation', in Proceedings of the thiry-fourth annual ACM symposium on Theory of computing, STOC '02 (19th May 2002), pp. 643–652, ISBN: 978-1-58113-495-7.

[37] M. Grassl, W. Geiselmann and T. Beth, 'Quantum Reed—Solomon Codes', in Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, Vol. 1719, edited by M. Fossorier, H. Imai, S. Lin and A. Poli, Lecture Notes in Computer Science (1999), pp. 231–244, ISBN: 978-3-540-46796-0.

[38] E. Villasenor and R. Malaney, 'A Three-Mode Erasure Code for Continuous Variable Quantum Communications', in GLOBECOM 2022 - 2022 IEEE Global Communications Conference (4th Dec. 2022), pp. 5231–5236, ISBN: 978-1-66543-540-6.

[39] M. Lassen, M. Sabuncu, A. Huck, J. Niset, N. Cerf, G. Leuchs and U. Andersen, 'Continuous-Variable Quantum Erasure Correcting Code', in Conference on Lasers and Electro-Optics 2010 (2010), JThE25, ISBN: 978-1-55752-889-6.

[40] R. Cleve, D. Gottesman and H.-K. Lo, 'How to Share a Quantum Secret', Physical Review Letters 83, 648–651 (1999).

[41] S.-B. Zheng, 'Splitting quantum information via W states', Physical Review A 74, 054303 (2006).

[42] S. Muralidharan and P. K. Panigrahi, 'Quantum-information splitting using multipartite cluster states', Physical Review A 78, 062333 (2008).

[43] T. Tyc and B. C. Sanders, 'How to share a continuous-variable quantum secret by optical interferometry', Physical Review A 65, 042310 (2002).

[44] A. M. Lance, T. Symul, W. P. Bowen, T. Tyc, B. C. Sanders and P. K. Lam, 'Continuous variable (2, 3) threshold quantum secret sharing schemes', New Journal of Physics 5, 4 (2003).

[45] Q. He, L. Rosales-Zárate, G. Adesso and M. D. Reid, 'Secure Continuous Variable Teleportation and Einstein-Podolsky-Rosen Steering', Physical Review Letters 115, 180502 (2015).

[46] T. Tyc, D. J. Rowe and B. C. Sanders, 'Efficient sharing of a continuous-variable quantum secret', Journal of Physics A: Mathematical and General 36, 7625–7637 (2003).

[47] B. C. Sanders, T. Tyc and D. J. Rowe, 'Sharing quantum secrets', in *Quantum Communications and Quantum Imaging*, Vol. 5161, edited by R. E. Meyers and Y. Shih (SPIE, San Diego, California, USA, 3rd Feb. 2004), pp. 116–126.

[48] A. M. Lance, T. Symul, W. P. Bowen, B. C. Sanders and P. K. Lam, 'Tripartite Quantum State Sharing', Physical Review Letters 92, 177903 (2004).

[49] A. M. Lance, T. Symul, W. P. Bowen, B. C. Sanders, T. Tyc, T. C. Ralph and P. K. Lam, 'Continuous-variable quantum-state sharing via quantum disentanglement', Physical Review A 71, 033814 (2005).

[50] C. Wilkinson, 'Characterising the resource state requirements for secure quantum state sharing', Project Report (University of St Andrews, May 2020).

[51] A. Unnikrishnan and D. Markham, 'Authenticated teleportation and verification in a noisy network', Physical Review A 102, 042401 (2020).

[52] C. Portmann and R. Renner, 'Security in quantum cryptography', Reviews of Modern Physics 94, 025008 (2022).

[53] K. G. Fedorov, M. Renger, S. Pogorzalek, R. Di Candia, Q. Chen, Y. Nojiri, K. Inomata, Y. Nakamura, M. Partanen, A. Marx, R. Gross and F. Deppe, 'Experimental quantum teleportation of propagating microwaves', Science Advances 7, eabk0891 (2021).

[54] F. Grosshans and P. Grangier, 'Quantum cloning and teleportation criteria for continuous quantum variables', Physical Review A 64, 010301(R) (2001).

[55] W. K. Wootters and W. H. Zurek, 'A single quantum cannot be cloned', Nature 299, 802–803 (1982).

[56] S. Olivares, M. G. A. Paris and U. L. Andersen, 'Cloning of Gaussian states by linear optics', Physical Review A 73, 062330 (2006).

[57] V. Bužek and M. Hillery, 'Quantum copying: Beyond the no-cloning theorem', Physical Review A 54, 1844–1852 (1996).

[58] V. Bužek and M. Hillery, 'Universal Optimal Cloning of Arbitrary Quantum States: From Qubits to Quantum Registers', Physical Review Letters 81, 5003–5006 (1998).

[59] W. P. Bowen, N. Treps, B. C. Buchler, R. Schnabel, T. C. Ralph, H.-A. Bachor, T. Symul and P. K. Lam, 'Experimental investigation of continuous variable quantum teleportation', Physical Review A 67, 032302 (2003).

[60] K. Hammerer, M. M. Wolf, E. S. Polzik and J. I. Cirac, 'Quantum Benchmark for Storage and Transmission of Coherent States', Physical Review Letters 94, 150503 (2005).

[61] P. T. Cochrane, T. C. Ralph and A. Dolińska, 'Optimal cloning for finite distributions of coherent states', Physical Review A 69, 042313 (2004).

[62] S. Pogorzalek, K. G. Fedorov, M. Xu, A. Parra-Rodriguez, M. Sanz, M. Fischer, E. Xie, K. Inomata, Y. Nakamura, E. Solano, A. Marx, F. Deppe and R. Gross, 'Secure quantum remote state preparation of squeezed microwave states', Nature Communications 10, 2604 (2019).

[63] M. Guţă and K. Matsumoto, 'Optimal cloning of mixed Gaussian states', Physical Review A 74, 032305 (2006).

[64] L. Ge, J. Xin and L. Zhang, 'Universal Global Cloning of Continuous Variables Entanglement', Annalen der Physik 535, 2200453 (2023).

[65] V. Nordgren, O. Leskovjanová, J. Provazník, A. Johnston, N. Korolkova and L. Mišta, 'Certifying emergent genuine multipartite entanglement with a partially blind witness', Physical Review A 106, 062410 (2022).

[66] M. V. Larsen, X. Guo, C. R. Breum, J. S. Neergaard-Nielsen and U. L. Andersen, 'Deterministic generation of a two-dimensional cluster state', Science 366, 369–372 (2019).

[67] P. Van Loock, 'Optical hybrid approaches to quantum information', Laser & Photonics Reviews 5, 167–200 (2011).

[68] M. He, R. Malaney and R. Aguinaldo, 'Teleportation of discrete-variable qubits via continuous-variable lossy channels', Physical Review A 105, 062407 (2022).

[69] R. E. S. Polkinghorne and T. C. Ralph, 'Continuous Variable Entanglement Swapping', Physical Review Letters 83, 10.1103/PhysRevLett.83.2095 (1999).

[70] S. H. Lie and H. Jeong, 'Limitations of teleporting a qubit via a two-mode squeezed state', Photonics Research 7, A7 (2019).

[71] J. S. Ivan, K. K. Sabapathy and R. Simon, 'Operator-sum representation for bosonic Gaussian channels', Physical Review A 84, 042311 (2011).

[72] C. M. Caves and K. Wódkiewicz, 'Fidelity of Gaussian Channels', Open Systems & Information Dynamics 11, 309–323 (2004).

[73] A. Serafini and G. Adesso, 'Standard forms and entanglement engineering of multimode Gaussian states under local operations', Journal of Physics A: Mathematical and Theoretical 40, 8041–8053 (2007).

[74] W. Schleich, *Quantum optics in phase space*, 1st edition (Wiley-VCH, Berlin Weinheim New York, 2001), ISBN: 978-3-527-29435-0.

[75] T. H. Koornwinder, R. Wong, R. Koekoek, R. F. Swarttouw and W. P. Reinhardt, 'Orthogonal polynomials', in *Nist handbook of mathematical functions*, edited by F. W. J. Olver, D. W. Lozier, R. F. Boisvert and C. W. Clark (Cambridge University Press, 2010), ISBN: 978-0-521-19225-5.

[76] A. Altland and B. Simons, *Condensed matter field theory*, 2nd ed (Cambridge University Press, Cambridge ; New York, 2010), ISBN: 978-0-521-76975-4.

[77] M. W. Wong, *Weyl transforms*, Universitext (Springer, New York, 1998), ISBN: 978-0-387-98414-8.

[78] W. R. Inc., *Mathematica, Version 14.2*, Champaign, IL, 2024.

[79] M. AbuGhanem, 'IBM quantum computers: evolution, performance, and future directions', The Journal of Supercomputing 81, 687 (2025).

[80] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven and J. M. Martinis, 'Quantum supremacy using a programmable superconducting processor', Nature 574, 505–510 (2019).

[81] L. K. Grover, 'A fast quantum mechanical algorithm for database search', in Proceedings of the twenty-eighth annual ACM symposium on Theory of computing - STOC '96 (1996), pp. 212–219.

[82] P. Shor, 'Algorithms for quantum computation: discrete logarithms and factoring', in Proceedings 35th Annual Symposium on Foundations of Computer Science (1994), pp. 124–134.

[83] X.-M. Hu, Y. Guo, B.-H. Liu, C.-F. Li and G.-C. Guo, 'Progress in quantum teleportation', Nature Reviews Physics 5, 339–353 (2023).

[84] H. Kristjánsson, R. Gardner and G. Chiribella, *Quantum Communications Report for Ofcom*, Policy Paper (Ofcom, 28th July 2021).

[85] S. L. Braunstein and P. van Loock, 'Quantum information with continuous variables', Reviews of Modern Physics 77, 513–577 (2005).

[86] A. Azzalini, 'A Class of Distributions Which Includes the Normal Ones', Scandinavian Journal of Statistics 12, 171–178 (1985).

[87] A. Azzalini and A. Capitanio, *The Skew-Normal and Related Families*, 1st ed. (Cambridge University Press, 19th Dec. 2013), ISBN: 978-1-139-24889-1.

[88] M. Tsagris, C. Beneki and H. Hassani, 'On the Folded Normal Distribution', Mathematics 2, 12–28 (2014).

[89] A. Wunsche, 'Displaced Fock states and their connection to quasiprobabilities', Quantum Optics: Journal of the European Optical Society Part B 3, 359–383 (1991).

[90] *The convergence of healthcare and pharmaceuticals with quantum computing: A new frontier in medicine*, Insights Paper (National Quantum Computing Centre, 25th Mar. 2025).

[91] S. Takeda, T. Mizuta, M. Fuwa, P. Van Loock and A. Furusawa, 'Deterministic quantum teleportation of photonic quantum bits by a hybrid technique', Nature 500, 315–318 (2013).

[92] D. Travers, *Secure Gaussian quantum state sharing for a network of five players*, Research Report (University of St Andrews, Aug. 2024).

[93] R. Gittings, 'Large-Scale Quantum Secret Sharing: Fidelity and Steering Constraints in the (3,5) Threshold Scheme', Project Report (University of St Andrews, Apr. 2025).

[94] K. Weber, 'Microwave quantum secret sharing', Project Report (Walther Meißner Institute, May 2025).

[95] K. Weber, W. K. Yam, K. Fedorov, C. Wilkinson and N. Korolkova, '[Manuscript in preparation]', (2025).

[96] V. Giovannetti, S. Lloyd and L. Maccone, 'Quantum Metrology', Physical Review Letters 96, 010401 (2006).

[97] A. Sunderland, L. Ju, D. G. Blair, W. McRae and A. V. Veryaskin, 'Optimizing a direct string magnetic gradiometer for geophysical exploration', Review of Scientific Instruments 80, 104705 (2009).

[98] T. R. Clem, 'Superconducting Magnetic Gradiometers For Underwater Target Detection', Naval Engineers Journal 110, 139–149 (1998).

[99] C. D. Gratta, V. Pizzella, F. Tecchio and G. L. Romani, 'Magnetoencephalography - a noninvasive brain imaging method with 1 ms time resolution', Reports on Progress in Physics 64, 1759–1814 (2001).

[100] S. Altenburg, M. Oszmaniec, S. Wölk and O. Gühne, 'Estimation of gradients in quantum metrology', Physical Review A 96, 042319 (2017).

[101] J. S. Sidhu and P. Kok, 'Geometric perspective on quantum parameter estimation', AVS Quantum Science 2, 014701 (2020).

[102] C. W. Helstrom, *Quantum detection and estimation theory*, Mathematics in Science and Engineering v. 123 (Academic Press, New York, 1976), ISBN: 978-0-12-340050-5.

[103] A. C. Atkinson, A. N. Donev and R. D. Tobias, *Optimum experimental designs, with SAS*, Oxford Statistical Science Series 34 (Oxford university press, Oxford, 2007), ISBN: 978-0-19-929660-6.

[104] T. A. Severini, *Likelihood methods in statistics*, Oxford Statistical Science Series 22 (Oxford University Press, Oxford ; New York, 2000), ISBN: 978-0-19-850650-8.

[105] M. G. A. Paris, 'Quantum estimation for quantum technology', International Journal of Quantum Information 07, 125–137 (2009).

[106] S. L. Braunstein, 'How large a sample is needed for the maximum likelihood estimator to be approximately Gaussian?', Journal of Physics A: Mathematical and General 25, 3813–3826 (1992).

[107] S. L. Braunstein and C. M. Caves, 'Statistical distance and the geometry of quantum states', Physical Review Letters 72, 3439–3443 (1994).

[108] S. Pang and T. A. Brun, 'Quantum metrology for a general Hamiltonian parameter', Physical Review A 90, 022117 (2014).

[109] J. Liu, X.-X. Jing and X. Wang, 'Quantum metrology with unitary parametrization processes', Scientific Reports 5, 8565 (2015).

[110] Z. Huang, C. Macchiavello and L. Maccone, 'Usefulness of entanglement-assisted quantum metrology', Physical Review A 94, 012101 (2016).

[111] J. Liu, H. Yuan, X.-M. Lu and X. Wang, 'Quantum Fisher information matrix and multiparameter estimation', Journal of Physics A: Mathematical and Theoretical 53, 023001 (2020).

[112] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. Van Der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. Van Mulbregt, SciPy 1.0 Contributors, A. Vijaykumar, A. P. Bardelli, A. Rothberg, A. Hilboll, A. Kloeckner, A. Scopatz, A. Lee, A. Rokem, C. N. Woods, C. Fulton, C. Masson, C. Häggström, C. Fitzgerald, D. A. Nicholson, D. R. Hagen, D. V. Pasechnik, E. Olivetti, E. Martin, E. Wieser, F. Silva, F. Lenders, F. Wilhelm, G. Young, G. A. Price, G.-L. Ingold, G. E. Allen, G. R. Lee, H. Audren, I. Probst,

J. P. Dietrich, J. Silterra, J. T. Webber, J. Slavič, J. Nothman, J. Buchner, J. Kulick, J. L. Schönberger, J. V. De Miranda Cardoso, J. Reimer, J. Harrington, J. L. C. Rodríguez, J. Nunez-Iglesias, J. Kuczynski, K. Tritz, M. Thoma, M. Newville, M. Kümmerer, M. Bolingbroke, M. Tartre, M. Pak, N. J. Smith, N. Nowaczyk, N. Shebanov, O. Pavlyk, P. A. Brodtkorb, P. Lee, R. T. McGibbon, R. Feldbauer, S. Lewis, S. Tygier, S. Sievert, S. Vigna, S. Peterson, S. More, T. Pudlik, T. Oshima, T. J. Pingel, T. P. Robitaille, T. Spura, T. R. Jones, T. Cera, T. Leslie, T. Zito, T. Krauss, U. Upadhyay, Y. O. Halchenko and Y. Vázquez-Baeza, 'SciPy 1.0: fundamental algorithms for scientific computing in Python', Nature Methods 17, 261–272 (2020).

[113] R. H. Byrd, P. Lu, J. Nocedal and C. Zhu, 'A Limited Memory Algorithm for Bound Constrained Optimization', SIAM Journal on Scientific Computing 16, 1190–1208 (1995).

[114] D. J. Wales and J. P. K. Doye, 'Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms', The Journal of Physical Chemistry A 101, 5111–5116 (1997).

[115] B. Olson, I. Hashmi, K. Molloy and A. Shehu, 'Basin Hopping as a General and Versatile Optimization Framework for the Characterization of Biological Macromolecules', Advances in Artificial Intelligence 2012, 1–19 (2012).

[116] M. Iwamatsu and Y. Okabe, 'Basin hopping with occasional jumping', Chemical Physics Letters 399, 396–400 (2004).

[117] A. Grosso, M. Locatelli and F. Schoen, 'A Population-based Approach for Hard Global Optimization Problems based on Dissimilarity Measures', Mathematical Programming 110, 373–404 (2007).

[118] A. Grosso, M. Locatelli and F. Schoen, 'An experimental analysis of a population based approach for global optimization', Computational Optimization and Applications 38, 351–370 (2007).

[119] M. Wilde, *Quantum information theory* (Cambridge University Press, Cambridge, 2013), ISBN: 978-1-139-52534-3.

[120] M.-K. Zhou, Z.-K. Hu, X.-C. Duan, B.-L. Sun, J.-B. Zhao and J. Luo, 'Precisely mapping the magnetic field gradient in vacuum with an atom interferometer', Physical Review A 82, 061602 (2010).

[121] F. Schmidt-Kaler and R. Gerritsma, 'Entangled states of trapped ions allow measuring the magnetic field gradient produced by a single atomic spin', Europhysics Letters 99, 53001 (2012).

[122] I. Urizar-Lanz, P. Hyllus, I. L. Egusquiza, M. W. Mitchell and G. Tóth, 'Macroscopic singlet states for gradient magnetometry', Physical Review A 88, 013626 (2013).

[123] I. Apellaniz, I. Urizar-Lanz, Z. Zimborás, P. Hyllus and G. Tóth, 'Precision bounds for gradient magnetometry with atomic ensembles', Physical Review A 97, 053603 (2018).

[124] T. Baumgratz and A. Datta, 'Quantum Enhanced Estimation of a Multidimensional Field', Physical Review Letters 116, 030801 (2016).

[125] S. Altenburg and S. Wölk, 'Multi-parameter estimation: global, local and sequential strategies', Physica Scripta 94, 014001 (2019).

[126] R. Demkowicz-Dobrzański and L. Maccone, 'Using Entanglement Against Noise in Quantum Metrology', Physical Review Letters 113, 250801 (2014).

[127] P. Sekatski, S. Wölk and W. Dür, 'Optimal distributed sensing in noisy environments', Physical Review Research 2, 023052 (2020).

[128] M. Cerezo, G. Verdon, H.-Y. Huang, L. Cincio and P. J. Coles, 'Challenges and opportunities in quantum machine learning', Nature Computational Science 2, 567–576 (2022).

[129] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio and P. J. Coles, 'Variational quantum algorithms', Nature Reviews Physics 3, 625–644 (2021).

[130] J. J. Meyer, J. Borregaard and J. Eisert, 'A variational toolbox for quantum multi-parameter estimation', npj Quantum Information 7, 89 (2021).

[131] J. L. Beckey, M. Cerezo, A. Sone and P. J. Coles, 'Variational quantum algorithm for estimating the quantum Fisher information', Physical Review Research 4, 013083 (2022).

[132] *Nist digital library of mathematical functions*, `https://dlmf.nist.gov/`, Release 1.2.3 of 2024-12-15, F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.

[133] D. S. Bernstein, *Matrix mathematics: theory, facts, and formulas with application to linear systems theory* (Princeton University Press, Princeton, N.J, 2005), ISBN: 978-0-691-11802-4.

[134] G. E. Andrews, R. Askey and R. Roy, *Special Functions*, Encyclopedia of Mathematics and Its Applications (Cambridge University Press, Cambridge, 1999), ISBN: 978-0-521-62321-6.

[135] A. B. O. Daalhuis, 'Hypergeometric function', in *Nist handbook of mathematical functions*, edited by F. W. J. Olver, D. W. Lozier, R. F. Boisvert and C. W. Clark (Cambridge University Press, 2010), ISBN: 978-0-521-19225-5.