Continuous variable quantum advantages and applications in quantum optics

Ulysse Chabaud

Continuous Variable Quantum Advantages

and Applications in Quantum Optics

By

ULYSSE CHABAUD

Laboratoire d'Informatique de Paris 6
SORBONNE UNIVERSITÉ

A dissertation submitted to Sorbonne Université in accordance with the requirements of the degree of DOCTOR OF PHILOSOPHY, under the supervision of Damian Markham and Elham Kashefi.

Members of the jury: Anthony Levrrier, Andreas Winter, Perola Milman, Gerardo Adesso, Sébastien Tanzili.

JUNE 2020

Abstract

Quantum physics has led to a revolution in our conception of the nature of our world and is now bringing about a technological revolution. The use of quantum information promises indeed applications that outperform those of today's so-called classical devices. Continuous variable quantum information theory refers to the study of quantum information encoded in continuous degrees of freedom of quantum systems. This theory extends the mathematical study of quantum information to quantum states in Hilbert spaces of infinite dimension. It offers different perspectives compared to discrete variable quantum information theory and is particularly suitable for the description of quantum states of light. Quantum optics is thus a natural experimental platform for developing quantum applications in continuous variable.

This thesis focuses on three main questions: where does a *quantum advantage*, that is, the ability of quantum machines to outperform classical machines, come from? How to ensure the proper functioning of a quantum machine? What advantages can be gained in practice from the use of quantum information? These three questions are at the heart of the development of future quantum technologies and we provide several answers within the frameworks of continuous variable quantum information and linear quantum optics.

Quantum advantage in continuous variable comes in particular from the use of so-called *non-Gaussian* quantum states. We introduce the stellar formalism to characterize these states. We then study the transition from classically simulable models to models universal for quantum computing. We show that *quantum computational supremacy*, the dramatic speedup of quantum computers over their classical counterparts, may be realised with non-Gaussian states and Gaussian measurements.

Quantum certification denotes the methods seeking to verify the correct functioning of a quantum machine. We consider certification of quantum states in continuous variable, introducing several protocols according to the assumptions made on the tested state. We develop efficient methods for the verification of a large class of multimode quantum states, including the output states of the *Boson Sampling* model, enabling the experimental verification of quantum supremacy with photonic quantum computing.

We give several new examples of practical applications of quantum information in linear quantum optics. Generalising the *swap test*, we highlight a connection between the ability to distinguish two quantum states and the ability to perform universal programmable quantum measurements, for which we give various implementations in linear optics, based on the use of single photons or coherent states. Finally, we obtain, thanks to linear optics, the first implementation of a quantum protocol for *weak coin flipping*, a building block for many cryptographic applications.

Table of Contents

INTRODUCTION

Quantum mechanics has deepened our understanding of the world. It has led us to rethink the very notion of reality—how can a cat be neither dead nor alive?—by putting forth intriguing properties such as *entanglement* and *superposition*. Nowadays, new information processing devices using quantum properties are being developed, such as quantum computers, and it is fascinating and maybe incumbent to see whether and to what extent these quantum technologies may outperform conventional technologies.

Motivation and context

While classical mechanics, as opposed to quantum, has been quite successful in describing the world at our scale, quantum mechanics has proven to be a very powerful tool for understanding the world at the particle scale. Interesting effects appear at this scale, and the challenge posed by the development of quantum technologies is not only to understand these effects but also to harness them. Quantum information—that is, information encoded in quantum degrees of freedom of physical systems—promises advantages over classical information notably for computing, communication, cryptography and sensing. That the use of quantum mechanics may provide an advantage over classical mechanics for information processing is an exciting perspective, which raises the following question:

What leads to a quantum advantage?

This profound question has attracted enormous attention and so far has only partial answers. From a foundational point of view, this question asks what differentiates the quantum from the classical and what makes nature fundamentally nonclassical. While shedding light on the very nature of our world, answering this question also enables the development of new technologies exploiting quantum properties to gain an advantage over classical machines.

In order to understand the possible origins of a quantum advantage it is worthwhile to highlight some of the differences between quantum and classical information and in particular quantum features that are inherently nonclassical.

Properties of quantum systems are intrinsically random prior to being measured and this randomness is lost whenever the quantum system is measured—hence the infamous Schrödinger's

cat thought experiment, in which a cat is locked in a box with a device that kills the animal with some probability: before opening the box, the cat is neither dead nor alive, but rather in a superposition of these two states, and opening the box collapses the state of the cat to either dead or alive. In a more general fashion, the state of a quantum system can be mathematically described by a wave function consisting of complex-valued probability amplitudes. The probabilities for the possible results of measurements made on the system can be derived from these amplitudes. As their name indicates, wave functions behave qualitatively like mechanical waves: they satisfy a linear wave equation and may interfere. This interference of probability amplitudes is a striking example of nonclassical phenomena. A quantum computer outperforming its classical counterpart would crucially interfere various branches of a computation.

The linear evolution of probability amplitudes also has striking consequences: it implies that an arbitrary quantum state cannot be perfectly cloned [WZ82]. This contrasts with the fact that classical information is trivial to copy. This quantum no-cloning property can also be derived from the uncertainty principle, which asserts that complementary quantum observables—such as position and momentum—cannot be simultaneously measured with arbitrary precision: measuring one of the two collapses the state of the measured quantum system such that the value of the other becomes uniformly random. If one was able to perfectly clone a quantum state, one could measure the position of the first copy and the momentum of the second and infer both quantities for the original state, thus contradicting the uncertainty principle. While uncertainty and no-cloning may be seen as limitations of quantum information, quantum advantage in cryptography notably comes from exploiting these properties to hide information from a possible eavesdropper [BB84b].

These quantum properties may be witnessed already for a single system. On the other hand, multiple systems may display correlations and it turns out that quantum systems may be correlated in a way classical systems cannot, as a consequence of entanglement. A quantum state over multiple subsystems is said to be entangled if it cannot be separated into the individual states of its subsystems. An important consequence of entanglement is the nonlocality of quantum theory, i.e., the fact that correlations displayed by spatially separated quantum systems cannot be reproduced locally by classical means [Bel64]—what Einstein famously described as “spooky action at a distance”. While these nonclassical correlations may be exploited for the so-called quantum teleportation [BBC$^+$93], they do not allow for superluminal communication, as a consequence of the no-signaling principle [PT04].

In theory, the nonclassical properties previously described may allow quantum devices to outperform their classical counterparts for a variety of information processing tasks, and in particular to demonstrate quantum computational supremacy [HM17]—a quantum computer performing efficiently a computational task which is provably intractable for classical computers—which marks a key milestone in the development of quantum technologies [AAB$^+$19].

However, a major obstacle to the use of the nonclassical properties of quantum information for

technological applications is decoherence, i.e., the loss of coherence of the information encoded in a physical system, due to the interaction of that system with its environment. Quantum devices will inevitably interact with their environment and suffer the effect of noise. How to mitigate the consequences of decoherence is an active domain of research [Pre98a]. In theory, quantum computations may be performed fault-tolerantly, even though this results in a huge overhead in terms of physical systems needed for the computation. It is also not obvious how one can mitigate noise in other quantum information processing tasks, such as sensing or simulations, where the fault-tolerant quantum computing approach is not natural. Hence, another question that arises when looking for an advantage using a quantum device is the following:

How do we check the correct functioning of a quantum device?

Answering this second question is a timely problem in the absence of fault-tolerant mechanisms, for benchmarking existing and upcoming quantum devices. It has also attracted a lot of attention [EHW+20], under different names: validation, benchmarking, certification, verification. We shall use certification in the following when no context is precised. The task of certification may indeed vary depending on the context: fundamental research, industrial quantum device, or even delegated quantum computing and quantum cryptography. In all these cases, what may vary is the level of trust one wants to guarantee, as well as the assumptions one is ready to make on the device being tested.

The challenge posed by the certification of quantum devices therefore depends on this context. What is more, the very properties of quantum information—entanglement, unclonability—add uniquely quantum challenges to the task of certification, and the way the information is encoded in physical systems also matters.

Information, both classical and quantum, may be encoded using either discrete degrees of freedom of a physical system—such as the presence or absence of an electrical signal, or the spin of an electron—or continuous degrees of freedom—such as the position of a particle, or quadratures of the electromagnetic field.

A great part of the theory already developed for discrete variable quantum information is still missing for continuous variable quantum information. The latter is based on the beautiful mathematics of quantum mechanics in infinite-dimensional Hilbert spaces and gives different perspectives on quantum information [BvL05]. In addition, continuous variable quantum information has an exciting experimental status, thanks to quantum optics in particular, which enables the scalable generation of large entangled quantum states [YUA+13b] and provides high efficiency measurements. Moreover, some continuous variable quantum technologies—such as continuous variable quantum key distribution [GG02]—compete with their discrete variable counterparts [JKJL+13]. This implies that the question of certification is of great importance for continuous variable quantum devices, which also allow for outperforming classical devices and demonstrating quantum computational supremacy [AA13, HKS+16].

The demonstration of quantum supremacy, that is the convincing demonstration of a quantum computation beating what is possible classically, is however only a milestone, and what is at stake in the development of quantum technologies is to obtain advantages for real-world applications. It is thus natural to ask the following question:

What useful advantages can we obtain from the use of quantum information?

Depending on the application considered, a quantum advantage may take different forms: to obtain the result of a computation faster [Sho94, Gro98], to communicate more messages within the same physical system [BW92] or in a more secured fashion [BB84a], or to perform a measurement with a better precision [GLM11], for example. Answering this third question amounts to developping new theoretical quantum algorithms as well as deriving realistic implementations for existing ones, for example with linear quantum optics and quantum states of light.

This section has provided an overview of the different contexts on which the work of this thesis is based. Motivated by the three very general questions above—origin of quantum advantage, certification of quantum devices and useful quantum advantages—this dissertation explores various directions, with particular emphasis on continuous variable quantum information theory and optical quantum information processing. The next section presents a technical summary of the content of the thesis.

Summary of results

After a preliminary chapter 1, chapters 2 and 3 deal with continuous variable quantum information theory and computing. Chapters 4 and 5 consider the probems of quantum state certification and testing, in the continuous variable regime and using quantum optics. Chapter 6 discusses the implementation of a quantum cryptography protocol with quantum optics. We detail the content of each chapter in what follows. The dependencies between the chapters are indicated in Fig. 0.1.

Chapter 1. After briefly introducing preliminary material on quantum information theory, this chapter presents the formalisms of continuous variable quantum information theory used in this thesis. Phase-space formalism is discussed. A description of Gaussian states and processes follows, together with the symplectic formalism. Then, quantum linear optics is presented with an exposition of Boson Sampling [AA13]. Finally, the Segal–Bargmann formalism is introduced.

Chapter 2. This chapter investigates the origin of quantum advantage for continuous variable quantum computing. Continuous variable quantum states are separated into two broad families: Gaussian and non-Gaussian. While Gaussian states feature interesting properties such as entanglement, non-Gaussian states are crucial for a variety of quantum information tasks [ESP02, Fiu02, GC02, WHG$^+$03, GPFC$^+$04, GS07, NFC09, ADDS$^+$09, BDE$^+$19]. Characterizing and understanding the properties of these states is thus of major importance [TZ18, ZSS18,

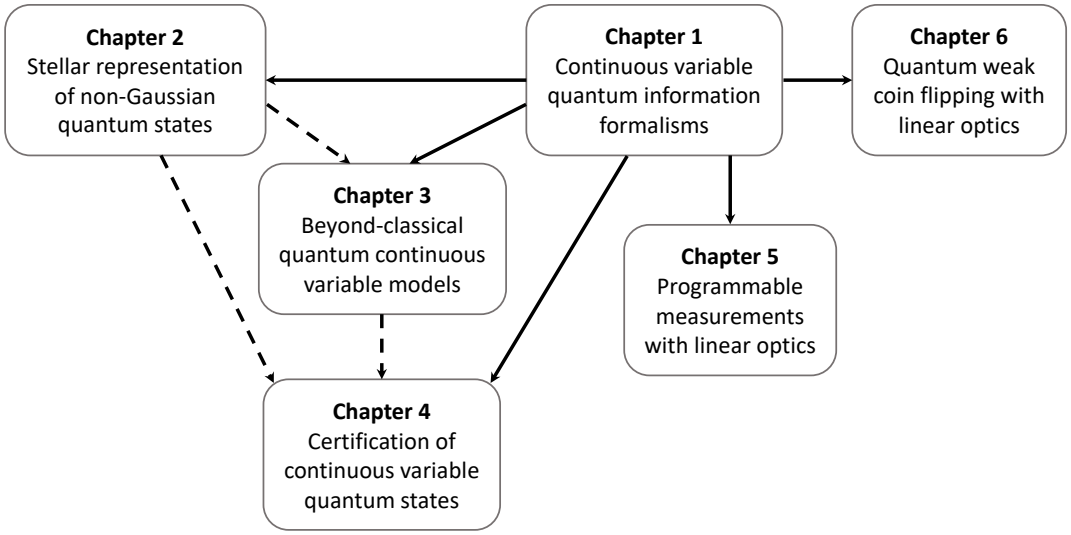

FIGURE 0.1. Dependencies between the chapters of this thesis. The solid arrows mean that one chapter depends on an other, while the dashed arrows indicate a partial dependence.

AGPF18, LRW$^+$18]. This chapter applies the Segal–Bargmann formalism to derive the stellar representation of single-mode non-Gaussian states. We define and study the stellar rank, using properties of holomorphic functions. This rank induces a hierarchy among continuous variable quantum states. The stellar representation is used to derive a criterion for Gaussian convertibility of states with finite stellar rank within this hierarchy. Its topology with respect to the trace norm is investigated and we show that the hierarchy is robust to small deviations and how to compute the robustness. The main result of this chapter is a classification of single-mode continuous variable quantum states with respect to their non-Gaussian properties, which can be experimentally witnessed and has consequences for non-Gaussian quantum state engineering.

Chapter 3. In this chapter, we explore the quantum advantage transition for continuous variables, i.e., the boundary between classically simulable quantum computational models and models capable of outperforming their classical counterparts, in terms of non-Gaussian resources. We give classical simulation algorithms for several quantum models and computational tasks, including linear optics with adaptive measurements and Gaussian circuits with non-Gaussian input states. Then, we introduce a subuniversal family of continuous variable circuits related to Boson Sampling called Continuous Variable Sampling from photon-added or photon-subtracted squeezed states (CVS) circuits. We show that the continuous output probability densities of these circuits are on average hard to sample exactly classically, by relating their output probabilities to permanents of $(0,1)$-matrices. The main results of this chapter are classical simulation algorithms for Gaussian circuits with weakly non-Gaussian input states, as well as showing how quantum supremacy may be achieved with non-Gaussian states, together with Gaussian operations and

measurements.

Chapter 4. This chapter considers the certification of continuous variable quantum states. Determining an unknown quantum state is difficult especially in continuous variables, where it is described by possibly infinitely many complex parameters. Existing methods like homodyne quantum state tomography require many different measurement settings and heavy classical post-processing [LR09]. This chapter shows how continuous variable quantum states can be efficiently verified: we introduce a reliable method for performing continuous variable quantum state state tomography using a single Gaussian measurement, namely heterodyne detection, which can be implemented with quantum optics ; then, we show how this tomography method may be promoted to a state certification protocol under the i.i.d. assumption, by adding an energy test. We also derive a similar protocol for continuous variable quantum state verification, making no assumption whatsoever on the state preparation method, using a de Finetti reduction for infinite-dimensional systems [RC09]. We further show that this protocol extends to the multimode case and allows us to efficiently verify output states of Boson Sampling and CVS interferometers. The main result of this chapter is a flexible protocol for building trust for a large class of multimode mode continuous variable quantum states with Gaussian measurements, which provides analytical confidence intervals and allow for a reliable verification of quantum computational supremacy with photonic quantum computing.

Chapter 5. On top of being a promising candidate for the demonstration of quantum supremacy with Boson Sampling, quantum optics provides an exciting experimental platform for near-term quantum applications, as well as for probing quantum behaviours. This chapter discusses the relations between quantum state discrimination, quantum state identity testing and universal programmable projective measurements and proposes implementations in linear optics. A generalisation of the swap test [BCWDW01] is introduced, together with its implementation in linear optics using single-photon encoding. We show how this allows us to construct universal quantum-programmable projective measurements, based on a simple classical post-processing of samples from number-resolving or parity detectors. In order to simplify the experimental requirements, an alternative scheme is derived which uses a coherent state encoding, a simpler interferometer and single-photon threshold detectors, with applications to optical quantum communication protocols.

Chapter 6. Cryptographic protocols are built from a selection of simpler functionalities, called primitives. Remarkably, quantum mechanics allows for the implementation of some primitives with information-theoretic security which can only be achieved with conditional security classically, i.e., by relying on computational assumptions. The so-called coin flipping by telephone [Blu83], or weak coin flipping, is one of such cryptographic primitives. It refers to the cryptographic scenario in which two mistrustful and distant parties want to agree on a random bit, while they favor opposite outcomes. The use of quantum mechanics allows for achieving better security than classical mechanics. However, even though various quantum weak coin flipping protocols have been theorised [SR02, KN04, Moc05, Moc07, ARV19], no practical implementation

has been proposed so far. This chapter introduces an implementation in linear optics of quantum weak coin flipping. The proposed implementation relies on adapting a theoretical protocol for quantum weak coin flipping [SR02] to linear optics, using the so-called dual-rail encoding, i.e., encoding a qubit with a photon in two spatial modes. The protocol can be implemented with current technology and may display quantum advantage over any classical protocol for the same task.

Additional remarks

This thesis is intended to be accessible to a reader familiar with the basics of quantum information and computing with discrete and continuous variables. Good introductions to the field of quantum information theory include [Pre98b] and [NC02], while [BvL05] provides a comprehensive review of quantum information with continuous variables. Pointers to the relevant literature are also displayed throughout the thesis.

This thesis is based on several previous works.

- **Chapter 2.** This chapter is mainly based on a joint work with D. Markham and F. Grosshans [CMG20b], and section 2.3.3 is based on a joint work with with G. Roland, M. Walschaers, V. Parigi, F. Grosshans, D. Markham and N. Treps [CRW$^+$20].

- **Chapter 3.** Section 3.2 is based on a joint work with A. Sohbi and D. Markham [CMS20], sections 3.3.1 and 3.3.2 on a joint work with D. Markham and F. Grosshans [CMG20a], and section 3.3.3 on a joint work with T. Douce, D. Markham, P. van Loock, E. Kashefi and G. Ferrini [CDM$^+$17].

- **Chapter 4.** Section 4.1 is based on a joint work with J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, and E. Kashefi [EHW$^+$20], sections 4.2 to 4.5 on a joint work with T. Douce, F. Grosshans, D. Markham and E. Kashefi [CDG$^+$19], section 4.6 on a joint work with with G. Roland, M. Walschaers, V. Parigi, F. Grosshans, D. Markham and N. Treps [CRW$^+$20], and section 4.7 on a joint work with F. Grosshans, D. Markham and E. Kashefi [CGKM20].

- **Chapter 5.** Sections 5.1 and 5.2 are based on a joint work with E. Diamanti, D. Markham, E. Kashefi and A. Joux [CDM$^+$18], and section 5.3 on a joint work with N. Kumar, E. Kashefi, D. Markham, and E. Diamanti [KCK$^+$20].

- **Chapter 6.** This chapter is based on a joint work with M. Bozzio, E. Diamanti and I. Kerenidis [BCKD20].

CHAPTER 1

CONTINUOUS VARIABLE QUANTUM INFORMATION FORMALISMS

Continuous variable quantum information theory refers to the study of information encoded in quantum physical systems with continuous degrees of freedom. The approach of the work presented in this dissertation for studying continuous variable quantum information is to use different mathematical formalisms as different ways of gaining intuition. Juggling several representations of the same mathematical object is indeed an excellent way to get insights about this object. In this chapter, we briefly review the formalisms for continuous variable quantum information theory used throughout the rest of the thesis. These include phase space formalism for continuous variable quantum states and operators, symplectic formalism for Gaussian states, quantum optics and Boson Sampling, and Segal–Bargmann formalism for continuous variable quantum states.

1.1 Preliminary material

1.1.1 Notations

The sets $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{C}$ are the usual sets of natural, real and complex numbers, with a $*$ exponent when 0 is removed from the set. The size of a set $\mathscr{X}$ is denoted by $|\mathscr{X}|$. The natural logarithm is denoted log.

We write complexity classes with sans serif font: P, NP...

The number of subsystems or modes will generally be denoted by $m \in \mathbb{N}^*$. Hilbert spaces are denoted by $\mathscr{H}$ or $\mathscr{K}$. The expressions $|\phi\rangle$, $|\psi\rangle$ denote pure states, and ρ and σ denote density operators of possibly mixed quantum states.

For vectors and operators, we denote by a $*$ exponent the complex conjugate, by a T exponent the transpose and by a $\dagger$ exponent the transpose complex conjugate (adjoint). Matrices are

denoted by capital letters and covariance matrices will be denoted by $\boldsymbol{V}$. Operators are indicated by a hat, with the exception of density operators, positive-operator valued measure elements and identity operator $\mathbb{1}$. In particular, $\hat{a}$ and $\hat{a}^\dagger$ denote the annihilation and creation operators and $\hat{q}$ and $\hat{p}$ denote the position-like and momentum-like quadrature operators. The identity matrix is also denoted $\mathbb{1}$, sometimes with an index indicating its size. The zero matrix is similarly denoted $\mathbb{0}$. The trace is denoted by Tr and the determinant by Det.

Pr denotes a probability, while $\mathbb{E}$ denotes an expected value. A function δ may stand for the Kronecker symbol or a Dirac delta, depending on the context. The letters α, β and γ are used for coherent state amplitudes or complex amplitudes, while the letters ξ and ζ are used for squeezing parameters. The letter z denotes a complex variable.

We write $\otimes$ and $\oplus$ for the tensor product and the direct sum, respectively. We use bold math for multimode states, vectors and multi-index notations. Let $m,n \in \mathbb{N}^*$. We define $\mathbf{0} = (0,\dots,0)$ and $\mathbf{1} = (1,\dots,1)$, and we write $\mathbf{0}^n = (0,\dots,0) \in \mathbb{N}^n$ or $\mathbf{1}^n = (1,\dots,1) \in \mathbb{N}^n$ to avoid ambiguity. For all $k \in \{1,\dots,m\}$, we also define $\mathbf{1}_k = (0,\dots,0,1,0,\dots,0)$, where the k^{th} entry is 1 and all the other $m-1$ entries are 0. For all $\boldsymbol{z} = (z_1,\dots,z_m) \in \mathbb{C}^m$, all $\boldsymbol{z}' = (z'_1,\dots,z'_m) \in \mathbb{C}^m$ and all $\boldsymbol{p} = (p_1,\dots,p_m) \in \mathbb{N}^m$ we write

$$
\begin{aligned}
&\boldsymbol{z}^* = (z_1^*,\dots,z_m^*)\\
&-\boldsymbol{z} = (-z_1,\dots,-z_m)\\
&\tilde{\boldsymbol{z}} = \boldsymbol{z} \oplus \boldsymbol{z}^* = (z_1,\dots,z_m,z_1^*,\dots,z_m^*)\\
&|\boldsymbol{z}\rangle = |z_1 \dots z_m\rangle\\
&\|\boldsymbol{z}\|^2 = |z_1|^2 + \dots + |z_m|^2\\
&\boldsymbol{z}^{\boldsymbol{p}} = z_1^{p_1} \dots z_m^{p_m}\\
&\boldsymbol{z} + \boldsymbol{z}' = (z_1 + z'_1,\dots,z_m + z'_m)\\
&\boldsymbol{z} \le \boldsymbol{z}' \Leftrightarrow z_k \le z'_k \quad \forall k \in \{1,\dots,m\}\\
&\boldsymbol{p}! = p_1! \dots p_m!\\
&|\boldsymbol{p}| = p_1 + \dots + p_m\\
&\partial^{\boldsymbol{p}} = \partial_1^{p_1} \dots \partial_m^{p_m}\\
&\left(\frac{\partial}{\partial \boldsymbol{z}}\right)^{\boldsymbol{p}} = \frac{\partial^{|\boldsymbol{p}|}}{\partial z_1^{p_1} \cdots \partial z_m^{p_m}}.
\end{aligned}
\tag{1.1}
$$

We will use for brevity the notations $c_\chi = \cosh\chi$, $s_\chi = \sinh\chi$ and $t_\chi = \tanh\chi$, for all $\chi \in \mathbb{C}$. The commutator is denoted by $[\,,\,]$ and the anticommutator by $\{\,,\,\}$. Finally we adopt the convention $\hbar = 1$ and use canonical conventions rather than optical ones.

Note that proofs of intermediate technical results will be indicated by a vertical bar running along the side of the page, with a square symbol marking the end of the proof.

1.1.2 Basics of quantum information theory

The presentation given here is very succinct and good introductions to the field of quantum information theory include [Pre98b] and [NC02].

In quantum information theory, we identify two notions of randomness. On the one hand, there is an inherent randomness in the formalism of quantum measurements, which we call quantum randomness. On the other hand, classical randomness corresponds to the usual notion of randomness to which we refer, for example, when we draw a card from a shuffled deck of cards or when we roll a die. In practice, a quantum system can manifest both classical and quantum randomness.

The properties of a quantum system are described by its quantum state. Quantum states with no classical randomness are called pure states. These pure quantum states are represented mathematically as normalised vectors in a separable Hilbert space $\mathscr{H}$. We adopt Dirac bra-ket notation [Dir81] in what follows: a column vector ψ is represented as the ket $|\psi\rangle$ and its adjoint (transpose complex conjugate) line vector is represented as the bra $\langle\psi|$. In particular, the projector onto $|\psi\rangle$ is expressed as $|\psi\rangle\langle\psi|$ and the inner product of two states $|\phi\rangle$ and $|\psi\rangle$ is denoted by $\langle\phi|\psi\rangle$, with $\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$. The quantity $|\langle\phi|\psi\rangle|^2$ is referred to as the overlap of the states $|\phi\rangle$ and $|\psi\rangle$.

The simplest nontrivial example is a Hilbert space of dimension 2. In that case, quantum states are referred to as qubit states, states in a Hilbert space of finite dimension $d > 2$ being referred to as qudit states. Given an orthonormal basis $(|0\rangle, |1\rangle)$ of a Hilbert space of dimension 2, a qubit state $|\psi\rangle$ is expressed as

$$|\psi\rangle = a|0\rangle + b|1\rangle, \qquad |a|^2 + |b|^2 = 1, \tag{1.2}$$

for $a, b \in \mathbb{C}$, with $\langle\psi| = a^*\langle 0| + b^*\langle 1|$. The coefficients a and b are the complex amplitudes of the qubit state $|\psi\rangle$. If $a \neq 0$ and $b \neq 0$, the state $|\psi\rangle$ is said to be in a superposition of the states $|0\rangle$ and $|1\rangle$.

The basis $(|0\rangle, |1\rangle)$ is referred to as the computational basis. On the other hand, setting $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$, the states $(|+\rangle, |-\rangle)$ also form an orthonormal basis, referred to as the diagonal basis.

Observable physical quantities, or simply observables, are represented mathematically by self-adjoint (hermitian) operators $\hat{O} = \hat{O}^\dagger$. Such operators have an orthonormal basis of eigenvectors and measuring the observable $\hat{O}$ gives an outcome sampled from the list of its eigenvalues. The probability of each outcome is determined by the Born rule:

$$\Pr[\lambda] = \langle\psi|\Pi_\lambda|\psi\rangle, \tag{1.3}$$

where λ is the eigenvalue, $|\psi\rangle$ is the state of the measured quantum system and Π_λ is a projector onto the eigenvector corresponding to the eigenvalue λ. Equivalently, we say that we measure in a specific orthonormal basis to say that we measure an observable which has this basis as

an eigenbasis. In particular, Eq. (1.2) may be interpreted as follows: $\langle\psi|\Pi_0|\psi\rangle = \langle\psi|0\rangle\langle 0|\psi\rangle = |\langle 0|\psi\rangle|^2 = |a|^2$ (resp. $|b|^2$) is the probability that we obtain the outcome 0 (resp. 1) when measuring the state $|\psi\rangle$ in the $(|0\rangle, |1\rangle)$ basis. The two probabilities sum to 1, corresponding to the fact that the measurement will yield an outcome, either 0 or 1. The measurement outcome is random when $a \neq 0$ and $b \neq 0$, i.e., quantum randomness manifests when the measured state is in a superposition of eigenvectors of the observable. Measuring a quantum state collapses the state onto the eigenvector corresponding to the outcome obtained. In particular, any subsequent measurement of the same observable will yield the same result with probability 1.

The most general notion of quantum measurement is captured by positive-operator valued measures (POVM). A POVM is a set of semidefinite operators $\{\Pi_i\}_{i\in\mathscr{I}}$ whose elements sum to the identity operator, indexed by a set of outcomes $\mathscr{I}$. The operator Π_i is associated to the measurement outcome $i \in \mathscr{I}$ and the probability for this outcome is given by Eq. (1.3), replacing λ by i. The case where the operators Π_i are projectors, as in Eq. (1.3), corresponds to projection-valued measures (PVM).

Quantum systems can also exhibit classical randomness. When that is the case, we refer to the quantum state as mixed. A mixed quantum state is represented mathematically by a so-called density operator, i.e., a hermitian operator with trace 1 acting on a Hilbert space. The density operator for a pure state $|\psi\rangle$ is simply a projector $|\psi\rangle\langle\psi|$. A mixed quantum state can be written as a convex combination, or mixture, of pure states. For example, the state obtained by flipping an unbiaised coin and choosing the state $|0\rangle$ for tails and $|1\rangle$ for heads is a mixed state expressed as $\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$, which is different from the pure superposition $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, whose density operator is given by $\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|0\rangle\langle 1| + \frac{1}{2}|1\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$. The Born rule for a mixed state ρ reads

$$\Pr[i] = \mathrm{Tr}(\Pi_i \rho), \tag{1.4}$$

where $\{\Pi_i\}_{i\in\mathscr{I}}$ is a POVM over a set of outcomes $\mathscr{I}$. Setting $\rho = |\psi\rangle\langle\psi|$, we retrieve the Born rule for pure states in Eq. (1.3). Writing the semidefinite operator $\Pi_i = M_i^\dagger M_i$, the state after a measurement with outcome $i \in \mathscr{I}$ is given by

$$\rho^{(i)} = \frac{M_i \rho M_i^\dagger}{\mathrm{Tr}(\Pi_i \rho)}. \tag{1.5}$$

Note that the choice of M_i is not unique and this choice reflects different possible ways of physically implementing the same POVM. Given an observable $\hat{O}$, the quantity $\mathrm{Tr}(\hat{O}\rho)$ is the expectation of the operator $\hat{O}$ for the quantum state ρ and is alternatively denoted $\langle\hat{O}\rangle_\rho$.

The global state of two independent quantum systems with states $|\phi\rangle$ and $|\psi\rangle$ in two Hilbert spaces $\mathscr{H}$ and $\mathscr{H}'$, respectively, lies in the tensor product $\mathscr{H}\otimes\mathscr{H}'$ and is obtained by taking the tensor product $|\phi\rangle\otimes|\psi\rangle$ of both states. We will usually write $|\phi\rangle\otimes|\psi\rangle = |\phi\psi\rangle$ when there is no ambiguity. The dimension of the Hilbert space $\mathscr{H}\otimes\mathscr{H}'$ is the product of the dimensions of the Hilbert spaces $\mathscr{H}$ and $\mathscr{H}'$, implying in particular that the computational basis of n-qubit states has size 2^n.

Two quantum systems may not be independent and a pure quantum state which cannot be written as a tensor product of quantum states is called entangled. For example, the state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ is entangled while the state $\frac{1}{2}(|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle)$ is separable. A (mixed) quantum state is called separable if it can be written as a mixture of separable pure states, and entangled otherwise.

Entanglement may be conceived as the quantum version of classical correlation [Wer89]: the mixed quantum state $\frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11|$ is classically correlated—the measurements of each subsystem in the $(|0\rangle, |1\rangle)$ basis will always yield the same outcomes—but not entangled, since it is a mixture of product states. On the other hand, the pure state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}(|{+}{+}\rangle + |{-}{-}\rangle)$ is entangled. This state is 'more' correlated than the previous one in the following sense: not only the measurements of each subsystem in the $(|0\rangle, |1\rangle)$ basis will always yield the same outcomes but measuring each subsystem in the $(|+\rangle, |-\rangle)$ basis will also always yield the same outcomes.

Given a state ρ over two subsystems in $\mathscr{H}$ and $\mathscr{H}'$, the reduced state of the first subsystem is obtained by tracing out, or taking the partial trace over, the second subsystem $\mathrm{Tr}_{\mathscr{H}'}(\rho)$. A separable state is fully described by the reduced states of its individual subsystems, while this is no longer the case for an entangled state.

The simplest example of evolution of a quantum system is a unitary evolution over a time t, described by a unitary operator $\hat{U}$ with $\hat{U}^\dagger \hat{U} = \mathbb{1}$, generated by a Hamiltonian H with $H^\dagger = H$, such that $\hat{U} = e^{-iHt}$. If the system is in a pure state $|\psi\rangle$, then the state after the evolution is a normalised pure state $\hat{U}|\psi\rangle$. If the system is in a mixed state ρ, then the state after the evolution is a mixed state with density operator $\hat{U}\rho\hat{U}^\dagger$.

More general quantum evolutions are described by quantum channels, i.e., completely positive trace-preserving maps (CPTP). By Stinespring dilation theorem, CPTP maps can be expressed as unitaries acting on a larger space. Formally, if $\mathscr{E}$ is a CPTP map acting on a Hilbert space $\mathscr{H}$, then there exist a Hilbert space $\mathscr{H}'$ and a unitary operator $\hat{U}$ such that for all density operators ρ,

$$\mathscr{E}(\rho) = \mathrm{Tr}_{\mathscr{H}'}[U(\rho \otimes |0\rangle\langle 0|)U^\dagger]. \tag{1.6}$$

In other words, any quantum channel can be obtained by tensoring with a second system in a fixed state, a unitary evolution and a reduction to a subsystem. Naimark's theorem provides a similar result for decomposing a POVM as a unitary followed by a PVM on a larger space.

The most general physical evolutions are described by quantum operations, i.e., completely positive trace-decreasing maps (CPTD). These operations can be obtained as obtained by tensoring with a second system in a fixed state, a unitary evolution, a PVM and a reduction to a subsystem. Non-CPTD maps are referred to as unphysical operations. Such operations can be approximated by quantum operations, for example when they act as CPTD maps on a subset of the Hilbert space.

A quantum computation is composed of the three following steps: input, evolution and measurement. With the above, one may conceive elaborate quantum computations as building a

highly entangled state from a simple input product state via a unitary evolution and sampling from a probability distribution given by the Born rule and the choice of measurement. Quantum computations can be looked at in the circuit picture, in which the unitary evolution is decomposed as a product of gates acting on at most two subsystems at a time.

Discrimination of quantum states is a central element in many quantum information processing tasks [NC02] and various measures are available [FVDG99]. We review two measures used extensively in the thesis: the fidelity and the trace distance. The properties outlined are independent of the dimension of the Hilbert space.

The fidelity between two states ρ and σ is defined as

$$F(\rho,\sigma) = \mathrm{Tr}\left(\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right)^2. \tag{1.7}$$

Note that the definition used here is the square of the definition in [FVDG99, NC02]. Even though it is not apparent with the above equation, the fidelity is symmetric in its arguments ρ and σ. When at least one of the two states is a pure state, this expression reduces to

$$F\left(\psi,\rho\right) = \mathrm{Tr}(|\psi\rangle\langle\psi|\,\rho) = \langle\psi|\rho|\psi\rangle\,. \tag{1.8}$$

In particular when both states are pure $F\left(\phi,\psi\right) = |\langle\phi|\psi\rangle|^2$.

We write the Schatten 1-norm of a bounded operator T as

$$\|T\|_1 = \mathrm{Tr}\left(\sqrt{T^\dagger T}\right) = \mathrm{Tr}(|T|). \tag{1.9}$$

The trace distance between two states ρ,σ is defined as

$$\begin{aligned} D(\rho,\sigma) &= \frac{1}{2}\|\rho-\sigma\|_1 \\ &= \frac{1}{2}\mathrm{Tr}(|\rho-\sigma|). \end{aligned} \tag{1.10}$$

It is jointly convex in its two arguments. The fidelity is related to the trace distance by the Fuchs-van de Graaf inequalities [FVDG99]

$$1-\sqrt{F(\rho,\sigma)} \le D(\rho,\sigma) \le \sqrt{1-F(\rho,\sigma)}. \tag{1.11}$$

When one of the states is pure, the lower bound may be refined as

$$1-F(\psi,\rho) \le D(\psi,\rho). \tag{1.12}$$

When both states are pure, the upper bound in Eq. (1.11) becomes an equality:

$$\begin{aligned} D(\phi,\psi) &= \sqrt{1-F(\phi,\psi)} \\ &= \sqrt{1-|\langle\phi|\psi\rangle|^2}. \end{aligned} \tag{1.13}$$

The fidelity is nondecreasing under quantum operations and the trace distance is nonincreasing under quantum operations. The total variation distance of two probability distributions P and Q over a sample space $\mathscr{S}$ is defined as

$$\|P-Q\|_{tvd} = \frac{1}{2}\sum_{s\in\mathscr{S}} |P(s)-Q(s)|. \tag{1.14}$$

A similar definition holds for probability densities over a continuous sample space, by replacing the discrete sum by a continuous sum. The trace distance verifies

$$D(\rho,\sigma) = \max_{\hat{O}} \|P_\rho^{\hat{O}} - P_\sigma^{\hat{O}}\|_{tvd}, \tag{1.15}$$

where $P_\rho^{\hat{O}}$ (resp. $P_\sigma^{\hat{O}}$) is the probability distribution associated to measuring the observable $\hat{O}$ for the state ρ (resp. σ) and where the maximum of the total variation distance is taken over all observables. The trace distance thus has an operational significance: if two states are close in trace distance, then any computation taking as input one of the two states is indistinguishable from the same computation taking as input the other state. Moreover, with Eq. (1.11), lower bounds on the fidelity also give upper bounds on the total variation distance, which are tight when the states are pure, by Eq. (1.13).

In what follows, we consider the case of infinite-dimensional Hilbert spaces, allowing for the description of quantum systems with continuous degrees of freedom. Discrete variables can be encoded in continuous degrees of freedom and finite-dimensional Hilbert spaces may be embedded in infinite-dimensional ones. Despite its discrete character, we will also refer to the study of such embedded discrete variable quantum information in an infinite-dimensional Hilbert space as continuous variable quantum information theory, since the same mathematical formalisms are employed in both case.

1.1.3 Continuous variable quantum information theory in a nutshell

We refer the reader to the first chapters of [BvL05, FOP05, ARL14] for a further introduction on the material presented in this section. While the presentation that follows is quite technical, it avoids many of the subtleties which appear when dealing with infinite-dimensional Hilbert spaces. The interested reader will find an example of a formal treatment in [DlM05].

The continuous variable equivalent of a qubit or qudit is the qumode, or simply mode. Single-mode continuous variable quantum states are mathematically described as normalised complex vectors in an infinite-dimensional separable Hilbert space, with an infinite countable orthonormal basis $\{|n\rangle\}_{n\in\mathbb{N}}$ referred to as the Fock basis, or photon-number basis in the context of optical quantum information processing. In particular, $|0\rangle$ is referred to as the vacuum state and $|1\rangle$ as the single-photon state. A single-mode pure state $|\psi\rangle$ can be written in Fock basis as

$$|\psi\rangle = \sum_{n\geq 0} \psi_n |n\rangle, \tag{1.16}$$

where $\psi_n \in \mathbb{C}$ for all $n \in \mathbb{N}$, with the normalisation condition $\sum_{n=0}^{+\infty} |\psi_n|^2 = 1$. The Fock basis comes with canonical adjoint operators $\hat{a}$ and $\hat{a}^\dagger$ referred to as annihilation and creation operators, respectively, or photon subtraction and photon addition operators in the context of optical quantum information processing. These operators are defined by their action on the Fock basis as

$$\begin{aligned} &\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle\,, && \text{for } n \in \mathbb{N}^*, \\ &\hat{a}\,|0\rangle = 0, && \\ &\hat{a}^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle\,, && \text{for } n \in \mathbb{N}, \end{aligned} \tag{1.17}$$

and follow the canonical commutation relation

$$[\hat{a}, \hat{a}^\dagger] = \mathbb{1}, \tag{1.18}$$

where $\mathbb{1}$ is the identity operator. The eigenstates of the annihilation operator are the coherent states $\{|\alpha\rangle\}_{\alpha\in\mathbb{C}}$, defined as

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n\geq 0} \frac{\alpha^n}{\sqrt{n!}} |n\rangle, \tag{1.19}$$

for all $\alpha \in \mathbb{C}$. Alternatively, defining the displacement operator as

$$\hat{D}(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}, \tag{1.20}$$

for all $\alpha \in \mathbb{C}$, the coherent state of amplitude $\alpha \in \mathbb{C}$ is obtained from the vacuum state as

$$|\alpha\rangle = \hat{D}(\alpha)\,|0\rangle\,. \tag{1.21}$$

The inner product of two coherent states $|\alpha\rangle$ and $|\beta\rangle$ is given by

$$\langle\alpha|\beta\rangle = e^{\alpha^*\beta - \frac{1}{2}(|\alpha|^2+|\beta|^2)}, \tag{1.22}$$

for all $\alpha, \beta \in \mathbb{C}$. In particular, two coherent states have nonzero overlap. These states form an overcomplete family:

$$\int_{\alpha\in\mathbb{C}} |\alpha\rangle\langle\alpha| \frac{d^2\alpha}{\pi} = \mathbb{1}, \tag{1.23}$$

where $d^2\alpha = d\Re(\alpha)d\Im(\alpha)$. The canonical position-like and momentum-like operators $\hat{q}$ and $\hat{p}$ are defined as

$$\begin{aligned} \hat{q} &= \frac{1}{\sqrt{2}}(\hat{a} + \hat{a}^\dagger), \\ \hat{p} &= \frac{1}{i\sqrt{2}}(\hat{a} - \hat{a}^\dagger). \end{aligned} \tag{1.24}$$

These hermitian operators, also referred to as quadrature operators in the context of optical quantum information processing, follow the canonical commutation relation

$$[\hat{q}, \hat{p}] = i\mathbb{1}. \tag{1.25}$$

They satisfy Heisenberg uncertainty principle [Hei85]

$$\sigma_{\hat{q}}\sigma_{\hat{p}} \geq \frac{1}{2}, \tag{1.26}$$

where $\sigma_{\hat{q}}$ and $\sigma_{\hat{p}}$ denote the standard deviation of position and momentum, respectively, i.e., they cannot be measured both with arbitrary precision for the same quantum state: measuring one randomises the other.

The eigenstates of $\hat{q}$ (resp. $\hat{p}$) form a continuous family of unnormalisable states $\{|q\rangle\}_{q\in\mathbb{R}}$ (resp. $\{|p\rangle\}_{p\in\mathbb{R}}$), thus technically lying outside of the Hilbert space. These states may be treated formally as an infinite uncountable basis of the Hilbert space, the so-called position basis (resp. momentum basis). Expanding a single-mode pure state $|\psi\rangle$ in the position basis gives

$$|\psi\rangle = \int_{q\in\mathbb{R}} \psi(q)|q\rangle\, dq, \tag{1.27}$$

where $\psi(q) = \langle q|\psi\rangle$ is the position wave function of the state $|\psi\rangle$, with the normalisation condition for the position probability distribution $\int_{q\in\mathbb{R}} |\psi(q)|^2 dq = 1$. A similar expansion holds in the momentum basis with the momentum wave function. The position and momentum bases are related by a Fourier transform:

$$|q\rangle = \frac{1}{\sqrt{2\pi}} \int_{p\in\mathbb{R}} e^{-iqp}\,|p\rangle\, dp, \tag{1.28}$$

and

$$|p\rangle = \frac{1}{\sqrt{2\pi}} \int_{q\in\mathbb{R}} e^{iqp}\,|q\rangle\, dq. \tag{1.29}$$

Note that the Fock state $|n=0\rangle$ and the coherent state $|\alpha=0\rangle$ are equal, but different from the position state $|q=0\rangle$ and the momentum state $|p=0\rangle$, themselves distinct.

1.2 Phase space formalism

We refer the reader to [CG69b, CG69a] for an introduction to the material presented in this section. In particular, we restrict to single-mode states and operators.

The expectation values of the position and momentum operators lie in the so-called phase space, which is the quantum analogue of classical phase space. Continuous variable quantum states and operators can be alternatively described by a phase space representation. This formulation identifies a quantum state with a normalised distribution over phase space.

This allows for a simple and experimentally relevant classification of quantum states: those with a Gaussian phase space distribution are called Gaussian states and the others non-Gaussian states. By extension, operations mapping Gaussian states to Gaussian states are also called Gaussian. These Gaussian operations and states are the ones implementable with linear optics and quadratic non-linearities [BvL05], and are hence relatively easy to construct experimentally.

Hereafter, we identify the single-mode phase space with $\mathbb{C}$, where the real part corresponds to expectation values of the position operator and the imaginary part to expectation values momentum operator. We adopt the convention $\alpha = \frac{1}{\sqrt{2}}(q+ip) \in \mathbb{C}$, with $\frac{d^2\alpha}{\pi} = \frac{d\Re(\alpha)d\Im(\alpha)}{\pi} = \frac{dqdp}{2\pi}$.

There exists a continuum of equivalent phase space distributions representing the same operator in phase space. This continuum of representations is parametrized by a real parameter $s \leq 1$. For all $s \leq 1$, let us define the operator

$$\hat{T}(\alpha,s) := \int_{\beta\in\mathbb{C}} \hat{D}(\beta)\exp\left(\alpha\beta^* - \alpha^*\beta + \frac{s}{2}|\beta|^2\right)\frac{d^2\beta}{\pi}, \tag{1.30}$$

for all $\alpha \in \mathbb{C}$. The phase space representation with parameter s of an operator $\hat{O}$ is defined as

$$W_{\hat{O}}(\alpha,s) = \mathrm{Tr}\left[\hat{T}(\alpha,s)\hat{O}\right]. \tag{1.31}$$

This expression should be treated formally for unbounded operators and the case $s = 1$ should be understood as the limit $s \to 1^-$. The same definition holds for density operators, in which case the representation is real-valued and corresponds to the expectation value of the operator $\hat{T}$. The phase space representations are normalised as

$$\int_{\alpha\in\mathbb{C}} W_\rho(\alpha,s)\frac{d^2\alpha}{\pi} = \mathrm{Tr}(\rho), \tag{1.32}$$

for any density operator ρ and any $s \leq 1$. As the parameter s decreases, the phase space representation smoothens. This is captured by the following relation:

$$W(\alpha,s) = \frac{2}{t-s}\int_{\beta\in\mathbb{C}} W(\beta,t)\exp\left(-\frac{2|\alpha-\beta|^2}{t-s}\right)\frac{d^2\beta}{\pi}, \tag{1.33}$$

for all $s < t \leq 1$, i.e., the representation with lower parameter is obtained from the representation with higher parameter by a Gaussian convolution. In particular, if one representation is a Gaussian function, then all representations are Gaussian. Moreover, for all operators $\hat{O}_1$ and $\hat{O}_2$,

$$\mathrm{Tr}\left(\hat{O}_1\hat{O}_2\right) = \int_{\alpha\in\mathbb{C}} W_{\hat{O}_1}(\alpha,-s)W_{\hat{O}_2}(\alpha,s)\frac{d^2\alpha}{\pi}, \tag{1.34}$$

for all $s \in [-1,1]$. This important property allows one to retrieve information about quantum systems by probing their phase space representation: if one of the two operators in the above equation is a density operator, the expectation value is obtained as

$$\mathrm{Tr}\left(\hat{O}\rho\right) = \int_{\alpha\in\mathbb{C}} W_{\hat{O}}(\alpha,-s)W_\rho(\alpha,s)\frac{d^2\alpha}{\pi}, \tag{1.35}$$

for all $s \in [-1,1]$.

In what follows, we detail some properties of the three most prominent representations in the literature: the Wigner W function [Wig97], the Glauber–Sudarshan P function [Sud63, Gla63] and the Husimi Q function [Hus40], corresponding to the values $s = 0$, $s = 1$ and $s = -1$,

respectively (Fig. 1.1). In particular, we will make extensive use of the Husimi representation throughout the first chapters of the thesis. We adopt the normalising conventions

$$\begin{aligned} W(\alpha) &= \frac{1}{\pi} W(\alpha, 0), \\ P(\alpha) &= \frac{1}{\pi} W(\alpha, 1), \\ Q(\alpha) &= \frac{1}{\pi} W(\alpha, -1), \end{aligned} \tag{1.36}$$

for all $\alpha \in \mathbb{C}$, so that the W, P and Q functions are normalised to 1 for normalised states (note the difference of normalisation with [CG69a] for the Wigner and Husimi functions).

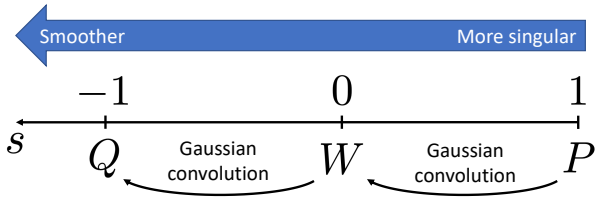

Figure 1.1: A pictorial representation of the continuum of phase space representations.

1.2.1 Wigner W function

The Wigner function is a nonsingular distribution for all states and is referred to as a quasiprobability distribution, as it is a normalised distribution which can take negative values. This contrasts with classical phase space probability distributions.

By virtue of Hudson's theorem [Hud74, SC83], a pure quantum state is non-Gaussian if and only if its Wigner function has negative values. In other words, if a pure quantum state has a positive Wigner function, then it is a Gaussian state. Various notions relating to negativity of the Wigner function have been introduced for measuring how much non-Gaussian a quantum state is [KŻ04, AGPF18].

The Wigner function can be expressed as [Roy77]

$$W_{\hat{O}}(\alpha) = \frac{2}{\pi} \mathrm{Tr}\left[\hat{D}(\alpha)\hat{\Pi}\hat{D}^\dagger(\alpha)\hat{O}\right], \tag{1.37}$$

for all $\alpha \in \mathbb{C}$ and for any operator $\hat{O}$, where $\hat{\Pi} = (-1)^{\hat{a}^\dagger \hat{a}} = \sum_{n\geq 0} (-1)^n |n\rangle\langle n|$ is the parity operator and $\hat{D}(\alpha) = e^{\alpha \hat{a}^\dagger - \alpha^* \hat{a}}$ is a displacement operator of amplitude $\alpha \in \mathbb{C}$. In particular, the Wigner function of a quantum state is related to the expectation value of displaced parity operators.

1.2.2 Glauber–Sudarshan P function

The Glauber–Sudarshan P function is the most singular phase space representation. For quantum states, it is actually always a singular distribution.

The P function gives a convenient diagonal representation of a state in coherent state basis as

$$\rho = \int_{\alpha\in\mathbb{C}} P_\rho(\alpha)\,|\alpha\rangle\langle\alpha|\,d^2\alpha, \tag{1.38}$$

and this representation is unique. The P function can be expressed formally as [Meh67]

$$P_{\hat{O}}(\alpha) = \frac{e^{|\alpha|^2}}{\pi}\int_{\beta\in\mathbb{C}} \langle -\beta|\hat{O}|\beta\rangle \exp\left(\alpha\beta^* - \alpha^*\beta + |\beta|^2\right)\frac{d^2\beta}{\pi}, \tag{1.39}$$

for all $\alpha\in\mathbb{C}$ and for any operator $\hat{O}$.

1.2.3 Husimi Q function

The Husimi Q function is a smoother version of the Wigner function and the Glauber–Sudarshan P function. It is given by

$$Q_{\hat{O}}(\alpha) = \frac{1}{\pi}\langle\alpha|\hat{O}|\alpha\rangle, \tag{1.40}$$

for all $\alpha\in\mathbb{C}$ and for any operator $\hat{O}$, where $|\alpha\rangle$ is the coherent state of amplitude $\alpha\in\mathbb{C}$. The Husimi Q function of a state thus is always nonnegative and normalised. However, it does not represent probabilities of mutually exclusive states since the overlap between two coherent states is always nonzero.

For any state ρ and any operator $\hat{O}$ we have, with Eq. (1.41), the so-called optical equivalence theorem for antinormal ordering:

$$\mathrm{Tr}\left(\hat{O}\rho\right) = \pi\int_{\alpha\in\mathbb{C}} Q_\rho(\alpha)P_{\hat{O}}(\alpha)\,d^2\alpha. \tag{1.41}$$

Hudson's theorem may be formulated as follows for the Husimi function [LB95]: a pure quantum state is non-Gaussian if and only if its Husimi function has zeros. In other words, a pure quantum state is non-Gaussian if and only if it is orthogonal to at least one coherent state.

1.3 Gaussian states and processes

Gaussian states and processes have been defined in the previous section, the former as the states having a Gaussian phase space representation and the latter as the processes mapping Gaussian states to Gaussian states. Ubiquitous in quantum physics, they are well understood theoretically [FOP05, WPGP$^+$12, ARL14] and routinely implemented experimentally [GCP07].

We review Gaussian processes and states in the following sections, restricting to pure states, unitary operations and projectors.

1.3.1 Gaussian unitary operations

The displacement operator of amplitude $\alpha\in\mathbb{C}$ has been introduced in the previous section and reads

$$\hat{D}(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}. \tag{1.42}$$

It satisfies the relations

$$\begin{aligned}
&\hat{D}^\dagger(\alpha) = \hat{D}(-\alpha),\\
&\hat{D}(\alpha)\hat{a}\hat{D}^\dagger(\alpha) = \hat{a} - \alpha\mathbb{1},\\
&\hat{D}(\alpha)\hat{a}^\dagger\hat{D}^\dagger(\alpha) = \hat{a}^\dagger - \alpha^*\mathbb{1},\\
&\hat{D}(\alpha)\hat{D}(\beta) = e^{\frac{1}{2}(\alpha\beta^* - \alpha^*\beta)}\hat{D}(\alpha+\beta),
\end{aligned} \tag{1.43}$$

for all $\alpha, \beta \in \mathbb{C}$. We denote a tensor product of m single-mode displacements by $\hat{D}(\boldsymbol{\alpha}) = \bigotimes_{i=1}^m \hat{D}(\alpha_i)$ for all $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m) \in \mathbb{C}$.

The squeezing operator is defined as

$$\hat{S}(\xi) = e^{\frac{1}{2}(\xi\hat{a}^2 - \xi^*\hat{a}^{\dagger 2})}, \tag{1.44}$$

for all $\xi \in \mathbb{C}$. The parameter ξ is called squeezing parameter. The squeezing operator satisfies the relations

$$\begin{aligned}
&\hat{S}^\dagger(\xi) = \hat{S}(-\xi),\\
&\hat{S}(\xi)\hat{a}\hat{S}^\dagger(\xi) = \cosh r\,\hat{a} + e^{-i\theta}\sinh r\,\hat{a}^\dagger,\\
&\hat{S}(\xi)\hat{a}^\dagger\hat{S}^\dagger(\xi) = \cosh r\,\hat{a}^\dagger + e^{i\theta}\sinh r\,\hat{a},
\end{aligned} \tag{1.45}$$

for all $\xi = re^{i\theta} \in \mathbb{C}$. We denote a tensor product of m single-mode squeezings by $\hat{S}(\boldsymbol{\xi}) = \bigotimes_{i=1}^m \hat{S}(\xi_i)$ for all $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_m) \in \mathbb{C}$.

The displacement and squeezing operators may be conceived as acting on a state by displacing and squeezing its phase space representation, respectively, as their name indicates. This geometrical intuition holds in particular for the Wigner quasiprobability distribution.

Any single-mode Gaussian unitary operation may be written as a squeezing and a displacement operator. The ordering is only a convention, since the displacement and squeezing operators satisfy the braiding relation [NT97]

$$\hat{D}(\alpha)\hat{S}(\xi) = \hat{S}(\xi)\hat{D}(\gamma), \qquad \gamma = \alpha\cosh r + \alpha^* e^{-i\theta}\sinh r, \tag{1.46}$$

for all $\alpha \in \mathbb{C}$ and all $\xi = re^{i\theta} \in \mathbb{C}$.

Passive linear transformation over m modes are defined as the unitary transformations $\hat{U}$ which act unitarily on the creation operators of the modes $\hat{a}_1^\dagger, \ldots, \hat{a}_m^\dagger$ as well as on the annihilation operators $\hat{a}, \ldots, \hat{a}_m$. Any such transformation $\hat{U}$ is associated to an $m \times m$ unitary matrix U which transforms the creation operators of the modes as

$$\begin{pmatrix} \hat{a}_1^\dagger \\ \vdots \\ \hat{a}_m^\dagger \end{pmatrix} \to U \begin{pmatrix} \hat{a}_1^\dagger \\ \vdots \\ \hat{a}_m^\dagger \end{pmatrix}, \tag{1.47}$$

and the annihilation operators of the modes as

$$\begin{pmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_m \end{pmatrix} \to U^* \begin{pmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_m \end{pmatrix}. \tag{1.48}$$

These transformations map the multimode vacuum state onto itself.

Finally, Gaussian projectors are identified with projections onto Gaussian pure states, which we review in what follows.

1.3.2 Single-mode Gaussian pure states

General single-mode Gaussian pure states are obtained from the vacuum with a Gaussian unitary operation. They are the squeezed coherent states (or alternatively the displaced squeezed vacuum states):

$$\hat{S}(\xi)\hat{D}(\alpha)|0\rangle\,, \tag{1.49}$$

for $\alpha,\xi \in \mathbb{C}$. Setting $\xi = 0$ we obtain a coherent state of amplitude $\alpha \in \mathbb{C}$, while setting $\alpha = 0$ we obtain a squeezed vacuum state with squeezing parameter $\xi \in \mathbb{C}$.

The phase space representation of a coherent state is a Gaussian displaced in phase space, while the phase space representation of a squeezed vacuum state is a Gaussian centered at 0, squeezed in a direction depending on the phase of the squeezing parameter. The strength of the squeezing depends on the modulus of the squeezing parameter (Fig. 1.2). In particular, position and momentum eigenstates can be conceived formally as infinitely squeezed vacuum states, displaced by a finite amplitude [SEMC13].

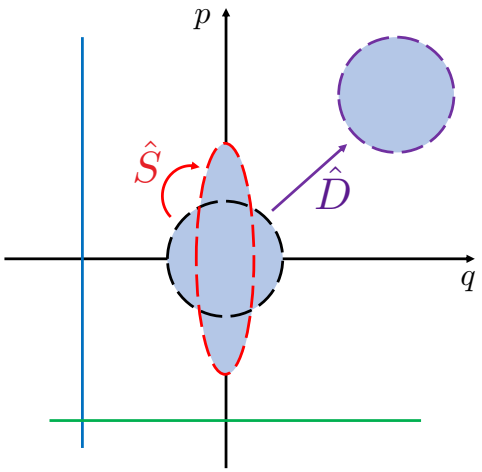

Figure 1.2: A pictorial representation of Gaussian states and processes in phase space. Circles are normalised Gaussian probability distributions—coherent states—viewed from the top and the ellipse represents a squeezed Gaussian probability distribution—a squeezed state. The vertical blue line and the horizontal green line correspond to position and momentum eigenstates, respectively.

1.3.3 Multimode case: the symplectic formalism

We present a short introduction to the symplectic formalism and refer to [ARL14] for a detailed exposition.

Any m-mode Gaussian state $\boldsymbol{\rho}$ can be described by a $2m \times 2m$ covariance matrix $\boldsymbol{V}^{\mathbb{R}}$ containing its second canonical moments and a displacement vector $\boldsymbol{d}^{\mathbb{R}}$ of size m containing its first canonical moments. The coefficients of the covariance matrix are defined, for $k,l \in \{1,\dots,2m\}$, by $V^{\mathbb{R}}_{kl} = \frac{1}{2}\langle R_k R_l + R_l R_k\rangle_\rho - \langle R_k\rangle_\rho \langle R_l\rangle_\rho$ where $\boldsymbol{R} = (\hat{q}_1,\dots,\hat{q}_m,\hat{p}_1,\dots,\hat{p}_m)$. The coefficients of the displacement vector are given by $d^{\mathbb{R}}_j = \langle R_j\rangle_\rho$ for all $j \in \{1,\dots,2m\}$. Alternatively and more conveniently, one can describe covariance matrices and displacement vectors in the complex basis $\boldsymbol{\lambda} = (\hat{a}_1,\dots,\hat{a}_m,\hat{a}_1^\dagger,\dots,\hat{a}_m^\dagger)$. We write $\boldsymbol{V}$ and $\tilde{\boldsymbol{d}}$ the covariance matrix and displacement vector in that basis, with

$$\boldsymbol{V} = \Omega \boldsymbol{V}^{\mathbb{R}} \Omega^\dagger, \qquad \tilde{\boldsymbol{d}} = \Omega \boldsymbol{d}^{\mathbb{R}}, \tag{1.50}$$

where

$$\Omega = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1}_m & i\mathbb{1}_m \\ \mathbb{1}_m & -i\mathbb{1}_m \end{pmatrix}. \tag{1.51}$$

The complex covariance matrix has the structure

$$\boldsymbol{V} = \begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix}, \tag{1.52}$$

with $A = A^\dagger$ and $B = B^T$, so that $\boldsymbol{V}^\dagger = \boldsymbol{V}$. The displacement vector has the structure

$$\tilde{\boldsymbol{d}} = \begin{pmatrix} \boldsymbol{d} \\ \boldsymbol{d}^* \end{pmatrix}. \tag{1.53}$$

We will also refer to the above vector $\boldsymbol{d}$ as the displacement vector.

Gaussian multimode unitary operations are generated by Hamiltonians that are at most quadratic in the annihilation and creation operators of the modes. As a consequence, they induce affine transformations of the annihilation and creation operators which preserve their canonical commutation relations, i.e., symplectic linear transfomations, together with displacements. The evolution of a Gaussian state during a Gaussian evolution (excluding displacements) is described by a complex symplectic transformation of its complex covariance matrix and its complex displacement vector:

$$(\boldsymbol{V}, \tilde{\boldsymbol{d}}) \to (S\boldsymbol{V}S^\dagger, S\tilde{\boldsymbol{d}}), \tag{1.54}$$

where a complex symplectic matrix S satisfies

$$S\Omega J \Omega^\dagger S^\dagger = \Omega J \Omega^\dagger, \tag{1.55}$$

where $J = \begin{pmatrix} \mathbb{0}_m & \mathbb{1}_m \\ -\mathbb{1}_m & \mathbb{0}_m \end{pmatrix}$ and where the matrix Ω is defined in Eq. (1.51). We will use the notations

$$S_{\boldsymbol{\xi}} \equiv \begin{pmatrix} D_c(\boldsymbol{\xi}) & D_s(\boldsymbol{\xi}) \\ D_s(\boldsymbol{\xi}) & D_c(\boldsymbol{\xi}) \end{pmatrix} \tag{1.56}$$

for all $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_m) \in \mathbb{C}^m$, with $D_c(\boldsymbol{\xi}) = \mathrm{Diag}(c_{\xi_1}, \ldots, c_{\xi_m})$ and $D_s(\boldsymbol{\xi}) = \mathrm{Diag}(s_{\xi_1}, \ldots, s_{\xi_m})$, where $c_\chi = \cosh\chi$ and $s_\chi = \sinh\chi$ for the symplectic matrices that implement squeezing and

$$S_U \equiv \begin{pmatrix} U^* & \mathbb{0}_m \\ \mathbb{0}_m & U \end{pmatrix} \tag{1.57}$$

for the symplectic matrix associated with a passive linear transformations with $m \times m$ unitary matrix U. A displacement does not affect the covariance matrix and only translates the displacement vector.

The so-called Bloch-Messiah or Euler decomposition implies that any $2m \times 2m$ complex symplectic matrix can be written as $S_U S_{\boldsymbol{\xi}} S_V$ for some $m \times m$ unitary matrices U and V and some squeezing parameters $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_m) \in \mathbb{C}^m$. In particular, any multimode Gaussian unitary operation can be decomposed as a passive linear transformation followed by a product of single-mode squeezings, followed by another passive linear transformation, together with single-mode displacements.

Since any multimode Gaussian pure quantum state may be engineered from the vacuum with a Gaussian unitary operation, by virtue of Williamson decomposition, and since the vacuum is mapped onto itself by passive linear transformations, this means that any multimode Gaussian pure quantum state can be written as a tensor product of single-mode Gaussian states (displaced squeezed vacuum states) followed by a single passive linear transformation.

1.4 Linear optics

Linear optics covers the manipulation of light by unitary transformations whose exponent is at most quadratic in the field operator [WM07], i.e., Gaussian unitaries. It induces transformations of quantum states of light which are divided in two categories, passive and active transformations, depending on whether these transformations change the total number of photons of the input state. In what follows, we review a few examples of quantum states of light and quantum optical measurements, and we detail passive linear optical transformations, implemented by unitary interferometers, with the examples of the Hong-Ou-Mandel effect [HOM87] and its generalisation Boson Sampling [AA13].

1.4.1 Quantum states of light

We briefly list single-mode quantum states that are common in the literature, some of which were already introduced in the previous sections, and which we will encounter in the following chapters.

- Photon-number states: these states form the orthonormal Fock basis and are obtained from the vacuum as

$$|n\rangle = \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}} |0\rangle , \tag{1.58}$$

for all $n \in \mathbb{N}$. Taking $n = 0$ gives the vacuum state and photon-number states are non-Gaussian for $n > 0$. They are the eigenstates of the photon-number operator $\hat{n} = \hat{a}^\dagger \hat{a}$: for all $n \in \mathbb{N}$, $\hat{n}\,|n\rangle = n\,|n\rangle$.

- Coherent states: these Gaussian states are expressed as

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n\geq 0} \frac{\alpha^n}{\sqrt{n!}} |n\rangle, \tag{1.59}$$

for all $\alpha \in \mathbb{C}$. These states are a good approximation of the quantum state of a laser and are sometimes referred to as classical states, because their behaviour resembles that of a classical harmonic oscillator. They are the eigenstates of the annihilation operator $\hat{a}$: for all $\alpha \in \mathbb{C}$, $\hat{a}\,|\alpha\rangle = \alpha\,|\alpha\rangle$.

- Squeezed vacuum states: these Gaussian states are expressed as

$$|\xi\rangle = \frac{1}{\sqrt{\cosh r}} \sum_{n\geq 0} (-e^{-i\theta} \tanh r)^n \frac{\sqrt{(2n)!}}{2^n n!} |2n\rangle, \tag{1.60}$$

for all $\xi = re^{i\theta} \in \mathbb{C}$. They display reduced variance for one quadrature, but increased variance for the conjugate quadrature, in accordance with the uncertainty principle.

- Photon-subtracted/added states: these states are obtained by applying the annihilation/creation operator to a state (and renormalising). These unphysical operations cannot be implemented determinisically and are implemented probabilistically in practice. For example, a photon subtraction may be implemented by mixing the input state with the vacuum on a beam splitter with near unity reflectance. Then, conditioned on a successful single-photon heralding of the transmitted light, the reflected state has been photon-subtracted.

- Cat states: named after Schrödinger's cat, these states are superpositions of two coherent states of equal amplitudes, $|\alpha\rangle$ and $|-\alpha\rangle$, like the cat in Schrödinger's thought experiment is in a superposition of two classical states, dead and alive. Varying the relative phase between the coherent states in the superposition gives different cat states. In particular, we introduce the cat$^+$ and cat$^-$ states:

$$|\mathrm{cat}_\alpha^\pm\rangle = \frac{1}{\sqrt{\mathcal{N}_\alpha^\pm}}(|\alpha\rangle \pm |-\alpha\rangle), \tag{1.61}$$

for all $\alpha \in \mathbb{C}$, where $\mathcal{N}_\alpha^\pm = 2(1 \pm e^{-2|\alpha|^2})$ is a normalisation factor.

- GKP states: finally, let us mention the Gottesman-Kitaev-Preskill (GKP) states which form a family of unphysical states with periodic wave functions [GKP01]. These states are formal periodic superpositions of infinitely squeezed states and their physical approximations have applications for continuous variable quantum error correction [TBMS20].

1.4.2 Quantum optical measurements

We list various (idealised) single-mode measurements in what follows: homodyne detection, balanced heterodyne detection, unbalanced heterodyne detection, single-photon threshold detection, photon number parity detection and photon-number resolving detection. Detailed information on these detection methods can be found, e.g., in [FOP05]. We will only consider multimode detections that are tensor products of such single-mode detections.

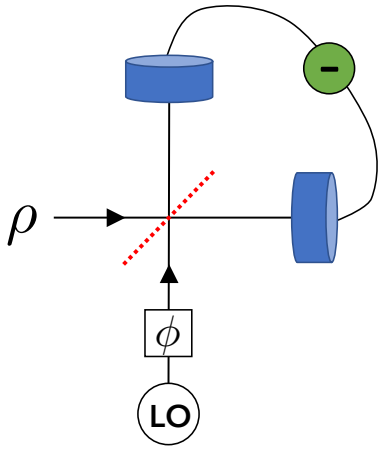

Figure 1.3: A schematic representation of homodyne detection of a state ρ. The dashed red line represents a balanced beamsplitter. LO stands for local oscillator, i.e., strong coherent state. The blue circles are photodiode detectors. Changing the phase ϕ of the local oscillator allows one to measure rotated quadratures.

Homodyne detection consists in a Gaussian measurement of a quadrature of the field, by mixing the state to be measured on a balanced beam splitter with a strongly excited coherent state, the local oscillator. Then, the intensities of both output arms are measured and their difference yields a value proportional to a quadrature of the input mode, rotated depending on the phase of the local oscillator (Fig. 1.3). The POVM elements for homodyne detection with phase ϕ are given by

$$\Pi_x^\phi = |x\rangle_\phi\langle x| \tag{1.62}$$

for all $x \in \mathbb{R}$, where $|x\rangle_\phi$ is the eigenstate of the rotated quadrature operator $\hat{x}_\phi = \cos\phi\,\hat{q} + \sin\phi\,\hat{p}$ corresponding to the eigenvalue $x \in \mathbb{R}$.

Balanced heterodyne detection, also called double homodyne or eight-port homodyne [FOP05], consists in splitting the measured state with a balanced beam splitter and measuring both ends with homodyne detection. This corresponds to a joint noisy measurement of quadratures $\hat{q}$ and $\hat{p}$. This is a Gaussian measurement which yields two real outcomes, corresponding to the real and

imaginary parts of $\alpha \in \mathbb{C}$. The POVM elements for balanced heterodyne detection are given by

$$\Pi_\alpha = \frac{1}{\pi} |\alpha\rangle\langle\alpha|, \tag{1.63}$$

for all $\alpha \in \mathbb{C}$, where $|\alpha\rangle$ is the coherent state of amplitude $\alpha \in \mathbb{C}$. Measuring a state with balanced heterodyne detection effectively amounts to sampling from its Q function.

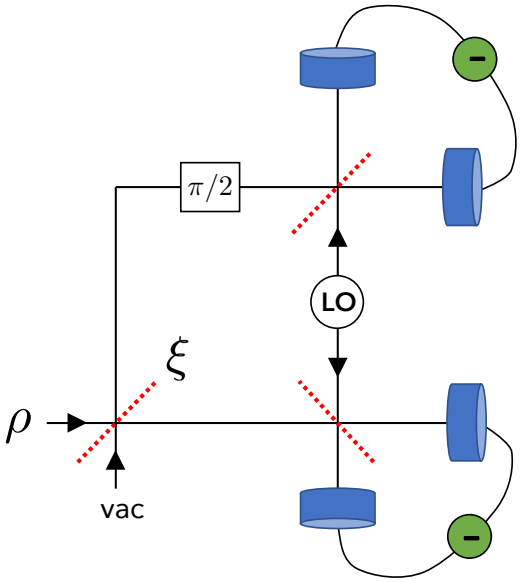

Figure 1.4: Schematic representation of unbalanced heterodyne detection with unbalancing parameter $\xi = re^{i\theta} \in \mathbb{C}$. LO stands for local oscillator, i.e., strong coherent state. The blue circles are photodiode detectors. The $\hat{q}$ and $\hat{p}$ measurements are each performed by balanced homodyne detection.

A straightforward generalisation is unbalanced heterodyne detection (Fig. 1.4), where the input beam splitter is no longer balanced but characterized instead by a reflectance R and a transmittance T, with $R^2 + T^2 = 1$. The POVM elements for unbalanced heterodyne detection with unbalancing parameter $\xi \in \mathbb{C}$ are given by

$$\Pi_\alpha^\xi = \frac{1}{\pi} |\alpha, \xi\rangle\langle\alpha, \xi|, \tag{1.64}$$

for all $\alpha \in \mathbb{C}$, where $|\alpha, \xi\rangle = \hat{S}(\xi)\hat{D}(\alpha)|0\rangle$ is a squeezed coherent state. Writing $\xi = re^{i\theta}$, the unbalalancing parameter is related to the optical setup by $r = \left|\log\left(\frac{T}{R}\right)\right|$, with θ being the phase of the local oscillator [CDM+17]. Measuring a state with unbalanced heterodyne detection effectively amounts to sampling from its squeezed Q function. Setting $\xi = 0$ gives balanced heterodyne detection, while sending $|\xi| = r$ to infinity gives homodyne detection. Any Gaussian measurement can thus be implemented by Gaussian unitary operations and heterodyne detection only, since it can be implemented by Gaussian unitary operations and homodyne detection only [GC02, EP03].

Additionnally, we introduce three non-Gaussian measurements, each giving more information about the photon number of the measured state. The first is single-photon threshold detection [Had09], or simply threshold detection, whose POVM elements are given by

$$\Pi_0 = |0\rangle\langle 0|, \quad \Pi_1 = \mathbb{1} - |0\rangle\langle 0|. \tag{1.65}$$

This binary measurement only distinguishes the vacuum state from other states. The second is photon number parity detection [HBR07], or simply parity detection, whose POVM elements are given by

$$\Pi_+ = \sum_{n\geq 0} |2n\rangle\langle 2n|, \quad \Pi_- = \sum_{n\geq 0} |2n+1\rangle\langle 2n+1|. \tag{1.66}$$

This is a binary measurement of the parity operator $\hat{\Pi} = (-1)^{\hat{a}^\dagger \hat{a}}$ yielding, as its name indicates, the parity of the number of photons of the measured state. The third is photon number-resolving detection [DMB$^+$08], whose POVM elements are given by

$$\Pi_n = |n\rangle\langle n|, \tag{1.67}$$

for all $n \in \mathbb{N}$, i.e., projections onto Fock states.

1.4.3 Linear interferometers

Linear optical unitary interferometers are composed of beam splitters and phase shifters and implement passive linear transformations of the modes. In particular, any passive linear transformation $\hat{U}$ over m modes with $m \times m$ unitary matrix U can be implemented by a linear interferometer with at most $\frac{m(m-1)}{2}$ balanced beam splitters and m phase shifters [RZBB94]. The corresponding unitary interferometer is described by the same unitary matrix $U = (u_{ij})_{1\leq i,j\leq m}$. Unlike in the circuit picture, the matrix U does not act on the computational basis, which in this case is the infinite multimode Fock basis, but rather describes the linear evolution of the creation operator of each mode. More precisely,

$$\begin{pmatrix} \hat{a}_1^\dagger \\ \vdots \\ \hat{a}_m^\dagger \end{pmatrix} \to U \begin{pmatrix} \hat{a}_1^\dagger \\ \vdots \\ \hat{a}_m^\dagger \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^m u_{1k} \hat{a}_k^\dagger \\ \vdots \\ \sum_{k=1}^m u_{mk} \hat{a}_m^\dagger \end{pmatrix}. \tag{1.68}$$

In that picture, the direct sum plays the role of the tensor product in the computational basis: taking the direct sum of two unitaries corresponds to putting linear optical elements in parallel, while multiplying unitaries corresponds to putting linear optical elements in sequence.

Multimode coherent states have a specific evolution through linear interferometers: they are mapped onto coherent states and do not become entangled, unlike other states. If U is the unitary matrix describing an interferometer which implements a passive linear transformation $\hat{U}$, an input coherent state $|\boldsymbol{\alpha}\rangle$ is mapped to an output coherent state $\hat{U}|\boldsymbol{\alpha}\rangle = |U\boldsymbol{\alpha}\rangle$, where the vector of

output amplitudes $U\boldsymbol{\alpha}$ is obtained by multiplying the vector of input amplitudes $\boldsymbol{\alpha}$ by the unitary matrix U.

Remarkable quantum effects may be witnessed when the input to linear optical unitary interferometers are single-photon Fock states instead of coherent states. The celebrated Knill–Laflamme–Milburn scheme [KLM01] shows that single photons and linear optics are enough to achieve universal quantum computing together with adaptive measurements (making the rest of the computation depend on the result of intermediate measurements). Already without adaptive measurements, interesting effects can be observed. We give two notable examples in the following sections: the Hong–Ou–Mandel effect and Boson Sampling.

1.4.4 Hong–Ou–Mandel effect

The Hong–Ou–Mandel effect, or photon bunching, refers to the bosonic behaviour of indistinguishable photons which bunch together when mixed on a balanced beamsplitter (Fig. 1.5). A balanced beam splitter is a unitary interferometer over two modes, with unitary matrix

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{1.69}$$

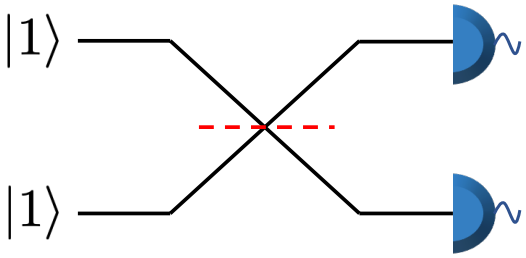

Figure 1.5: Hong-Ou-Mandel effect. The dashed red line represents a balanced beam splitter. The number of photons is detected for both output arms. If the input single photons are indistinguishable, the outcomes (20) and (02) occur with the same probability $\frac{1}{2}$ and the outcome (11) never occurs.

The input state is composed of two single photons, one in each mode. Labelling the modes u and d, for 'up' and 'down', let $\hat{a}_u^\dagger$, $\hat{a}_d^\dagger$ and $\hat{b}_u^\dagger$, $\hat{b}_d^\dagger$ be the creation operators of the input and output modes, respectively. The balanced beam splitter acts on the input creation operators as

$$\begin{pmatrix} \hat{b}_u^\dagger \\ \hat{b}_d^\dagger \end{pmatrix} = H \begin{pmatrix} \hat{a}_u^\dagger \\ \hat{a}_d^\dagger \end{pmatrix}. \tag{1.70}$$

The input state thus evolves as

$$\begin{aligned} |11\rangle &= \hat{a}_u^\dagger \hat{a}_d^\dagger |00\rangle \\ &\xrightarrow{H} \frac{1}{2}(\hat{b}_u^\dagger + \hat{b}_d^\dagger)(\hat{b}_u^\dagger - \hat{b}_d^\dagger)|00\rangle \\ &= \frac{1}{2}(\hat{b}_u^{\dagger 2} - \hat{b}_d^{\dagger 2})|00\rangle \\ &= \frac{1}{\sqrt{2}}(|20\rangle - |02\rangle), \end{aligned} \tag{1.71}$$

where we used $\hat{b}_u^\dagger \hat{b}_d^\dagger = \hat{b}_d^\dagger \hat{b}_u^\dagger$. In particular, measuring the photon number in both output modes will always yield 0 for one of the modes: the outcome (11) is never witnessed if the photons are indistinguishable, i.e., the photons have bunched together.

1.4.5 Boson Sampling

Let $m \in \mathbb{N}^*$ and $n \in \mathbb{N}$, with $m \geq n$. Boson Sampling, introduced in [AA13], is a generalisation of the Hong–Ou–Mandel setup, where the balanced beam splitter is replaced by a general unitary interferomer over m modes with $m \times m$ unitary matrix U and the input is composed of n single photons in the first n modes and vacuum in the remaining $m-n$ modes, the photon number of all output modes being measured (Fig. 1.6).

Even though Boson Sampling has been formulated for general bosonic particles, linear optics provides a convenient way of looking at it. Boson Sampling is a subuniversal model of quantum computation, believed to be hard to simulate by classical computers while not possessing the computational power of a universal quantum computer. We review this model in what follows and we refer to [AA13] for a detailed version of the material presented in this section. In particular, we do not discuss the theoretical use of postselection.

We denote photon number states over m modes by

$$|\boldsymbol{s}\rangle = |s_1 \dots s_m\rangle = \frac{(\hat{a}_1^\dagger)^{s_1}}{\sqrt{s_1!}} \cdots \frac{(\hat{a}_m^\dagger)^{s_m}}{\sqrt{s_m!}} |0\rangle^{\otimes m}, \tag{1.72}$$

where s_k and $\hat{a}_k^\dagger$ are respectively the number of photons and the creation operator of the k^{th} mode. We identify these states with m-tuples of integers $\boldsymbol{s} = (s_1, \dots, s_m) \in \mathbb{N}^m$ (see section 1.1.1 for multi-index notations). The input state with n single photons in the first n modes and vacuum in the other modes is denoted $|\boldsymbol{t}\rangle$, with $\boldsymbol{t} = (\mathbf{1}^n, \mathbf{0}^{m-n})$. We introduce,

$$\Phi_{m,n} := \{\boldsymbol{s} \in \mathbb{N}^m,\ |\boldsymbol{s}| = n\}. \tag{1.73}$$

This set corresponds to the m-mode Fock states with total number of photons equal to n. We have $|\Phi_{m,n}| = \binom{m+n-1}{n}$ and $\boldsymbol{t} \in \Phi_{m,n}$.

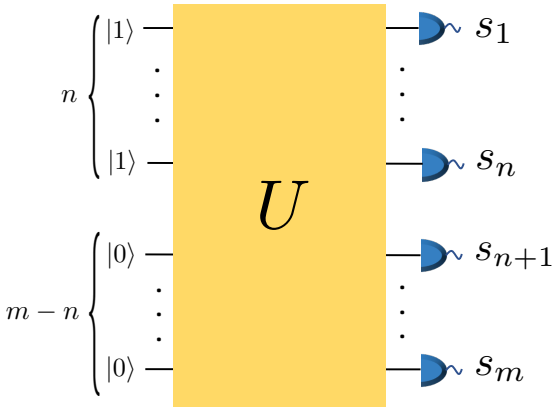

Figure 1.6: BosonSampling with n photons over m modes. The outcomes $s_1,\dots,s_m$ denote the measured photon number for each mode.

We consider a unitary interferometer of size m, described by an $m \times m$ unitary matrix $U = (u_{ij})_{1\leq i,j\leq m}$ acting on the creation and annihilation operators of the modes as in Eq. (1.68). We write $\hat{U}$ the unitary action of the interferometer on the multimode Fock basis. Its entries are indexed by elements of $\Phi_{m,n}$, for all $n \in \mathbb{N}$. Because the interferometer conserves the total number of photons, for all $p,q \in \mathbb{N}$, all $\boldsymbol{s} \in \Phi_{m,p}$ and all $\boldsymbol{s}' \in \Phi_{m,q}$,

$$\langle \boldsymbol{s}|\hat{U}|\boldsymbol{s}'\rangle = 0 \tag{1.74}$$

whenever $p \neq q$. In particular, it may be written as the direct sum of its action on the various fixed energy subspaces. We write

$$\hat{U} = \bigoplus_{n=1}^{+\infty} \hat{U}_n, \tag{1.75}$$

where $\hat{U}_n$ is the $|\Phi_{m,n}| \times |\Phi_{m,n}|$ unitary submatrix of $\hat{U}$ obtained by only keeping the rows $\boldsymbol{s}$ and the columns $\boldsymbol{s}'$ for all $\boldsymbol{s},\boldsymbol{s}' \in \Phi_{m,n}$. We have $\hat{U}_0 = \begin{pmatrix}1\end{pmatrix}$, and $\hat{U}_1 = U$ up to a reordering of the basis states.

Let $n \in \mathbb{N}$ and $\boldsymbol{s},\boldsymbol{t} \in \Phi_{m,n}$. Combining Eq. (1.68) and Eq. (1.72) we obtain [AA13]

$$\langle \boldsymbol{s}|\hat{U}|\boldsymbol{t}\rangle = \frac{\mathrm{Per}(U_{\boldsymbol{s},\boldsymbol{t}})}{\sqrt{\boldsymbol{s}!\boldsymbol{t}!}}, \tag{1.76}$$

where $U_{\boldsymbol{s},\boldsymbol{t}}$ is the $n \times n$ matrix obtained from U by repeating s_i times its i^{th} row and t_j times its j^{th} column for $i,j = 1,\dots,m$, and where the permanent of an $n \times n$ matrix $A = (a_{ij})_{1\leq i,j\leq n}$ is defined as

$$\mathrm{Per}\, A = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i\sigma(i)}, \tag{1.77}$$

where the sum is over the permutations of the set $\{1,\dots,n\}$.

We write $\Pr_{m,n}[.|\boldsymbol{t}]$ the probability distribution of the outputs over $\Phi_{m,n}$ of the unitary interferometer U acting on the input $|\boldsymbol{t}\rangle$. With the previous notations we obtain, for all $\boldsymbol{s},\boldsymbol{t} \in \Phi_{m,n}$,

$$\Pr_{m,n}[\boldsymbol{s}|\boldsymbol{t}] = \frac{\left|\mathrm{Per}\left(U_{\boldsymbol{s},\boldsymbol{t}}\right)\right|^2}{\boldsymbol{s}!\boldsymbol{t}!}. \tag{1.78}$$

With $\boldsymbol{t} = (\mathbf{1}^n, \mathbf{0}^{m-n})$ we have $\boldsymbol{t}! = 1$ and thus

$$\Pr_{m,n}[\boldsymbol{s}|\boldsymbol{t}] = \frac{1}{\boldsymbol{s}!}\left|\mathrm{Per}\left(U_{\boldsymbol{s},\boldsymbol{t}}\right)\right|^2. \tag{1.79}$$

The output photon-number distribution of a Boson Sampling interferometer with n photons over m modes thus is related to the modulus squared of the permanent of an $n \times n$ matrix with complex entries. This matrix is obtained from the unitary matrix U describing the interferometer by discarding its last $m-n$ columns and repeating its lines according to the detection pattern $\boldsymbol{s}$.

The permanent defined in Eq. (1.77) is a 'hard' quantity to compute. In order to appreciate this hardness, let us take a brief and informal detour through the realm of complexity theory [Man01]. A formal introduction to the complexity classes presented here is given in [AA13].

A complexity class is a set of computational problems. These problems may be of different types: in particular, a decision problem is a problem with yes or no answers, a function problem is a problem with more general answers (e.g., natural, real or complex numbers), and a sampling problem consists in outputting samples from a target probability distribution, either exactly or approximately.

In the language of complexity theory, an efficient algorithm is an algorithm which takes a number of steps which is polynomial in the size of its input (its number of bits), and the generic model for a classical computer is a deterministic Turing machine.

Given a complexity class C, a problem p is said to be C-hard if any problem in C can be rephrased efficiently as an instance of the problem p. Roughly speaking, this means that the problem p is harder than any of the problems in C. If the problem p is also in C, it is referred to as C-complete.

The class of decision problems that can be solved efficiently by a classical computer is denoted P. The class of decision problems whose solution can be verified efficiently by a classical computer is denoted NP. A great open problem in complexity theory is whether these two complexity classes are equal or if $\mathsf{P} \neq \mathsf{NP}$, the latter being widely believed.

An oracle for a given computational problem is a black box which is able to produce a solution for any instance of this problem. An oracle for a complexity class is a black box which, given any problem in the complexity class, is able to produce a solution for any instance of this problem. The access to an oracle is denoted with an exponent. For example, a problem which can be solved efficiently when given access to an oracle for an NP-complete problem is in the class P^{NP}.

The polynomial hierarchy PH is a tower of complexity classes generalising P and NP. It can be defined inductively based on an oracle construction, where the level 0 is P, the level 1 contains NP, the level 2 contains $\mathsf{NP}^{\mathsf{NP}}$, and so on. Each level is contained in the next one and if two

consecutive levels are equal, then they are also equal to all of the above levels—we talk about a collapse of the polynomial hierarchy. The conjecture that the polynomial hierarchy does not collapse, i.e., that all levels within the hierarchy are distinct, is a stronger version of the $\mathsf{P} \neq \mathsf{NP}$ conjecture.

The class of decision problems that can be solved efficiently by a classical computer with access to a genuine random number source is denoted BPP. It lies at the second level of the polynomial hierarchy PH_2 [Lau83].

The class of function problems which consist in counting the number of solutions of an NP problem is denoted #P. Its equivalent complexity class of decision problems is denoted $\mathsf{P}^{\#\mathsf{P}}$ and by Toda's theorem [Tod91] we have $\mathsf{PH} \subset \mathsf{P}^{\#\mathsf{P}}$, i.e., counting the solutions of NP problems is harder than any problem in the whole polynomial hierarchy of complexity classes.

With these elements introduced, we are now in position to discuss the hardness of the permanent: computing exactly the permanent of matrices with $(0,1)$ entries is a #P-complete problem [Val79] and hence PH-hard. Moreover, approximating the permanent of real matrices up to multiplicative error, i.e., outputting an estimate $\tilde{P}$ such that $(1-1/\text{poly}\, m)P \leq \tilde{P} \leq (1+/\text{poly}\, m)P$ where P is the permanent of a square matrix of size m with real entries, is also #P-hard [AA13].

The computational problem 'Boson Sampling' corresponds to the task of sampling from the output probability distribution in Eq. (1.79), given the description U of the Boson Sampling interferometer.

Making use of the hardness of the permanent and the connection between the output probabilities of a Boson Sampling interferometer and the permanent, two main results are derived in [AA13] about the hardness of classically solving two versions of the Boson Sampling problem, which we refer to as exact hardness and approximate hardness.

Exact hardness corresponds to the following result: let $\mathcal{O}$ be an oracle which, given a unitary matrix U and a random string as its unique source of randomness, samples exactly from the output probability distribution of the Boson Sampling interferometer U. Then $\mathsf{PH} \subset \mathsf{BPP}^{\mathsf{NP}^{\mathcal{O}}}$. In particular, an efficient classical simulation of exact Boson Sampling collapses the polynomial hierarchy to its third level.

This result uses the fact that a single output probability of a Boson Sampling interferometer is hard to approximate up to multiplicative error and that being able to sample efficiently from a probability distribution allows one to obtain a multiplicative approximation of the probability of any outcome in $\mathsf{FBPP}^{\mathsf{NP}}$ (where FBPP is the class of function problems that can be solved efficiently using a BPP machine) thanks to Stockmeyer's approximate counting algorithm [Sto85]. In that case, an oracle which samples from an exact Boson Sampling probability distribution is required.

On the other hand, approximate sampling refers to the task of sampling from a probability distribution which has a given constant total variation distance with a target distribution (see Eq. (1.14)). Approximate hardness of Boson Sampling is more elaborate than exact hardness

and relies on two plausible but unproven conjectures, even though the statement of the result is nearly identical: let $\mathcal{O}$ be an oracle which, given a unitary matrix U and a random string as its unique source of randomness, samples approximately from the output probability distribution of the Boson Sampling interferometer U. Then $\mathsf{PH} \subset \mathsf{BPP}^{\mathsf{NP}^{\mathcal{O}}}$. In particular, an efficient classical simulation of approximate Boson Sampling collapses the polynomial hierarchy to its third level.

Unlike for exact sampling, one cannot apply directly Stockmeyer's approximate counting algorithm in order to obtain multiplicative estimates of the probabilities of the target distribution. This is because the oracle now only outputs samples from an approximate probability distribution, i.e., a probability distribution which is very close to the correct one for most of the samples but not all samples. In the worst case, the probability that we are trying to estimate could be the probability of one of these 'bad samples', and estimating this probability would merely give us a very bad estimate of the permanent, which is not hard to achieve. The trick to get around that problem is to hide the instance of the permanent that we wish to estimate into the probability of a random output of a Boson Sampling interferometer: given a classical machine which correctly performs the sampling for most of the samples, it would then correctly sample our instance with high probability. In the worst case, this effectively averages the constant total variation error over the sample space, allowing for a much more precise approximation of the permanent using Stockmeyer's algorithm.

This hiding procedure is based on the fact that small enough submatrices of random unitary matrices are very close to random complex Gaussian matrices. In order to restrict to matrices that do not have repeated lines, the so-called antibunching regime $n = O(\sqrt{m})$ is chosen, which ensures a negligible probability of detecting more than one photon in the same output mode. The procedure outlined above then allows one to prove that the problem $|\mathrm{GPE}|^2_{\pm}$ which consists in approximating up to additive error the square modulus of the permanent of random complex Gaussian matrices is in $\mathsf{FBPP}^{\mathsf{NP}^{\mathcal{O}}}$, where $\mathcal{O}$ is an oracle for approximate Boson Sampling.

The proof of approximate hardness then relies on two conjectures about the permanent of random complex Gaussian matrices in order to bridge the gap between additive approximations of the square modulus of the permanent of random complex Gaussian matrices and collapse of the polynomial hierarchy: the permanent of Gaussians conjecture and the permanent anti-concentration conjecture. The former conjecture states that the problem $\mathrm{GPE}_{\times}$ which consists in approximating the permanent of random complex Gaussian matrices up to multiplicative error is #P-hard. The latter conjecture states that with high probability the permanent of a randomly chosen complex Gaussian matrix is not too small. This implies in turn that the problem $|\mathrm{GPE}|^2_{\pm}$ of additive approximation of the square modulus of the permanent of random complex Gaussian matrices is as hard as the problem $\mathrm{GPE}_{\times}$ of multiplicative approximation of the permanent of random complex Gaussian matrices. With these two conjectures and the above argument, we obtain $\mathsf{PH} \subset \mathsf{P}^{\#\mathsf{P}} \subset \mathrm{GPE}_{\times} = |\mathrm{GPE}|^2_{\pm} \subset \mathsf{FBPP}^{\mathsf{NP}^{\mathcal{O}}}$, which concludes the proof.

Assuming that the polynomial hierarchy does not collapse, Boson Sampling is hard to simu-

late exactly classically, and even approximately with additional mathematical conjectures. The approximate hardness of Boson Sampling is important since it opens the way for an experimental demonstration of quantum supremacy. Indeed, it is unrealistic to expect that an experimental Boson Sampling device would sample exactly from the ideal Boson Sampling distribution. Moreover, given the nature of the computational task at hand, i.e., outputting samples from a given probability distribution, there is no hope of being able to verify that an exact sampling has been performed. On the other hand, verifying that approximate Boson Sampling has been performed could be possible and indeed we derive such a verification protocol in chapter 4.

1.5 Segal–Bargmann formalism

The Segal–Bargmann formalism [Bar61, SM63] associates to every quantum state an analytical function over the complex plane. It has been used to study quantum chaos [LV90, ABB96, KMW97, BS99], and the completeness of sequences of coherent states [Per71, BGZ75, BZ78]. We give hereafter a quick introduction to this formalism. Further details may be found in chapter 2 and in [Vou06].

1.5.1 Definition

We introduce below the analytical function, which we refer to as the stellar function. This function has been recently studied, in the context of non-Gaussian quantum state engineering [GG19], in order to simplify calculations related to photon-subtracted Gaussian states.

Definition 1.1 (Stellar function)**.** Let $|\psi\rangle = \sum_{n\geq 0} \psi_n |n\rangle \in \mathscr{H}$ be a normalised state. The stellar function of the state $|\psi\rangle$ is defined as

$$F_\psi^\star(z) = e^{\frac{1}{2}|z|^2} \langle z^* | \psi \rangle = \sum_{n\geq 0} \psi_n \frac{z^n}{\sqrt{n!}}, \tag{1.80}$$

for all $z \in \mathbb{C}$, where $|z\rangle = e^{-\frac{1}{2}|z|^2} \sum_{n\geq 0} \frac{z^n}{\sqrt{n!}} |n\rangle$ is the coherent state of amplitude z.

The stellar function is a holomorphic function over the complex plane, which provides an analytic representation of a quantum state.

1.5.2 Properties of holomorphic functions

A holomorphic function is a complex-valued function of one or more complex variables that is, at every point of its domain, complex differentiable in a neighbourhood of the point. As it turns out, the set of holomorphic functions is equal to the set of analytic functions, i.e., the functions that can be written as a convergent power series in a neighbourhood of each point of their domain. When their domain is the whole complex plane, they are called entire functions. In what follows we consider univariate entire functions.

These functions provide a natural extension of univariate complex polynomials and various properties of polynomials extend to entire functions. In particular, Liouville's theorem states that any bounded entire function is constant. The principle of permanence asserts that the zeros of an analytic function are isolated or this function is identically 0. Furthermore, the number of zeros of an analytic function f inside some contour is given by Cauchy's argument principle.

Theorem 1.1 (Cauchy's argument principle)**.** *Let f be an analytic function and let C be a contour in the domain of f. Then,*

$$Z_C(f) = \frac{1}{2i\pi} \oint_C \frac{f'(z)}{f(z)} dz, \tag{1.81}$$

where $Z_C(f)$ is the number of zeros of f inside the contour C, counted with multiplicity.

The growth of an analytic function is described by a pair of non-negative numbers ρ, σ called the order and the type. They are defined as [Boa54]

$$\rho = \lim_{r \to +\infty} \sup \frac{\ln \ln M(r)}{\ln r}, \quad \sigma = \lim_{r \to +\infty} \sup \frac{\ln M(r)}{r^\rho}, \tag{1.82}$$

where $M(r)$ is the maximum value of the modulus of the funtion on the circle $|z| = r$. For polynomials, the growth is deeply related to the number of zeros—the degree. For entire functions, the growth is related to the density of zeros (see, e.g., [SS10] for more details). An entire function can also be factorized into a possibly infinite product involving its zeros, thanks to Weierstrass factorization theorem. For entire functions of finite order, this result is refined by Hadamard-Weierstrass factorization theorem.

Theorem 1.2 (Hadamard-Weierstrass factorization theorem)**.** *Let f be an entire function of finite order ρ. Let $m \in \mathbb{N}$ be the multiplicity of 0 as a root of f. Let $\{z_n\} n \in \mathbb{N}$ be the non-zero roots of f, counted with multiplicity. Then, there exist $p, q \in \mathbb{N}$, with $p, q \leq \rho$, and a polynomial P of degree q such that, for all $z \in \mathbb{C}$,*

$$f(z) = z^m e^{P(z)} \prod_{n=1}^{+\infty} E_p\left(\frac{z}{z_n}\right), \tag{1.83}$$

where

$$E_p(z) = (1-z)\, e^{z + z^2/2 + \cdots + z^p/p}. \tag{1.84}$$

CHAPTER 2

STELLAR REPRESENTATION OF NON-GAUSSIAN QUANTUM STATES

Non-Gaussian states are crucial for a variety of quantum information tasks [ESP02, Fiu02, GC02, WHG+03, GPFC+04, GS07, NFC09, ADDS+09, BDE+19]. In particular, non-Gaussian states may be conceived as a resource for quantum computational advantage, Gaussian processes being classically simulable [BSBN02]. Hence, the characterisation of non-Gaussian states is of great importance and has attracted a lot of attention recently [TZ18, ZSS18, AGPF18, LRW+18].

In this chapter, building on the Segal–Bargmann formalism, we introduce the stellar representation, which allows for the representation of the non-Gaussian properties of single-mode continuous variable quantum states by the distribution of the zeros of their Husimi Q function in phase space. We use of this representation in order to derive an infinite hierarchy of single-mode states based on the number of zeros of the Husimi Q function, the stellar hierarchy. We give an operational characterisation of the states in this hierarchy with the minimal number of single-photon additions needed to engineer them and derive equivalence classes under Gaussian unitary operations. We study in detail the topological properties of this hierarchy with respect to the trace norm, and discuss implications for the robustness of the states in the stellar hierarchy and for non-Gaussian state engineering.

This chapter is based on [CMG20b, CRW+20].

2.1 The stellar function

In continuous variable quantum information, quantum states are mathematically described by vectors in a separable Hilbert space of infinite dimension (see section 1.1.3). Alternatively, phase space formalism allows us to describe quantum states conveniently using generalised quasi-probability distributions [CG69a], among which are the Husimi Q function, the Wigner

W function, and the Glauber–Sudarshan P function (see section 1.2). The states that have a Gaussian Wigner or Husimi function are called Gaussian states, while all the other states are called non-Gaussian. By extension, the operations mapping Gaussian states to Gaussian states are called Gaussian operations, and measurements projecting onto Gaussian states are called Gaussian measurements (see section 1.3).

Hudson [Hud74] has notably shown that a single-mode pure quantum state is non-Gaussian if and only if its Wigner function has negative values and this result has been generalised to multimode states by Soto and Claverie [SC83]. This characterization is an interesting starting point for studying non-Gaussian states. From this result, one can introduce measures of a state being non-Gaussian using Wigner negativity, e.g., the negative volume [KŻ04], that are invariant under Gaussian operations. However, computing these quantities from experimental data is complicated in practice. Other measures and witnesses for non-Gaussian states have been derived [GPB07, FMJ11, GPT$^+$13, HGT$^+$14], which allow us to discriminate non-Gaussian states from mixtures of Gaussian states from experimental data, but they do not address the structure of non-Gaussian states and answer the question *how much?* rather than *how?*.

In order to adress the latter question, we will make use of another characterization of Gaussian states: the Wigner function having negative values is actually equivalent to the Husimi function having zeros, as shown by Lütkenhaus and Barnett [LB95]. Informally,

Theorem 2.1. *A pure quantum state is non-Gaussian if and only if its Husimi Q function has zeros.*

Since the values of the Q function are the overlaps with coherent states, this result may be understood as follows: a pure quantum state is non-Gaussian if and only if it is orthogonal to at least one coherent state.

An interesting point is that for single-mode states, the zeros of the Husimi Q function form a discrete set, as we will show in the next section. The non-Gaussian properties of single-mode states may thus be described by the distribution of these zeros in phase space. Based on this observation, we classify single-mode continuous variable quantum states with respect to their non-Gaussian properties in the following sections, using the so-called stellar representation, or Segal–Bargmann formalism (see section 1.5), its link with the Husimi Q function and properties of holomorphic functions.

2.1.1 Definition and uniqueness

In what follows, $\mathscr{H}$ denotes a single-mode infinite-dimensional Hilbert space. We recall the definition of the stellar function [Bar61, SM63] and prove a few important properties.

Definition 2.1 (Stellar function)**.** Let $|\psi\rangle = \sum_{n\geq 0} \psi_n |n\rangle \in \mathscr{H}$ be a normalised state. The *stellar*

function of the state $|\psi\rangle$ is defined as

$$F^\star_\psi(z) = e^{\frac{1}{2}|z|^2} \langle z^* | \psi \rangle = \sum_{n\geq 0} \frac{\psi_n}{\sqrt{n!}} z^n, \tag{2.1}$$

for all $z \in \mathbb{C}$, where $|z\rangle = e^{-\frac{1}{2}|z|^2} \sum_{n\geq 0} \frac{z^n}{\sqrt{n!}} |n\rangle \in \mathscr{H}$ is the coherent state of amplitude z.

We now develop the formalism further, analysing the zeros of the stellar function to characterise states. The stellar function is a holomorphic function over the complex plane. For any normalised state $|\psi\rangle \in \mathscr{H}$ and all $z \in \mathbb{C}$,

$$\begin{aligned} \left|F^\star_\psi(z)\right|^2 &\leq \left| \sum_{n\geq 0} \psi_n \frac{z^n}{\sqrt{n!}} \right|^2 \\ &\leq \sum_{n\geq 0} |\psi_n|^2 \sum_{n\geq 0} \frac{|z|^{2n}}{n!} \\ &= e^{|z|^2} \end{aligned} \tag{2.2}$$

by Cauchy-Schwarz inequality. This implies that the stellar function of a normalised state is of finite order less or equal to 2 and type less or equal to $\frac{1}{2}$.

From the definition of the stellar function, for any state $|\psi\rangle \in \mathscr{H}$ we may write

$$|\psi\rangle = \sum_{n\geq 0} \psi_n |n\rangle = F^\star_\psi(\hat{a}^\dagger)|0\rangle\,. \tag{2.3}$$

From this equation one may understand the stellar function as an operational recipe for engineering a state from the vacuum, using the creation operator $\hat{a}^\dagger$. This intuition will be made more precise in the following sections. An important property is that the stellar representation is unique:

Lemma 2.1. *Let $|\phi\rangle$ and $|\psi\rangle$ be pure normalised single-mode states such that $F^\star_\phi = F^\star_\psi$. Then $|\phi\rangle = |\psi\rangle$. Moreover, let $|\chi\rangle = f(\hat{a}^\dagger)|0\rangle$ be a single-mode normalised pure state, where f is analytic. Then $f = F^\star_\chi$.*

Proof. With the notations of the Lemma, $F^\star_\phi(z) = \sum_{n\geq 0} \phi_n \frac{z^n}{\sqrt{n!}}$ and $F^\star_\psi(z) = \sum_{n\geq 0} \psi_n \frac{z^n}{\sqrt{n!}}$. The functions $F^\star_\phi$ and $F^\star_\psi$ are analytic, so $F^\star_\phi(z) = F^\star_\psi(z)$ implies that $\phi_n = \psi_n$ for all $n \geq 0$. Hence $|\phi\rangle = |\psi\rangle$.

Now with $|\chi\rangle = \sum_{n\geq 0} \chi_n |n\rangle = f(\hat{a}^\dagger)|0\rangle$, let us write $f(z) = \sum_{n\geq 0} f_n z^n$. We obtain

$$\begin{aligned} |\chi\rangle &= \sum_{n\geq 0} f_n (\hat{a}^\dagger)^n |0\rangle \\ &= \sum_{n\geq 0} f_n \sqrt{n!}\, |n\rangle, \end{aligned} \tag{2.4}$$

so $\chi_n = f_n\sqrt{n!}$ for all $n \geq 0$. On the other hand, for all $z \in \mathbb{C}$,

$$\begin{aligned} F_\chi^\star(z) &= e^{\frac{1}{2}|z|^2} \langle z^* | \psi \rangle \\ &= \sum_{n\geq 0} \chi_n \frac{z^n}{\sqrt{n!}} \\ &= \sum_{n\geq 0} f_n z^n \\ &= f(z). \end{aligned} \tag{2.5}$$

■

The stellar function of a state $|\psi\rangle \in \mathscr{H}$ is related to its Husimi Q function, a smoothed version of the Wigner function [CG69a], given by

$$Q_\psi(z) = \frac{1}{\pi} |\langle z|\psi\rangle|^2 = \frac{e^{-|z|^2}}{\pi} \left|F_\psi^\star(z^*)\right|^2, \tag{2.6}$$

for all $z \in \mathbb{C}$. The zeros of the Husimi Q function are the complex conjugates of the zeros of $F_\psi^\star$. Hence, by Theorem 2.1, a single-mode pure quantum state is non-Gaussian if and only if its stellar function has zeros. These zeros form a discrete set, as the stellar function is a non-zero analytic function. The non-Gaussian properties of a single-mode pure state are then described by the distribution of the zeros over the complex plane.

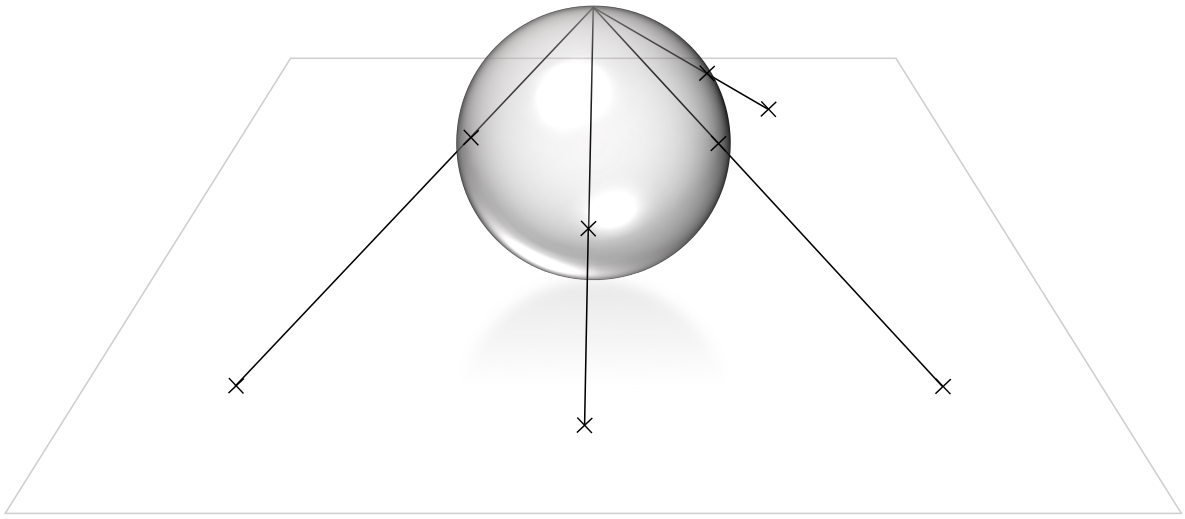

Figure 2.1: Antistereographic projection of four points onto the sphere

Using anti-stereographic projection [SB07], this amounts to describing the non-Gaussian properties of a pure state with a set of points on the sphere (Fig. 2.1), hence the name stellar

representation, where the points on the sphere looked at from the center of the sphere are seen as stars on the celestial vault [TV95, KMW97].

In all the chapter we will use for brevity the notations $c_\chi = \cosh\chi$, $s_\chi = \sinh\chi$ and $t_\chi = \tanh\chi$, for all $\chi \in \mathbb{C}$.

2.1.2 Examples

In this section we give the stellar functions of various states and operators.

2.1.2.1 Gaussian states and Fock states

The displacement operator of amplitude $\alpha \in \mathbb{C}$ is given by $\hat{D}(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}$. Its action on the vacuum state yields the coherent state $|\alpha\rangle$. The squeeze operator of parameter $\xi = re^{i\theta} \in \mathbb{C}$ is given by $\hat{S}(\xi) = e^{\frac{1}{2}(\xi\hat{a}^2 - \xi^*\hat{a}^{\dagger 2})}$. Its action on the vacuum state yields the squeezed state $|\xi\rangle$. All single-mode Gaussian operations may be decomposed as a squeezing operation and a displacement (see section 1.3).

For any single-mode Gaussian state $\hat{S}(\xi)\hat{D}(\alpha)|0\rangle$, where $\xi = re^{i\theta}$, the corresponding stellar function is [Vou06]:

$$G^\star_{\xi,\alpha}(z) = (1-|a|^2)^{1/4} e^{-\frac{1}{2}az^2 + bz + c}, \tag{2.7}$$

where

$$a := e^{-i\theta}\tanh r,\ b := \alpha\sqrt{1-|a|^2} = \frac{\alpha}{\cosh r},\ c := \frac{1}{2}a^*\alpha^2 - \frac{1}{2}|\alpha|^2. \tag{2.8}$$

In particular, we obtain

$$G^\star_{0,\alpha}(z) = e^{\alpha z - \frac{1}{2}|\alpha|^2}, \tag{2.9}$$

for a coherent state of amplitude $\alpha \in \mathbb{C}$, and

$$G^\star_{\xi,0}(z) = \frac{1}{\sqrt{\cosh r}} e^{-\frac{1}{2}(e^{-i\theta}\tanh r)z^2}, \tag{2.10}$$

for a squeezed vacuum state with squeezing parameter $\xi = re^{i\theta} \in \mathbb{C}$.

For Fock states $|n\rangle$ with $n \in \mathbb{N}$, the stellar function is simply given by

$$F^\star_n(z) = \frac{z^n}{\sqrt{n!}}. \tag{2.11}$$

2.1.2.2 Cat states

Let us define for $\alpha \in \mathbb{C}$ the cat$^+$ and cat$^-$ states:

$$|\mathrm{cat}^\pm_\alpha\rangle = \frac{1}{\sqrt{\mathcal{N}^\pm_\alpha}}(|\alpha\rangle \pm |-\alpha\rangle), \tag{2.12}$$

where $|\alpha\rangle$ is a coherent state, and $\mathcal{N}^\pm_\alpha$ is a normalisation factor.

Lemma 2.2. *The stellar functions of cat states are given by*

$$F^{\star}_{cat^{+}_{\alpha}}(z) = \frac{\cosh(\alpha z)}{\sqrt{\cosh(|\alpha|^2)}}, \tag{2.13}$$

and

$$F^{\star}_{cat^{-}_{\alpha}}(z) = \frac{\sinh(\alpha z)}{\sqrt{\sinh(|\alpha|^2)}}, \tag{2.14}$$

for all $z, \alpha \in \mathbb{C}$.

Proof. The overlap between two coherent states is given by

$$\langle z|\alpha\rangle = e^{-\frac{1}{2}(|z|^2+|\alpha|^2-2z^*\alpha)}, \tag{2.15}$$

for $z, \alpha \in \mathbb{C}$. Hence, with $\langle \mathrm{cat}^{\pm}_{\alpha}|\mathrm{cat}^{\pm}_{\alpha}\rangle = 1$ we have

$$\mathcal{N}^{\pm}_{\alpha} = 2(1 \pm e^{-2|\alpha|^2}). \tag{2.16}$$

We then obtain for $z, \alpha \in \mathbb{C}$,

$$\begin{aligned} F^{\star}_{\mathrm{cat}^{\pm}_{\alpha}}(z) &= e^{-\frac{1}{2}|z|^2}\,\langle z^*|\mathrm{cat}^{\pm}_{\alpha}\rangle \\ &= \frac{1}{\sqrt{\mathcal{N}^{\pm}_{\alpha}}}\left(e^{-\frac{1}{2}|\alpha|^2+\alpha z} \pm e^{-\frac{1}{2}|\alpha|^2-\alpha z}\right) \\ &= \frac{1}{\sqrt{2\left(e^{|\alpha|^2} \pm e^{-|\alpha|^2}\right)}}\left(e^{\alpha z} \pm e^{-\alpha z}\right). \end{aligned} \tag{2.17}$$

We finally obtain

$$F^{\star}_{\mathrm{cat}^{+}_{\alpha}}(z) = \frac{\cosh(\alpha z)}{\sqrt{\cosh(|\alpha|^2)}}, \tag{2.18}$$

and

$$F^{\star}_{\mathrm{cat}^{-}_{\alpha}}(z) = \frac{\sinh(\alpha z)}{\sqrt{\sinh(|\alpha|^2)}}. \tag{2.19}$$

■

2.1.2.3 GKP states

The set of Gottesman-Kitaev-Preskill (GKP) states have been proposed as a means for encoding a qubit in an oscillator, in a way which is fault-tolerant to small shifts in position and momentum [GKP01]. An example of such states is the simultaneous +1 eigenstate of the two commuting displacements operators $e^{-i\sqrt{2\pi}\hat{p}}$ and $e^{i\sqrt{2\pi}\hat{q}}$. The corresponding encoding may correct for comparable shifts in $\hat{q}$ and $\hat{p}$. An expression for this unphysical state (it has infinite norm) is given

by

$$\begin{aligned} |\mathrm{GKP}\rangle &= \sum_{s\in\mathbb{Z}} e^{-i\sqrt{2\pi}s\hat{p}} \sum_{t\in\mathbb{Z}} e^{i\sqrt{2\pi}t\hat{q}} |0\rangle \\ &= \sum_{s,t\in\mathbb{Z}^2} (-1)^{st} \hat{D}\left(2\sqrt{\pi}(s+it)\right)|0\rangle \\ &= \sum_{s,t\in\mathbb{Z}^2} (-1)^{st} |2\sqrt{\pi}(s+it)\rangle, \end{aligned} \tag{2.20}$$

as an infinite superposition of coherent states. The stellar function of this state is then given by

$$\begin{aligned} F^\star_{\mathrm{GKP}}(z) &= \sum_{s,t\in\mathbb{Z}^2} (-1)^{st} F^\star_{2\sqrt{\pi}(s+it)}(z) \\ &= \sum_{s,t\in\mathbb{Z}^2} (-1)^{st} e^{-2\pi(s^2+t^2)} e^{2\sqrt{\pi}(s+it)z}, \end{aligned} \tag{2.21}$$

where we used Eq. (2.9) in the second line. This stellar function may be expressed as a Riemann theta function [Rie57]. Using properties of these functions, we obtain the following result:

Lemma 2.3. *$F^\star_{GKP}$ has an infinite number of zeros and has exactly* 16 *zeros counted with multiplicity in each square region of the complex plane of size $4\sqrt{\pi}\times 4\sqrt{\pi}$.*

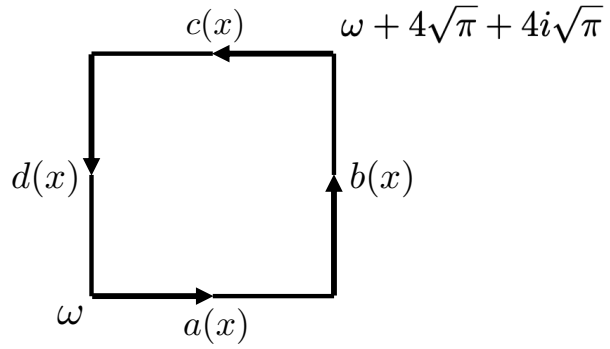

Figure 2.2: The contour used for the argument principle

Proof. We first derive a few invariance properties of $F^\star_{\mathrm{GKP}}$ and conclude with the argument principle (Theorem 1.1). For all $z\in\mathbb{C}$, we have

$$\begin{aligned} F^\star_{\mathrm{GKP}}(iz) &= \sum_{s,t\in\mathbb{Z}^2} (-1)^{st} e^{-2\pi(s^2+t^2)} e^{2\sqrt{\pi}(-t+is)z} \\ &\underset{t\to -t}{=} \sum_{s,t\in\mathbb{Z}^2} (-1)^{st} e^{-2\pi(s^2+t^2)} e^{2\sqrt{\pi}(t+is)z} \\ &\underset{s\leftrightarrow t}{=} F^\star_{\mathrm{GKP}}(z). \end{aligned} \tag{2.22}$$

Now for all $z \in \mathbb{C}$,

$$\begin{aligned} F^\star_{\mathrm{GKP}}(z) &= \sum_{s,t\in\mathbb{Z}^2} (-1)^{st} e^{-2\pi(s^2+t^2)} e^{2\sqrt{\pi}(s+it)z} \\ &\underset{s\to s-2}{=} \sum_{s,t\in\mathbb{Z}^2} (-1)^{(s+2)t} e^{-2\pi((s+2)^2+t^2)} e^{2\sqrt{\pi}((s+2)+it)z} \\ &= \sum_{s,t\in\mathbb{Z}^2} (-1)^{st} e^{-2\pi(s^2+t^2)} e^{-8\pi s-8\pi} e^{2\sqrt{\pi}(s+it)z} e^{4\sqrt{\pi}z} \\ &= e^{4\sqrt{\pi}z-8\pi} \sum_{s,t\in\mathbb{Z}^2} (-1)^{st} e^{-2\pi(s^2+t^2)} e^{2\sqrt{\pi}(s+it)(z-4\sqrt{\pi})} \\ &= e^{4\sqrt{\pi}z-8\pi} F^\star_{\mathrm{GKP}}(z-4\sqrt{\pi}) \end{aligned} \tag{2.23}$$

Combining this with Eq. (2.22) we also obtain for all $z \in \mathbb{C}$,

$$\begin{aligned} F^\star_{\mathrm{GKP}}(z) &= F^\star_{\mathrm{GKP}}(iz) \\ &= e^{4\sqrt{\pi}iz-8\pi} F^\star_{\mathrm{GKP}}(iz-4\sqrt{\pi}) \\ &= e^{4\sqrt{\pi}iz-8\pi} F^\star_{\mathrm{GKP}}(z+4i\sqrt{\pi}). \end{aligned} \tag{2.24}$$

This means that $F^\star_{\mathrm{GKP}}$ is quasiperiodic along the horizontal and vertical directions in the complex plane, with period $4\sqrt{\pi}$. The functions $z \mapsto e^{4\sqrt{\pi}z-8\pi}$ and $z \mapsto e^{4\sqrt{\pi}iz-8\pi}$ do not vanish and it is thus sufficient for our purpose to prove that $F^\star_{\mathrm{GKP}}$ has at least one zero: the quasiperiodicity implies that $F^\star_{\mathrm{GKP}}$ would also vanish on the lattice with square cells of size $4\sqrt{\pi}$ containing this zero.

By Theorem 1.1, the number of zeros of $F^\star_{\mathrm{GKP}}$ inside a closed contour C counted with multiplicity is given by

$$Z_C\left(F^\star_{\mathrm{GKP}}\right) = \frac{1}{2i\pi} \oint_C \frac{\partial_z F^\star_{\mathrm{GKP}}(z)}{F^\star_{\mathrm{GKP}}(z)} dz. \tag{2.25}$$

For all $\omega \in \mathbb{C}$, we consider the square contour C_ω with corners ω, $\omega+4\sqrt{\pi}$, $\omega+4\sqrt{\pi}+4i\sqrt{\pi}$ and $\omega+4i\sqrt{\pi}$, parametrised by

$$\begin{aligned} a(x) &= \omega + 4\sqrt{\pi}x \\ b(x) &= \omega + 4\sqrt{\pi} + 4i\sqrt{\pi}x \\ c(x) &= \omega + 4\sqrt{\pi}(1-x) + 4i\sqrt{\pi} \\ d(x) &= \omega + 4i\sqrt{\pi}(1-x), \end{aligned} \tag{2.26}$$

for $x \in [0,1]$ (Fig. 2.2). We have $a'(x) = 4\sqrt{\pi}$, $b'(x) = 4i\sqrt{\pi}$, $c'(x) = -4\sqrt{\pi}$ and $d'(x) = -4i\sqrt{\pi}$ for all $x \in [0,1]$.

The quasiperiodicity of $F^\star_{\mathrm{GKP}}$ may be rewritten as

$$\begin{cases} F^\star_{\mathrm{GKP}}(z+4\sqrt{\pi}) = e^{8\pi+4\sqrt{\pi}z} F^\star_{\mathrm{GKP}}(z), \\ F^\star_{\mathrm{GKP}}(z+4i\sqrt{\pi}) = e^{8\pi-4i\sqrt{\pi}z} F^\star_{\mathrm{GKP}}(z), \end{cases} \tag{2.27}$$

for all $z \in \mathbb{C}$. Taking the derivative with respect to z we obtain

$$\begin{cases} \partial_z F^\star_{\mathrm{GKP}}(z+4\sqrt{\pi}) = e^{8\pi+4\sqrt{\pi}z} \left[\partial_z F^\star_{\mathrm{GKP}}(z) + 4\sqrt{\pi} F^\star_{\mathrm{GKP}}(z)\right], \\ \partial_z F^\star_{\mathrm{GKP}}(z+4i\sqrt{\pi}) = e^{8\pi-4i\sqrt{\pi}z} \left[\partial_z F^\star_{\mathrm{GKP}}(z) - 4i\sqrt{\pi} F^\star_{\mathrm{GKP}}(z)\right], \end{cases} \tag{2.28}$$

and thus

$$\begin{cases} \frac{\partial_z F^\star_{\mathrm{GKP}}(z+4\sqrt{\pi})}{F^\star_{\mathrm{GKP}}(z+4\sqrt{\pi})} = 4\sqrt{\pi} + \frac{\partial_z F^\star_{\mathrm{GKP}}(z)}{F^\star_{\mathrm{GKP}}(z)}, \\ \frac{\partial_z F^\star_{\mathrm{GKP}}(z+4i\sqrt{\pi})}{F^\star_{\mathrm{GKP}}(z+4i\sqrt{\pi})} = -4i\sqrt{\pi} + \frac{\partial_z F^\star_{\mathrm{GKP}}(z)}{F^\star_{\mathrm{GKP}}(z)}. \end{cases} \tag{2.29}$$

With Eq. (2.25), for all $\omega \in \mathbb{C}$,

$$\begin{aligned} Z_{C_\omega}(F^\star_{\mathrm{GKP}}) &= \frac{1}{2i\pi} \oint_{C_\omega} \frac{\partial_z F^\star_{\mathrm{GKP}}(z)}{F^\star_{\mathrm{GKP}}(z)} dz \\ &= \frac{1}{2i\pi} \int_0^1 \left[\frac{\partial_z F^\star_{\mathrm{GKP}}(a(x))}{F^\star_{\mathrm{GKP}}(a(x))} a'(x) + \frac{\partial_z F^\star_{\mathrm{GKP}}(b(x))}{F^\star_{\mathrm{GKP}}(b(x))} b'(x) \right. \\ &\qquad \left. + \frac{\partial_z F^\star_{\mathrm{GKP}}(c(x))}{F^\star_{\mathrm{GKP}}(c(x))} c'(x) + \frac{\partial_z F^\star_{\mathrm{GKP}}(d(x))}{F^\star_{\mathrm{GKP}}(d(x))} d'(x) \right] dx. \end{aligned} \tag{2.30}$$

Given that $c(x) = a(1-x) + 4i\sqrt{\pi}$, we have

$$\begin{aligned} \int_0^1 \frac{\partial_z F^\star_{\mathrm{GKP}}(c(x))}{F^\star_{\mathrm{GKP}}(c(x))} c'(x) dx &= \int_0^1 -4\sqrt{\pi} \frac{\partial_z F^\star_{\mathrm{GKP}}(a(1-x)+4i\sqrt{\pi})}{F^\star_{\mathrm{GKP}}(a(1-x)+4i\sqrt{\pi})} dx \\ &= \int_0^1 -4\sqrt{\pi} \left(-4i\sqrt{\pi} + \frac{\partial_z F^\star_{\mathrm{GKP}}(a(1-x))}{F^\star_{\mathrm{GKP}}(a(1-x))} \right) \\ &= 16i\pi - \int_0^1 \frac{\partial_z F^\star_{\mathrm{GKP}}(a(1-x))}{F^\star_{\mathrm{GKP}}(a(1-x))} a'(1-x) dx \\ &= 16i\pi - \int_0^1 \frac{\partial_z F^\star_{\mathrm{GKP}}(a(x))}{F^\star_{\mathrm{GKP}}(a(x))} a'(x) dx, \end{aligned} \tag{2.31}$$

where we used Eq. (2.29) in the second line with $z = a(1-x)$. Hence,

$$\frac{1}{2i\pi} \int_0^1 \left[\frac{\partial_z F^\star_{\mathrm{GKP}}(a(x))}{F^\star_{\mathrm{GKP}}(a(x))} a'(x) + \frac{\partial_z F^\star_{\mathrm{GKP}}(c(x))}{F^\star_{\mathrm{GKP}}(c(x))} c'(x) \right] dx = 8. \tag{2.32}$$

Similarly $b(x) = d(1-x) + 2\sqrt{2\pi}$ gives

$$\frac{1}{2i\pi} \int_0^1 \left[\frac{\partial_z F^\star_{\mathrm{GKP}}(b(x))}{F^\star_{\mathrm{GKP}}(b(x))} b'(x) + \frac{\partial_z F^\star_{\mathrm{GKP}}(d(x))}{F^\star_{\mathrm{GKP}}(d(x))} d'(x) \right] dx = 8. \tag{2.33}$$

With Eq. (2.30) we finally obtain

$$Z_{C_\omega}\left(F^\star_{\mathrm{GKP}}\right) = 16. \tag{2.34}$$

The function $F^\star_{\mathrm{GKP}}$ thus has an infinite number of zeros. Since Eq. (2.34) is independent of the choice of $\omega \in \mathbb{C}$, $F^\star_{\mathrm{GKP}}$ has exactly 16 zeros (counted with multiplicity) in each square region of the complex plane of size $4\sqrt{\pi} \times 4\sqrt{\pi}$. Moreover, with the property of invariance under rotation in Eq. (2.22), it has exactly 4 zeros in each square region of size $2\sqrt{\pi} \times 2\sqrt{\pi}$ whose corners have coordinates in $2\sqrt{\pi}\mathbb{Z}$ (by considering the square region of size $4\sqrt{\pi} \times 4\sqrt{\pi}$ centered on the origin).

■

2.1.2.4 Operators

While operators have their own treatment in the Segal–Bargmann formalism [Vou06], it is sufficient for our purpose to consider the following correspondences: the creation and annihilation operators have the stellar representations

$$\hat{a}^\dagger \to z, \quad \hat{a} \to \partial_z, \tag{2.35}$$

i.e., the operator corresponding to $\hat{a}^\dagger$ in the stellar representation is the multiplication by z and the operator in the stellar representation corresponding to $\hat{a}$ is the derivative with respect to z. This implies that the stellar function of a photon-added state $\hat{a}^\dagger|\psi\rangle$ is given by $z \mapsto zF^\star_\psi(z)$, while the stellar function of a photon-subtracted state $\hat{a}|\psi\rangle$ is given by $z \mapsto \partial_z F^\star_\psi(z)$. In particular, photon-added states are always non-Gaussian, since 0 is a root of their stellar function, while photon-subtracted states can be Gaussian (e.g., the Fock state $|1\rangle$, for which $\hat{a}|1\rangle = |0\rangle$, or the coherent states $|\alpha\rangle$, for $\alpha \in \mathbb{C}$, for which $\hat{a}|\alpha\rangle = \alpha|\alpha\rangle$).

Any operator written as a power series in $\hat{a}^\dagger$ and $\hat{a}$ thus has a stellar representation obtained by taking the same power series in the operator multiplication by z and the operator derivative with respect to z, which corresponds to its effect on the stellar function of a state it is acting on. For example, the photon number operator $\hat{n} = \hat{a}^\dagger\hat{a}$ acts on the stellar function as

$$F^\star_\psi(z) \mapsto z\partial_z F^\star_\psi(z). \tag{2.36}$$

For various operators however, the corresponding stellar representation may be expressed more concisely than with a power series. We give a few examples in what follows.

The displacement and squeeze operators satisfy the following commutation rules (see section 1.3)

$$\begin{aligned} \hat{D}(\alpha)\hat{a}^\dagger\hat{D}^\dagger(\alpha) &= \hat{a}^\dagger - \alpha^* \\ \hat{S}(\xi)\hat{a}^\dagger\hat{S}^\dagger(\xi) &= c_r\hat{a}^\dagger + s_r e^{i\theta}\hat{a}, \end{aligned} \tag{2.37}$$

where $\alpha, \xi = re^{i\theta} \in \mathbb{C}$, with $c_r = \cosh r$ and $s_r = \sinh r$. For all $|\psi\rangle = \sum_{n\geq 0} \psi_n |n\rangle$ we thus have

$$\begin{aligned}
\hat{D}(\alpha)|\psi\rangle &= \hat{D}(\alpha) F^\star_\psi(\hat{a}^\dagger)|0\rangle \\
&= \sum_{n\geq 0} \frac{\psi_n}{\sqrt{n!}} \hat{D}(\alpha)(\hat{a}^\dagger)^n |0\rangle \\
&= \sum_{n\geq 0} \frac{\psi_n}{\sqrt{n!}} (\hat{a}^\dagger - \alpha^*)^n \hat{D}(\alpha)|0\rangle \\
&= F^\star_\psi(\hat{a}^\dagger - \alpha^*)|\alpha\rangle \\
&= F^\star_\psi(\hat{a}^\dagger - \alpha^*) e^{\alpha\hat{a}^\dagger - \frac{1}{2}|\alpha|^2} |0\rangle ,
\end{aligned} \tag{2.38}$$

where we used Eq. (2.3) in the first line, Eq. (2.1) in the second line, Eq. (2.37) in the third line and Eq. (2.9) in the last line. Hence, with Lemma 2.1, the displacement operator $\hat{D}(\alpha)$ acts on the stellar function as

$$F^\star_\psi(z) \mapsto e^{\alpha z - \frac{1}{2}|\alpha|^2} F^\star_\psi(z - \alpha^*), \tag{2.39}$$

for all $\alpha \in \mathbb{C}$. Similarly, for all $|\psi\rangle$ we have

$$\begin{aligned}
\hat{S}(\xi)|\psi\rangle &= \hat{S}(\xi) F^\star_\psi(\hat{a}^\dagger)|0\rangle \\
&= F^\star_\psi(c_r\hat{a}^\dagger + s_r e^{i\theta}\hat{a})\hat{S}(\xi)|0\rangle \\
&= F^\star_\psi(c_r\hat{a}^\dagger + s_r e^{i\theta}\hat{a})|\xi\rangle \\
&= \frac{1}{\sqrt{c_r}} F^\star_\psi(c_r\hat{a}^\dagger + s_r e^{i\theta}\hat{a}) e^{-\frac{1}{2}e^{-i\theta}t_r(\hat{a}^\dagger)^2} |0\rangle ,
\end{aligned} \tag{2.40}$$

where $\xi = re^{i\theta}$ with $c_r = \cosh r$, $s_r = \sinh r$ and $t_r = \tanh r$, and where we used Eq. (2.3) in the first line, Eq. (2.37) in the second line and Eq. (2.10) in the last line. Hence, with Lemma 2.1, the squeezing operator $\hat{S}(\xi)$ acts on the stellar function as

$$F^\star_\psi(z) \mapsto \frac{1}{\sqrt{c_r}} F^\star_\psi(c_r z + s_r e^{i\theta}\partial_z) e^{-\frac{1}{2}e^{-i\theta}t_r z^2}, \tag{2.41}$$

for all $\xi = re^{i\theta} \in \mathbb{C}$.

The POVM corresponding to a threshold detection is $\{|0\rangle\langle 0|, \mathbb{1} - |0\rangle\langle 0|\}$. The projector onto the vacuum acts as

$$|\psi\rangle \mapsto \langle 0|\psi\rangle |0\rangle , \tag{2.42}$$

so it maps the stellar function of a state $|\psi\rangle$ as

$$F^\star_\psi(z) \mapsto \langle 0|\psi\rangle F^\star_{|0\rangle}(z). \tag{2.43}$$

We have $\langle 0|\psi\rangle = F^\star_\psi(0)$ and $F^\star_{|0\rangle}(z) = 1$, so the projector $|0\rangle\langle 0|$ acts on the stellar function as

$$F^\star_\psi(z) \mapsto F^\star_\psi(0), \tag{2.44}$$

while the projector $\mathbb{1} - |0\rangle\langle 0|$ acts as

$$F^\star_\psi(z) \mapsto F^\star_\psi(z) - F^\star_\psi(0). \tag{2.45}$$

In particular, the click of a threshold detector projects the measured state onto a non-Gaussian state for which 0 is a root of the stellar function. This is consistent with the fact that the measured state is orthogonal to the vacuum state—a coherent state of amplitude 0—after being projected onto the support of $\mathbb{1} - |0\rangle\langle 0|$.

Finally, the parity operator $\hat{\Pi} = (-1)^{\hat{a}^\dagger \hat{a}} = e^{i\pi\hat{n}}$ maps the Fock state $|n\rangle$ to $(-1)^n |n\rangle$, for all $n \in \mathbb{N}$. Hence, it acts on the stellar function as

$$F_\psi^\star(z) \mapsto F_\psi^\star(-z). \tag{2.46}$$

by Eq. (2.1).

2.2 The stellar hierarchy

2.2.1 The stellar rank

The Hilbert space $\mathscr{H}$ is naturally partitioned into sets of states whose stellar functions—or equivalently Husimi Q function—have the same number of zeros counted with multiplicity. We introduce the following related definition:

Definition 2.2 (Stellar rank)**.** The *stellar rank* $r^\star(\psi)$ of a pure single-mode normalised quantum state $|\psi\rangle \in \mathscr{H}$ is defined as the number of zeros of its stellar function $F_\psi^\star$, counted with multiplicity.

By analogy with the Schmidt rank in entanglement theory [TH00], we define the stellar rank of a mixed state ρ as

$$r^\star(\rho) := \inf_{p_i, \psi_i} \sup r^\star(\psi_i), \tag{2.47}$$

where the infimum is over the statistical ensembles $\{p_i, \psi_i\}$ such that $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. In particular, a mixed quantum state has nonzero rank if and only if it cannot be written as a mixture of Gaussian states.

We introduce hereafter the notation $\overline{\mathbb{N}} = \mathbb{N} \cup \{+\infty\}$, so that $r^\star(\psi) \in \overline{\mathbb{N}}$, and extend naturally the ordering from $\mathbb{N}$ to $\overline{\mathbb{N}}$, with the convention $N < +\infty \Leftrightarrow N \in \mathbb{N}$. For $N \in \overline{\mathbb{N}}$, we define

$$R_N := \{|\psi\rangle \in \mathscr{H},\ r^\star(\psi) = N\} \tag{2.48}$$

the set of states with stellar rank equal to N. The *stellar hierarchy* is the hierarchy of states induced by the stellar rank (Fig 2.3). The following properties are easily obtained:

- By Lemma 2.1, if $M \neq N$ then $R_M \cap R_N = \varnothing$, for all $M, N \in \overline{\mathbb{N}}$, so all the ranks in the stellar hierarchy are disjoint.
- We have $\mathscr{H} = \bigcup_{N \in \overline{\mathbb{N}}} R_N$, i.e., the stellar hierarchy covers the whole space of normalised states, and the set of states of finite stellar rank is given by $\bigcup_{N \in \mathbb{N}} R_N$.

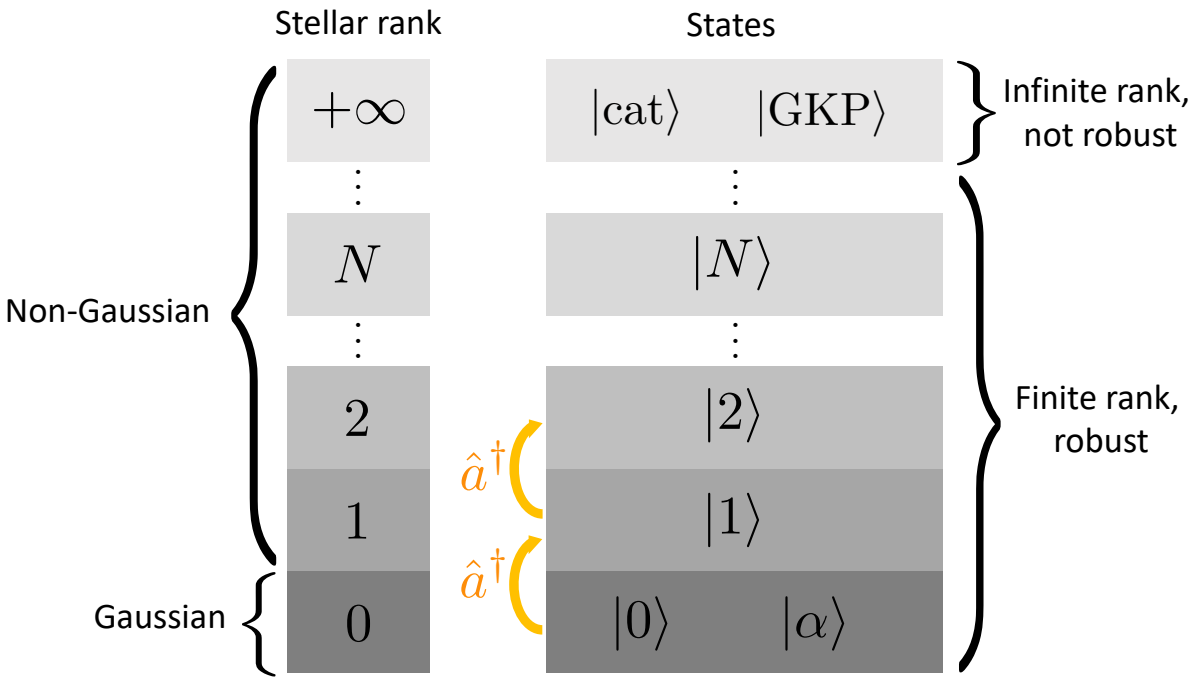

Figure 2.3: The stellar hierarchy of single-mode normalised quantum states. Each rank N contains states obtained from the vacuum with N single photon additions and Gaussian unitary operations (Theorem 2.2). The states of finite rank are robust, while the states of infinite rank are not (section 2.3).

- By Theorem 2.1, the rank zero of the stellar hierarchy R_0 is the set of single-mode normalised pure Gaussian states, and non-Gaussian states populate all higher ranks.
- For all $N \in \mathbb{N}$,the Fock state $|N\rangle$ is of stellar rank N, by Eq. (2.11), while cat states are of infinite stellar rank, by Lemma 2.2, so all ranks are non empty.

In the following, we investigate further properties of the stellar hierarchy. We prove a first general decomposition result for pure states of finite stellar rank:

Theorem 2.2. *Let $|\psi\rangle \in \bigcup_{N\in\mathbb{N}} R_N$ be a pure state of finite stellar rank. Let $\{\alpha_1, \ldots, \alpha_{r^\star(\psi)}\}$ be the roots of the Husimi Q function of $|\psi\rangle$, counted with multiplicity. Then,*

$$|\psi\rangle = \frac{1}{\mathcal{N}} \left[\prod_{n=1}^{r^\star(\psi)} \hat{D}(\alpha_n) \hat{a}^\dagger \hat{D}^\dagger(\alpha_n) \right] |G_\psi\rangle \,, \tag{2.49}$$

where $\hat{D}(\alpha)$ is a displacement operator, $|G_\psi\rangle$ is a Gaussian state, and $\mathcal{N}$ is a normalisation constant. Moreover, this decomposition is unique up to reordering of the roots.

Proof. We consider a state $|\psi\rangle$ of finite stellar rank $r^\star(\psi) \in \mathbb{N}$. Its stellar function is an analytic function over the complex plane of order less or equal to 2, so by Hadamard-Weierstrass factorization theorem (Theorem 1.2),

$$F^\star_\psi(z) = z^k \left[\prod_{n=1}^{r^\star(\psi)-k} \left(1 - \frac{z}{z_n^*}\right) e^{\frac{z}{z_n^*} + \frac{1}{2}\left(\frac{z}{z_n^*}\right)^2} \right] e^{g_0 + g_1 z + g_2 z^2}, \tag{2.50}$$

for all $z \in \mathbb{C}$, where $k \in \mathbb{N}$ is the multiplicity of 0 as a root of $F^\star_\psi$, where the $\{z_n\}$ are the non-zero roots of Q_ψ counted with multiplicity (i.e., the $\{z_n^*\}$ are the non-zero roots of $F^\star_\psi$ counted with multiplicity), and where $g_0, g_1, g_2 \in \mathbb{C}$. Let us introduce for brevity $m = r^\star(\psi) - k \in \mathbb{N}$. Because the product in the above equation is finite, we need not worry about convergence of individual factors, and we may reorder the expression at will. We obtain

$$\begin{aligned} F^\star_\psi(z) &= z^k \prod_{n=1}^{m} \left(1 - \frac{z}{z_n^*}\right) \cdot \prod_{n=1}^{m} e^{\frac{z}{z_n^*} + \frac{1}{2}\left(\frac{z}{z_n^*}\right)^2} \cdot e^{g_0 + g_1 z + g_2 z^2} \\ &= z^k \prod_{n=1}^{m} \left(1 - \frac{z}{z_n^*}\right) \cdot e^{g_0 + \left(g_1 + \sum_{n=1}^m \frac{1}{z_n^*}\right) z + \left(g_2 + \frac{1}{2}\sum_{n=1}^m \frac{1}{(z_n^*)^2}\right) z^2} \\ &= \frac{(-1)^m}{\prod_{n=1}^m z_n^*} \left[z^k \prod_{n=1}^{m} \left(z - z_n^*\right) \right] \cdot e^{g_0 + \left(g_1 + \sum_{n=1}^m \frac{1}{z_n^*}\right) z + \left(g_2 + \frac{1}{2}\sum_{n=1}^m \frac{1}{(z_n^*)^2}\right) z^2}. \end{aligned} \tag{2.51}$$

With Eqs. (2.3) and (2.37), we obtain, for all $\alpha \in \mathbb{C}$,

$$\begin{aligned} |\psi\rangle &= F^\star_\psi(\hat{a}^\dagger)|0\rangle \\ &= \frac{(-1)^m}{\prod_{n=1}^m z_n^*} \left[(\hat{a}^\dagger)^k \prod_{n=1}^{m} \left(\hat{a}^\dagger - z_n^*\right) \right] \cdot e^{g_0 + \left(g_1 + \sum_{n=1}^m \frac{1}{z_n^*}\right)\hat{a}^\dagger + \left(g_2 + \frac{1}{2}\sum_{n=1}^M \frac{1}{(z_n^*)^2}\right)(\hat{a}^\dagger)^2} |0\rangle \\ &= \frac{(-1)^m}{\prod_{n=1}^m z_n^*} \left[(\hat{a}^\dagger)^k \prod_{n=1}^{m} \hat{D}(z_n)\hat{a}^\dagger\hat{D}^\dagger(z_n) \right] \cdot e^{g_0 + \left(g_1 + \sum_{n=1}^m \frac{1}{z_n^*}\right)\hat{a}^\dagger + \left(g_2 + \frac{1}{2}\sum_{n=1}^m \frac{1}{(z_n^*)^2}\right)(\hat{a}^\dagger)^2} |0\rangle. \end{aligned} \tag{2.52}$$

Gouping the non-zero roots $\{z_n\}$ and the k zero roots into the set of zeros counted with multiplicity $\{\alpha_n\}$, we obtain

$$|\psi\rangle = \frac{(-1)^m}{\prod_{n=1}^m z_n^*} \left[\prod_{n=1}^{r^\star(\psi)} \hat{D}(\alpha_n)\hat{a}^\dagger\hat{D}^\dagger(\alpha_n) \right] \cdot e^{g_0 + \left(g_1 + \sum_{n=1}^m \frac{1}{z_n^*}\right)\hat{a}^\dagger + \left(g_2 + \frac{1}{2}\sum_{n=1}^m \frac{1}{(z_n^*)^2}\right)(\hat{a}^\dagger)^2} |0\rangle. \tag{2.53}$$

The state

$$e^{g_0 + \left(g_1 + \sum_{n=1}^m \frac{1}{z_n^*}\right)\hat{a}^\dagger + \left(g_2 + \frac{1}{2}\sum_{n=1}^m \frac{1}{(z_n^*)^2}\right)(\hat{a}^\dagger)^2} |0\rangle \tag{2.54}$$

is a (non normalised) Gaussian state, by Eq. (2.7) and Lemma 2.1. We finally obtain

$$|\psi\rangle = \frac{1}{\mathcal{N}} \left[\prod_{n=1}^{r^\star(\psi)} \hat{D}(\alpha_n)\hat{a}^\dagger\hat{D}^\dagger(\alpha_n) \right] |G_\psi\rangle, \tag{2.55}$$

where $\mathcal{N}$ is a normalisation constant, and $|G_\psi\rangle$ is a Gaussian state. The decomposition is unique by Lemma 2.1 (up to a reordering of the roots).

■

This decomposition implies that any state of finite stellar rank may be obtained from a Gaussian state by successive applications of the creation operator at different locations in phase space, given by the zeros of the Husimi Q function. Experimentally, this corresponds to the probabilistic non-Gaussian operation of single-photon addition [ZVB04, MA10, WSPT18]. Using this decomposition, we obtain the following property for the stellar rank:

Theorem 2.3. *A unitary operation is Gaussian if and only if it leaves the stellar rank invariant.*

Proof. If a unitary operation leaves the stellar rank invariant, it maps in particular all pure states of stellar rank zero to pure states of stellar rank zero, i.e., all Gaussian states to Gaussian states, so it is a Gaussian operation.

Reciprocally, let us show that Gaussian unitary operations leave the stellar rank invariant. We first consider finite stellar rank pure states. Let $|\psi\rangle$ be such a state. By Theorem 2,

$$|\psi\rangle = P_\psi(\hat{a}^\dagger)|G_\psi\rangle\,, \tag{2.56}$$

where P_ψ is a polynomial of degree $r^\star(\psi)$ and $|G_\psi\rangle$ is a Gaussian state. By Eq. (2.37) and by linearity we have

$$|\psi_\alpha\rangle := \hat{D}(\alpha)|\psi\rangle = \hat{P}_\psi(\hat{a}^\dagger - \alpha^*)\hat{D}(\alpha)|G_\psi\rangle\,, \tag{2.57}$$

and

$$|\psi_\xi\rangle := \hat{S}(\xi)|\psi\rangle = P_\psi(c_r\hat{a}^\dagger + s_r e^{i\theta}\hat{a})\hat{S}(\xi)|G_\psi\rangle\,, \tag{2.58}$$

where $\xi = re^{i\theta}$. By Eq. (2.39), during a displacement of α, the stellar function of $|\psi\rangle$ is modified as

$$F^\star_\psi(z) \to F^\star_{\psi,\alpha}(z) = e^{z\alpha - \frac{1}{2}|\alpha|^2}F^\star_\psi(z-\alpha^*) = P_\psi(z-\alpha^*)G^\star_\alpha(z), \tag{2.59}$$

where $G^\star_\alpha(z)$ is the Gaussian stellar function corresponding to the Gaussian state $\hat{D}(\alpha)|G_\psi\rangle$. Moreover, by Eq. (2.41), during a squeezing of ξ, the stellar function of $|\psi\rangle$ is modified as

$$\begin{aligned} F^\star_\psi(z) \to F^\star_{\psi,\xi}(z) &= P_\psi\left(c_r z + s_r e^{i\theta}\partial_z\right)G^\star_\xi(z) \\ &= Q_{\psi,r}(z)G^\star_\xi(z), \end{aligned} \tag{2.60}$$

where $G^\star_\xi(z)$ is the Gaussian stellar function corresponding to the Gaussian state $\hat{S}(\xi)|G_\psi\rangle$, and where $Q_{\psi,\xi}(z) = G^{\star-1}_\xi(z)\left[P_\psi\left(c_r z + s_r e^{i\theta}\partial_z\right)G^\star_\xi(z)\right]$ is a polynomial. Let us compute the leading coefficient of $Q_{\psi,\xi}$. Writing p the leading coefficient of P_ψ, and $N = r^\star(\psi)$ its degree

for brevity, the leading coefficient of $Q_{\psi,\xi}$ is given by the leading coefficient of

$$G_\xi^{\star -1}(z)\left[p\left(c_r z + s_r e^{i\theta}\partial_z\right)^N G_\xi^\star(z)\right]. \tag{2.61}$$

Let us write $G_\xi^\star(z) = e^{-\frac{1}{2}az^2+bz+c}$, as in Eq. (2.7). The leading coefficient of $Q_{\psi,\xi}$ may then be obtained as the leading coefficient of

$$e^{\frac{1}{2}az^2}\left[p\left(c_r z + s_r e^{i\theta}\partial_z\right)^N e^{-\frac{1}{2}az^2}\right]. \tag{2.62}$$

For all x, λ, we have [Wys17]

$$(x+\lambda\partial_x)^N = \sum_{n=0}^{\lfloor \frac{N}{2}\rfloor} \frac{N!\lambda^n}{(N-2n)!n!2^n} \sum_{k=0}^{N-2n} \binom{N-2n}{k} x^k \partial_x^{N-2n-k}, \tag{2.63}$$

so the leading coefficient of $Q_{\psi,\xi}$ is equal to the leading coefficient of:

$$p c_r^N \sum_{n=0}^{\lfloor \frac{N}{2}\rfloor} z^{N-2n}(1-a)^{N-2n} \frac{N! t_r^n e^{in\theta}}{(N-2n)!n!2^n}, \tag{2.64}$$

where $t_r = \tanh r$. Finally, taking the leading coefficient in z of this expression, corresponding to $n = 0$, gives

$$p c_r^N (1-a)^N. \tag{2.65}$$

It is non-zero unless $a = 1$, which corresponds to an infinite value for the modulus r of the squeezing parameter $\xi = re^{i\theta}$ by Eq. (2.7). Hence the polynomials P_ψ and $Q_{\psi,\xi}$ have the same degree. This shows that a finite number of zeros is not modified by Gaussian operations.

Gaussian operations also map states with infinite number of zeros to states with infinite number of zeros. Indeed, assuming there exist a state $|\phi\rangle$ with an infinite number of zeros which is mapped by a Gaussian operation $\hat{G}$ to a state $|\psi\rangle$ with a finite number of zeros, then $\hat{G}^\dagger$ would map $|\psi\rangle$ to $|\phi\rangle$, thus changing the (finite) number of zeros of $F_\psi^\star$, which would be in contradiction with the previous proof. Hence Gaussian unitary operations leave the stellar rank of pure states invariant.

Now by Eq. (2.47), the stellar rank of a mixed state ρ is given by

$$r^\star(\rho) = \inf_{p_i,\psi_i} \sup r^\star(\psi_i), \tag{2.66}$$

where the infimum is over the statistical ensembles such that $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. For $\hat{G}$ a unitary Gaussian operation,

$$\begin{aligned} r^\star(\hat{G}\rho\hat{G}^\dagger) &= \inf_{p_i,\psi_i} \sup r^\star(\hat{G}\psi_i) \\ &= \inf_{p_i,\psi_i} \sup r^\star(\psi_i) \end{aligned} \tag{2.67}$$

$$= r^\star(\rho),$$

where we used in the second line the fact that Gaussian unitary operations leave the stellar rank of pure states invariant. Hence, Gaussian unitary operations leave the stellar rank invariant.

■

An interesting consequence is that the number of single-photon additions in the decomposition of Theorem 2.2 is minimal. Indeed, if a quantum state is obtained from the vacuum by successive applications of Gaussian operations and single-photon additions, then its stellar rank is exactly the number of photon additions, because each single-photon addition increases by one its stellar rank—it adds a zero to the stellar function at zero—while each Gaussian operation leaves the stellar rank invariant by Theorem 2.3. Hence, the stellar rank is a measure of the non-Gaussian properties of a quantum state which may be interpreted as a minimal non-Gaussian operational cost, in terms of single-photon additions, for engineering the state from the vacuum.

2.2.2 Gaussian convertibility

Now that the first properties of the stellar hierarchy are laid out, we consider the convertibility of quantum states using Gaussian unitary operations:

Definition 2.3 (Gaussian convertibility)**.** Two states $|\phi\rangle$ and $|\psi\rangle$ are *Gaussian-convertible* if there exists a Gaussian unitary operation $\hat{G}$ such that $|\psi\rangle = \hat{G}\,|\phi\rangle$.

Note that this notion is different from the notion of Gaussian conversion introduced in [YBT+18], which denotes the conversion of Gaussian states with passive linear optics, and a subclass of Gaussian measurements and feed-forward.

Gaussian convertibility defines an equivalence relation in $\mathscr{H}$. By Theorem 2.3, having the same stellar rank is a necessary condition for Gaussian convertibility. However, this condition is not sufficient. In order to derive the equivalence classes for Gaussian convertibility, we introduce the following definition:

Definition 2.4 (Core state)**.** *Core states* are defined as the single-mode normalised pure quantum states which have a polynomial stellar function.

By Eq. (2.3) and Lemma 2.1, core states are the states with a bounded support over the Fock basis, i.e., finite superpositions of Fock states. These correspond to the minimal non-Gaussian core states introduced in [MF09], in the context of non-Gaussian state engineering.

With this definition, we obtain our following result.

Theorem 2.4. *Let $|\psi\rangle \in \bigcup_{N\in\mathbb{N}} R_N$ be a state of finite stellar rank. Then, there exists a unique core state $|C_\psi\rangle$ such that $|\psi\rangle$ and $|C_\psi\rangle$ are Gaussian-convertible.*

By Theorem 2.2, $|\psi\rangle = P_\psi(\hat{a}^\dagger)|G_\psi\rangle$, where P_ψ is a polynomial of degree $r^\star(\psi)$ and $|G_\psi\rangle = \hat{S}(\xi)\hat{D}(\alpha)|0\rangle$ is a Gaussian state, where $\hat{D}(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}$ is a displacement operator, and $\hat{S}(\xi) = e^{\frac{1}{2}(\xi\hat{a}^2 - \xi^*\hat{a}^{\dagger 2})}$ is a squeezing operator, with $\xi = re^{i\theta}$. Then,

$$|\psi\rangle = \hat{S}(\xi)\hat{D}(\alpha)|C_\psi\rangle = \hat{S}(\xi)\hat{D}(\alpha)F^\star_{C_\psi}(\hat{a}^\dagger)|0\rangle\,, \tag{2.68}$$

where the (polynomial) stellar function of $|C_\psi\rangle$ is given by

$$F^\star_{C_\psi}(z) = P_\psi\left(c_r z - s_r e^{i\theta}\partial_z + c_r\alpha^* - s_r e^{i\theta}\alpha\right)\cdot 1, \tag{2.69}$$

for all $z \in \mathbb{C}$.

Proof. Let $|\psi\rangle \in \bigcup_{N\in\mathbb{N}} R_N$ be a state of finite stellar rank. By Theorem 2.2,

$$|\psi\rangle = P_\psi(\hat{a}^\dagger)|G_\psi\rangle\,, \tag{2.70}$$

where P_ψ is a polynomial of degree $r^\star(\psi)$ and $|G_\psi\rangle = \hat{S}(\xi)\hat{D}(\alpha)|0\rangle$ is a Gaussian state, with $\xi = re^{i\theta}$. Let us define $|C_\psi\rangle = \hat{D}^\dagger(\alpha)\hat{S}^\dagger(\xi)|\psi\rangle$. The states $|\psi\rangle$ and $|C_\psi\rangle$ are Gaussian-convertible. Moreover, from the commutation relations in Eq. (2.37) and by linearity we obtain

$$\begin{aligned}
|C_\psi\rangle &= \hat{D}^\dagger(\alpha)\hat{S}^\dagger(\xi)P_\psi(\hat{a}^\dagger)|G_\psi\rangle \\
&= \hat{D}^\dagger(\alpha)P_\psi\left(c_r\hat{a}^\dagger - s_r e^{i\theta}\hat{a}\right)\hat{S}^\dagger(\xi)|G_\psi\rangle \\
&= P_\psi\left[c_r(\hat{a}^\dagger + \alpha^*) - s_r e^{i\theta}(\hat{a}+\alpha)\right]|0\rangle \\
&= P_\psi\left(c_r\hat{a}^\dagger - s_r e^{i\theta}\hat{a} + c_r\alpha^* - s_r e^{i\theta}\alpha\right)|0\rangle\,,
\end{aligned} \tag{2.71}$$

where we used Eq. (2.70) in the first line. By Eq. (2.35), the stellar operator corresponding to $\hat{a}^\dagger$ is the multiplication by z and the stellar operator corresponding to $\hat{a}$ is the derivative with respect to z. Hence,

$$F^\star_{C_\psi}(z) = P_\psi\left(c_r z - s_r e^{i\theta}\partial_z + c_r\alpha^* - s_r e^{i\theta}\alpha\right)\cdot 1, \tag{2.72}$$

for all $z \in \mathbb{C}$, which is a polynomial function, so the state $|C_\psi\rangle$ is a core state.

In order to conclude the proof, we need to show that $|C_\psi\rangle$ is the unique core state Gaussian-convertible to $|\psi\rangle$. Let $|C\rangle = P_C(\hat{a}^\dagger)|0\rangle$ be a core state Gaussian-convertible to $|\psi\rangle$. The states $|C_\psi\rangle$ and $|C\rangle$ are Gaussian-convertible so there exist $\xi, \alpha \in \mathbb{C}$ such that

$$\begin{aligned}
|C_\psi\rangle &= \hat{S}(\xi)\hat{D}(\alpha)|C\rangle \\
&= \hat{S}(\xi)\hat{D}(\alpha)P_C(\hat{a}^\dagger)|0\rangle \\
&= P_C(c_r\hat{a}^\dagger + s_r e^{i\theta}\hat{a} - \alpha^*)\hat{S}(\xi)\hat{D}(\alpha)|0\rangle\,,
\end{aligned} \tag{2.73}$$

where we used Eq. (2.37). Hence,

$$F^\star_{C_\psi}(z) = P_C(c_r z + s_r e^{i\theta}\partial_z - \alpha^*) G^\star_{\xi,\alpha}(z). \tag{2.74}$$

With Eq. (2.7), this function may be expressed as a polynomial multiplied by a Gaussian function $G^\star_{\xi,\alpha}$. On the other hand $F^\star_{C_\psi}$ is a polynomial, since $|C_\psi\rangle$ is a core state. By comparison of the speed of convergence, this implies that the Gaussian function $G^\star_{\xi,\alpha}$ is constant, i.e., that

$$e^{-i\theta}\tanh r = 0 \quad \text{and} \quad \alpha\sqrt{1-\tanh^2 r} = 0, \tag{2.75}$$

by Eq. (2.7). This in turn implies $\xi = \alpha = 0$, and $|C\rangle = \hat{S}(\xi)\hat{D}(\alpha)|C\rangle = |C_\psi\rangle$.

■

This result has several consequences:

- It implies a second general decomposition result, in addition to Theorem 2.2: by Eq. (2.68), any state of finite stellar rank can be uniquely decomposed as a finite superposition of equally displaced and equally squeezed number states. This shows that the stellar hierarchy matches the genuine n-photon hierarchy introduced in [LSH+18]: a pure state exhibits *genuine n-photon quantum non-Gaussianity* if and only if it has a stellar rank greater or equal to n. Formally, for all $N \in \mathbb{N}$, the set R_N of states of stellar rank equal to N is obtained by the free action of the group of single-mode Gaussian unitary operations on the set of core states of stellar rank N, which is isomorphic to the set of normalised complex polynomials of degree N.

- It shows that two different core states are never Gaussian-convertible, while any state of finite stellar rank is always Gaussian-convertible to a unique core state. This implies that equivalence classes for Gaussian convertibility for states of finite stellar rank correspond to the orbits of core states under Gaussian operations.

- It gives an analytic way to check if two states of finite stellar rank are Gaussian-convertible, given their stellar functions, by checking with Eq. (2.69) if they share the same core state.

- It shows that photon-subtracting a state of finite stellar rank, which amounts to derivating its stellar function, can either decrease its stellar rank by 1, leave it invariant, or increase it by 1, depending on whether the Gaussian operation which converts the state to its core state is either the identity, a displacement, or a Gaussian operation with nonzero squeezing parameter. In particular, this implies that the stellar rank is a lower bound on the number of photon subtractions necessary to enginer a state from the vacuum, together with Gaussian unitary operations.

We consider the following simple example to illustrate the use of Theorem 2.4 for determining Gaussian convertibility: a photon-subtracted squeezed state, a photon-added squeezed state and a single-photon Fock state. We write $|\phi\rangle = -\frac{1}{s_\xi}\hat{a}\,|\xi\rangle$ a normalised photon-subtracted squeezed vacuum state and $|\psi\rangle = \frac{1}{c_\xi}\hat{a}^\dagger\,|\xi\rangle$ a normalised photon-added squeezed vacuum state, with $\xi \in \mathbb{R}^*$. We write also $|\chi\rangle = |1\rangle$ a single-photon Fock state. Using Eq. (2.7) and Eq. (2.35), we obtain for all $z \in \mathbb{C}$

$$\begin{aligned} F^\star_\phi(z) &= -\frac{1}{s_\xi}\partial_z\left[e^{-\frac{1}{2}t_\xi z^2}\right] \\ &= \frac{z}{c_\xi}e^{-\frac{1}{2}t_\xi z^2}, \end{aligned} \tag{2.76}$$

and

$$F^\star_\psi(z) = \frac{z}{c_\xi}e^{-\frac{1}{2}t_\xi z^2}, \tag{2.77}$$

where $c_\xi = \cosh\xi$ and $t_\xi = \tanh\xi$. Hence $F^\star_\phi = F^\star_\psi$, so the states $|\phi\rangle$ and $|\psi\rangle$ are actually equal.

We also have $F^\star_\chi(z) = z$. With the notations of Theorem 2.4, we have $r_\phi = \xi$, $r_\chi = 0$, $\theta_\phi = \theta_\chi = \alpha_\phi = \alpha_\chi = 0$, $\hat{G}_\phi = \hat{S}(\xi)$, $\hat{G}_\chi = \mathbb{1}$, $P_\phi(z) = \frac{z}{c_\xi}$, and $P_\chi(z) = z$, so for all $z \in \mathbb{C}$,

$$\begin{aligned} P_\phi\left(c_{r_\phi}z - s_{r_\phi}e^{i\theta_\phi}\partial_z + c_{r_\phi}\alpha^*_\phi - s_{r_\phi}e^{i\theta_\phi}\alpha_\phi\right)\cdot 1 &= \frac{1}{c_\xi}\left(c_\xi z - s_\xi\partial_z\right)\cdot 1 \\ &= z, \end{aligned} \tag{2.78}$$

and

$$\begin{aligned} P_\chi\left(c_{r_\chi}z - s_{r_\chi}e^{i\theta_\chi}\partial_z + c_{r_\chi}\alpha^*_\chi - s_{r_\chi}e^{i\theta_\chi}\alpha_\chi\right)\cdot 1 = z\cdot 1 \\ = z, \end{aligned} \tag{2.79}$$

thus $|\phi\rangle$ and $|\chi\rangle$ share the same core state. By Theorem 2.4, this means that $|\phi\rangle$ and $|\chi\rangle$ are Gaussian-convertible, and we have $|\phi\rangle = \hat{G}_\phi\hat{G}^\dagger_\chi\,|\chi\rangle$, where

$$\hat{G}_\phi\hat{G}^\dagger_\chi = \hat{S}(\xi). \tag{2.80}$$

Using Eq. (2.37) confirms indeed that

$$-\frac{1}{s_\xi}\hat{a}\,|\xi\rangle = \frac{1}{c_\xi}\hat{a}^\dagger\,|\xi\rangle = \hat{S}(\xi)|1\rangle\,. \tag{2.81}$$

2.3 Robustness of non-Gaussian states

The stellar hierarchy provides a ranking of non-Gaussian states, in terms of the minimal number of photons additions necessary to engineer them. However, for this hierarchy to be relevant in realistic experimental scenarios, it has to be robust to small deviations. We consider this formally in what follows and analyse the robustness properties of the stellar hierarchy.

2.3.1 Definitions

We introduce the following definition:

Definition 2.5 (Stellar robustness)**.** Let $|\psi\rangle \in \mathscr{H}$. The *stellar robustness* of the state $|\psi\rangle$ is defined as

$$R^\star(\psi) := \inf_{r^\star(\phi)<r^\star(\psi)} D(\phi,\psi), \tag{2.82}$$

where D denotes the trace distance and where the infimum is over all states $|\phi\rangle \in \mathscr{H}$ such that $r^\star(\phi) < r^\star(\psi)$.

The stellar robustness quantifies how much one has to deviate from a quantum state in trace distance to find another quantum state of lower stellar rank: states with a positive stellar robustness will be referred to as *robust*. The stellar robustness inherits the property of invariance under Gaussian operations of the stellar rank, because the trace distance between two states is invariant under unitary operations. Because of its operational properties, the choice of the trace distance is especially relevant in the context of non-Gaussian state engineering and quantum computing with non-Gaussian states.

A similar notion, though more restricted, is the quantum non-Gaussian depth [SLH+18] which quantifies the maximum attenuation applicable on a quantum state, after which quantum non-Gaussianity can still be witnessed. A natural generalisation of the notion of stellar robustness is the following:

Definition 2.6 (k-robustness)**.** Let $|\psi\rangle \in \mathscr{H}$. For all $k \in \overline{\mathbb{N}}^*$, the *$k$-robustness* of the state $|\psi\rangle$ is defined as

$$R_k^\star(\psi) := \inf_{r^\star(\phi)<k} D(\phi,\psi), \tag{2.83}$$

where D denotes the trace distance and where the infimum is over all states $|\phi\rangle \in \mathscr{H}$ such that $r^\star(\phi) < k$.

For all $k \in \mathbb{N}^*$, the k-robustness quantifies how much one has to deviate from a quantum state in trace distance to find another quantum state which as a stellar rank between 0 and $k-1$. States with a positive $R_k^\star$ will be referred to as *robust with respect to states of stellar rank lower than k*. When $k = +\infty$, the ∞-robustness quantifies how much one has to deviate from a quantum state in trace distance to find another quantum state of finite stellar rank. Note that the stellar robustness satisfies $R^\star(\psi) = R^\star_{r^\star(\psi)}(\psi)$. We introduce the related definition:

Definition 2.7 (Robustness profile)**.** Let $|\psi\rangle \in \mathscr{H}$. The *robustness profile* of the state $|\psi\rangle$ is defined as

$$\mathscr{R}(\psi) := \left(R_k^\star(\psi)\right)_{k\in\mathbb{N}^*}. \tag{2.84}$$

The robustness profile is the sequence of k-robustnesses for all $k \in \mathbb{N}^*$. This profile describes how hard a non-Gaussian state is to produce experimentally, using photon additions.

A dual notion to the robustness is the following:

Definition 2.8 (Smoothed non-Gaussianity of formation)**.** Let ρ be a single-mode normalised state, and let $\epsilon > 0$. The *ϵ-smoothed non-Gaussianity of formation* $\mathcal{NGF}_\epsilon(\rho)$ is defined as the minimal stellar rank of the states σ that are ϵ-close to ρ in trace distance. Formally,

$$\mathcal{NGF}_\epsilon(\rho) := \inf_\sigma \left\{ r^\star(\sigma), \text{ s.t. } D(\rho,\sigma) \leq \epsilon \right\}, \tag{2.85}$$

where D denotes the trace distance.

The infimum in the definition is also a minimum, since the set considered only contains integer values and is lower bounded by zero. That minimum is not necessarily attained for the energy cut-off state (consider, e.g., a Gaussian pure state).

The smoothed non-Gaussianity of formation can be obtained directly from the robustness profile and gives a smoothed version of the stellar rank, dual to the robustness. By Theorem 2.2, it quantifies the minimal number of single-photon additions that need to be applied to a Gaussian state in order to obtain a state ϵ-close to a target state, and provides an operational cost measure for non-Gaussian resource states, which is also invariant under Gaussian operations.

The robustness is related to the fidelity by the following result:

Lemma 2.4. *Let $|\psi\rangle \in \mathcal{H}$. For all $k \in \overline{\mathbb{N}}^*$,*

$$\sup_{r^\star(\rho)<k} F(\rho,\psi) = 1 - [R_k^\star(\psi)]^2, \tag{2.86}$$

where F is the fidelity.

Proof. For any pure state $|\psi\rangle \in \mathcal{H}$, and any set of pure states $\mathcal{X}$, we have

$$\begin{aligned} \sup_{\substack{\rho=\sum p_i |\phi\rangle_i\langle\phi_i| \\ \sum p_i = 1, \phi_i \in \mathcal{X}}} F(\rho,\psi) &= \sup_{\substack{\rho=\sum p_i |\phi\rangle_i\langle\phi_i| \\ \sum p_i = 1, \phi_i \in \mathcal{X}}} \langle\psi|\rho|\psi\rangle \\ &= \sup_{\sum p_i = 1} \sup_{\phi_i \in \mathcal{X}} \sum p_i |\langle\phi_i|\psi\rangle|^2 \\ &= \sup_{\phi\in\mathcal{X}} |\langle\phi|\psi\rangle|^2 \\ &= \sup_{\phi\in\mathcal{X}} F(\phi,\psi). \end{aligned} \tag{2.87}$$

Hence, for $\mathcal{X}$ the set of pure states of stellar rank less than k,

$$\begin{aligned} R_k^\star(\psi) &= \inf_{r^\star(\phi)<k} D(\phi,\psi) \\ &= \inf_{r^\star(\phi)<k} \sqrt{1 - |\langle\phi|\psi\rangle|^2} \\ &= \sqrt{1 - \sup_{r^\star(\phi)<k} F(\phi,\psi)} \end{aligned} \tag{2.88}$$

$$= \sqrt{1 - \sup_{r^\star(\rho)<k} F(\rho,\psi)},$$

where D denotes the trace distance, where we used the definition of the stellar rank for mixed states (2.47). We finally obtain

$$\sup_{r^\star(\rho)<k} F(\rho,\psi) = 1 - [R_k^\star(\psi)]^2. \tag{2.89}$$

■

Certifying that a (mixed) state ρ has a fidelity greater than $1-[R_k^\star(\psi)]^2$ with a given target pure state $|\psi\rangle$ thus ensures that the state ρ has stellar rank greater or equal to k. However, this is only possible if the two following conditions are met:

- The target state $|\psi\rangle$ is robust with respect to states of stellar rank less than k, i.e., $R_k^\star(\psi) > 0$.
- The value of the k-robustness $R_k^\star(\psi)$ is known.

We consider these two problems in what follows. First, we determine for all $k \in \overline{\mathbb{N}}^*$ which states are robust with respect to states of stellar rank less than k. Then, we show how to compute their k-robustness.

2.3.2 Topology of the stellar hierarchy

Determining which states are robust amounts to characterizing the topology of the stellar hierarchy, with respect to the trace norm. Formally, this topology is summarised by the following result for states of finite stellar rank:

Theorem 2.5. *For all $N \in \mathbb{N}$,*

$$\overline{R_N} = \bigcup_{0\le K\le N} R_K, \tag{2.90}$$

where $\overline{X}$ denotes the closure of X for the trace norm in the set of normalised states of $\mathscr{H}$.

Proof. Recall that the set of normalised pure single-mode states is closed for the trace norm in the whole Hilbert space, since it is the reciprocal image of $\{1\}$ by the trace norm, which is Lipschitz continuous—with Lipschitz constant 1—hence continuous.

For the proof, we fix $N \in \mathbb{N}$. We prove the theorem by showing a double inclusion. We first show that $\bigcup_{K=0}^N R_K \subset \overline{R_N}$, and then that the set $\bigcup_{K=0}^N R_K$ is closed in $\mathscr{H}$ for the trace norm. Since the closure of a set X is the smallest closed set containing X, and given that $R_N \subset \bigcup_{K=0}^N R_K$, this will prove the other inclusion and hence the result.

We have $R_N \subset \overline{R_N}$. Let $|\psi\rangle \in \bigcup_{K=0}^{N-1} R_K$. There exists $K \in \{0,\dots,N-1\}$ such that $r^\star(\psi) = K$. By Theorem 4, there exists a core state $|C_\psi\rangle$, with a polynomial stellar function of degree K,

and a Gaussian operation $\hat{G}_\psi$ such that $|\psi\rangle = \hat{G}_\psi |C_\psi\rangle$. We define the sequence of normalised states

$$|\psi_m\rangle = \sqrt{1-\frac{1}{m}}\,|\psi\rangle + \frac{1}{\sqrt{m}}\hat{G}_\psi |N\rangle\,, \tag{2.91}$$

for $m \geq 1$. We have

$$|\psi_m\rangle = \hat{G}_\psi \left(\sqrt{1-\frac{1}{m}}\,|C_\psi\rangle + \frac{1}{\sqrt{m}}\,|N\rangle\right), \tag{2.92}$$

and the state $\sqrt{1-\frac{1}{m}}\,|C_\psi\rangle + \frac{1}{\sqrt{m}}\,|N\rangle$ is a normalised core state whose stellar function is a polynomial of degree N, hence $|\psi_m\rangle \in R_N$. Moreover, $\{|\psi_m\rangle\}_{m\geq 1}$ converges to $|\psi\rangle$ in trace norm. This shows that $\bigcup_{K=0}^{N} R_K \subset \overline{R_N}$.

We now prove that the set $\bigcup_{K=0}^{N} R_K$ is closed in $\mathscr{H}$ for the trace norm. For $N = 0$ (i.e., showing that the set of Gaussian states is a closed set), this is already a nontrivial result, and a proof may be found, e.g., in [LRW+18].

For all $N \geq 0$, the sketch of the proof is the following: given a converging sequence in $\bigcup_{K=0}^{N} R_K$, we want to show that its limit has a stellar rank less or equal to N. We first use the decomposition result of Theorem 2.4, in order to obtain a sequence of Gaussian operations acting on a sequence of core states of rank less or equal to N. We make use of the compactness of this set of core states to restrict to a unique core state. Then, we show that the squeezing and the displacement parameters of the sequence of Gaussian operations cannot be unbounded. This allows us to conclude by extracting converging subsequences from these parameters.

The trace distance D is induced by the trace norm. Let $\{|\psi_m\rangle\}_{m\in\mathbb{N}} \in \bigcup_{K=0}^{N} R_K$ be a converging sequence for the trace norm, and let $|\psi\rangle \in \mathscr{H}$ be its limit. By Theorem 2.4, there exist a sequence of core states $\{|C_m\rangle\}_{m\in\mathbb{N}}$, with polynomial stellar functions of degrees less or equal to N, and a sequence of Gaussian operations $\{\hat{G}_m\}_{m\in\mathbb{N}}$ such that for all $m \in \mathbb{N}$, $|\psi_m\rangle = \hat{G}_m |C_m\rangle$.

The set of normalised core states with a polynomial stellar function of degree less or equal to N corresponds to the set of normalised states with a support over the Fock basis truncated at N, and is compact for the trace norm in $\mathscr{H}$ (isomorphic to the set of norm 1 vectors in $\mathbb{C}^{N+1}$). Hence, the sequence $\{|C_m\rangle\}_{m\in\mathbb{N}}$ admits a converging subsequence $\{|C_{m_k}\rangle\}_{k\in\mathbb{N}}$. Let the core state $|C\rangle$, with a polynomial stellar function of degree less or equal to N, be its limit. Along this subsequence,

$$|\psi_{m_k}\rangle = \hat{G}_{m_k} |C_{m_k}\rangle\,, \tag{2.93}$$

and we have $\lim_{k\to+\infty} D(|\psi_{m_k}\rangle, |\psi\rangle) = 0$ and $\lim_{k\to+\infty} D(|C_{m_k}\rangle, |C\rangle) = 0$. Moreover, for all

$k \in \mathbb{N}$,

$$\begin{aligned} D(\hat{G}_{m_k}|C\rangle, |\psi\rangle) &\leq D(\hat{G}_{m_k}|C\rangle, |\psi_{m_k}\rangle) + D(|\psi_{m_k}\rangle, |\psi\rangle) \\ &= D(\hat{G}_{m_k}|C\rangle, \hat{G}_{m_k}|C_{m_k}\rangle) + D(|\psi_{m_k}\rangle, |\psi\rangle) \\ &= D(|C\rangle, |C_{m_k}\rangle) + D(|\psi_{m_k}\rangle, |\psi\rangle), \end{aligned} \tag{2.94}$$

where we used the triangular inequality in the first line, Eq. (2.93) in the second line, and the invariance of the trace distance under unitary transformations in the third line. Hence, the sequence $\{\hat{G}_{m_k}|C\rangle\}_{k\in\mathbb{N}}$ converges in trace norm to $|\psi\rangle$. This shows that we can restrict without loss of generality to a unique core state, with a polynomial stellar function of degree less or equal to N, instead of a sequence of such core states.

Let $|C\rangle$ thus be a core state, with a polynomial stellar function of degree K less or equal to N. We write

$$|C\rangle = P_C(\hat{a}^\dagger)|0\rangle = \sum_{n=0}^{K} \frac{p_n}{\sqrt{n!}}|n\rangle, \tag{2.95}$$

with $\sum_{n=0}^{K} \frac{|p_n|^2}{n!} = 1$. Let us consider a converging sequence $\{\hat{G}_m|C\rangle\}_{m\in\mathbb{N}}$, where $\hat{G}_m$ are Gaussian operations, and denote $|\psi\rangle$ its limit. There exists two sequences $\{\xi_m\}_{m\in\mathbb{N}}$ and $\{\alpha_m\}_{m\in\mathbb{N}}$, such that for all $m \in \mathbb{N}$,

$$\hat{G}_m = \hat{S}(\xi_m)\hat{D}(\alpha_m). \tag{2.96}$$

We write $\xi_m = r_m e^{i\theta_m}$, with $r_m \geq 0$, for all $m \in \mathbb{C}$. We may rewrite $\hat{G}_m = \hat{D}(\gamma_m)\hat{S}(\xi_m)$, where for all $m \in \mathbb{N}$,

$$\gamma_m = c_{r_m}\alpha_m + s_{r_m}e^{i\theta_m}\alpha_m^*, \tag{2.97}$$

where $c_{r_m} = \cosh(r_m)$ and $s_{r_m} = \sinh(r_m)$. With these notations, we prove the following result:

Lemma 2.5. *The sequences $\{\xi_m\}_{m\in\mathbb{N}}$ and $\{\gamma_m\}_{m\in\mathbb{N}}$ are bounded.*

Proof. We first compute an upper bound for the Q function of the state $\hat{G}_m|C\rangle$, which we obtain in Eq. (2.114). For $m \in \mathbb{N}$, we have:

$$\begin{aligned} Q_{\hat{G}_m|C\rangle}(z) &= Q_{\hat{D}(\gamma_m)\hat{S}(\xi_m)|C\rangle}(z) \\ &= Q_{\hat{S}(\xi_m)|C\rangle}(z-\gamma_m) \\ &= \frac{e^{-|z-\gamma_m|^2}}{\pi}\left|F^\star_{\hat{S}(\xi_m)|C\rangle}(z^*-\gamma_m^*)\right|^2, \end{aligned} \tag{2.98}$$

for all $z \in \mathbb{C}$.

We have

$$\begin{aligned} \hat{S}(\xi_m)|C\rangle &= \hat{S}(\xi_m)P_C(\hat{a}^\dagger)|0\rangle \\ &= P_C(c_{r_m}\hat{a}^\dagger + s_{r_m}e^{i\theta_m}\hat{a})\hat{S}(\xi_m)|0\rangle. \end{aligned} \tag{2.99}$$

Hence, with Eq. (2.7) and (2.35),

$$\begin{aligned} F^{\star}_{\hat{S}(\xi_m)|C\rangle}(z) &= (1-|t_{r_m}|^2)^{1/4} P_C(c_{r_m} z + s_{r_m} e^{i\theta_m}\partial_z)\cdot e^{-\frac{1}{2}t_{r_m}e^{-i\theta_m}z^2} \\ &= \frac{1}{\sqrt{c_{r_m}}}\sum_{n=0}^{K}\frac{p_n}{\sqrt{n!}}(c_{r_m} z + s_{r_m} e^{i\theta_m}\partial_z)^n\cdot e^{-\frac{1}{2}t_{r_m}e^{-i\theta_m}z^2}. \end{aligned} \tag{2.100}$$

where $t_{r_m} = \tanh(r_m)$.

The Hermite polynomials [AS65] satisfy the following recurrence relation

$$He_{n+1}(z) = z He_n(z) - \partial_z He_n(z), \tag{2.101}$$

for all $n \geq 0$ and all $z \in \mathbb{C}$, and $He_0 = 1$. Setting

$$f_n(z) := e^{\frac{1}{2}t_{r_m}e^{-i\theta_m}z^2}(c_{r_m} z + s_{r_m} e^{i\theta_m}\partial_z)^n\cdot e^{-\frac{1}{2}t_{r_m}e^{-i\theta_m}z^2}, \tag{2.102}$$

we obtain $f_0(z) = 1$, and

$$\begin{aligned} f_{n+1}(z) &= e^{\frac{1}{2}t_{r_m}e^{-i\theta_m}z^2}(c_{r_m} z + s_{r_m} e^{i\theta_m}\partial_z)\left[e^{-\frac{1}{2}t_{r_m}e^{-i\theta_m}z^2} f_n(z)\right] \\ &= \frac{z}{c_{r_m}} f_n(z) + s_{r_m} e^{i\theta_m}\partial_z f_n(z). \end{aligned} \tag{2.103}$$

Hence, with Eq. (2.101), for all $n \geq 0$ and all $z \in \mathbb{C}$,

$$f_n(z) = \lambda_m^{n/2} He_n\left(\frac{z}{c_{r_m}\sqrt{\lambda_m}}\right), \tag{2.104}$$

where we have set $\lambda_m = -e^{i\theta_m}t_{r_m}$. With Eq. (2.100) we thus obtain

$$\begin{aligned} F^{\star}_{\hat{S}(\xi_m)|C\rangle}(z) &= \frac{1}{\sqrt{c_{r_m}}}\sum_{n=0}^{K}\frac{p_n}{\sqrt{n!}} f_n(z)\cdot e^{-\frac{1}{2}t_{r_m}e^{-i\theta_m}z^2} \\ &= \frac{1}{\sqrt{c_{r_m}}}\sum_{n=0}^{K}\frac{p_n\lambda_m^{n/2}}{\sqrt{n!}} He_n\left(\frac{z}{c_{r_m}\sqrt{\lambda_m}}\right) e^{-\frac{1}{2}t_{r_m}e^{-i\theta_m}z^2}. \end{aligned} \tag{2.105}$$

From this and Eq. (2.98) we deduce

$$\begin{aligned} Q_{\hat{G}_m|C\rangle}(z) &= \frac{e^{-|z-\gamma_m|^2}}{\pi c_{r_m}}\left|\sum_{n=0}^{K}\frac{p_n\lambda_m^{n/2}}{\sqrt{n!}} He_n\left(\frac{z^*-\gamma_m^*}{c_{r_m}\sqrt{\lambda_m}}\right) e^{-\frac{1}{2}t_{r_m}e^{-i\theta_m}(z^*-\gamma_m^*)^2}\right|^2 \\ &\leq \frac{e^{-|z-\gamma_m|^2}}{\pi c_{r_m}}\left|e^{-\frac{1}{2}t_{r_m}e^{-i\theta_m}(z^*-\gamma_m^*)^2}\right|^2\sum_{n=0}^{K}\frac{|p_n|^2}{n!}\cdot\sum_{n=0}^{K}\left|\lambda_m^{n/2} He_n\left(\frac{z^*-\gamma_m^*}{c_{r_m}\sqrt{\lambda_m}}\right)\right|^2 \\ &= \frac{1}{\pi c_{r_m}} e^{-|z-\gamma_m|^2-\frac{1}{2}t_{r_m}[e^{i\theta_m}(z-\gamma_m)^2+e^{-i\theta_m}(z^*-\gamma_m^*)^2]}\sum_{n=0}^{K}\left|t_{r_m}^{n/2} He_n\left(\frac{z-\gamma_m}{c_{r_m}\sqrt{\lambda_m^*}}\right)\right|^2, \end{aligned} \tag{2.106}$$

where we used Cauchy-Schwarz inequality in the second line, $|\lambda_m| = t_{r_m}$ and the fact that the coefficients of He_n are real in the third line.

Setting

$$\alpha_m(z) := -\frac{ie^{\frac{1}{2}i\theta_m}}{c_{r_m}}(z-\gamma_m), \tag{2.107}$$

for all $m \in \mathbb{N}$ and for all $z \in \mathbb{C}$, we obtain

$$\begin{aligned} Q_{\hat{G}_m|C\rangle}(z) &\leq \frac{1}{\pi c_{r_m}} e^{-|z-\gamma_m|^2 - \frac{1}{2}t_{r_m}[e^{i\theta_m}(z-\gamma_m)^2 + e^{-i\theta_m}(z^*-\gamma_m^*)^2]} \sum_{n=0}^{K} \left| t_{r_m}^{n/2} He_n\left(\frac{e^{-i\theta_m}\alpha_m(z)}{\sqrt{t_{r_m}}}\right)\right|^2 \\ &= \frac{1}{\pi c_{r_m}} e^{-c_{r_m}^2|\alpha_m(z)|^2 + \frac{1}{2}c_{r_m}s_{r_m}[\alpha_m^2(z)+\alpha_m^{*2}(z)]} \sum_{n=0}^{K} \left| t_{r_m}^{n/2} He_n\left(\frac{e^{-i\theta_m}\alpha_m(z)}{\sqrt{t_{r_m}}}\right)\right|^2 \\ &= \frac{1}{\pi c_{r_m}} e^{-c_{r_m}(c_{r_m}-s_{r_m})x_m^2(z)} e^{-c_{r_m}(c_{r_m}+s_{r_m})y_m^2(z)} \sum_{n=0}^{K} \left| t_{r_m}^{n/2} He_n\left(\frac{e^{-i\theta_m}\alpha_m(z)}{\sqrt{t_{r_m}}}\right)\right|^2, \end{aligned} \tag{2.108}$$

where $\alpha_m(z) = x_m(z) + iy_m(z)$. For all $r \in \mathbb{R}$,

$$c_r(c_r - s_r) = \frac{1}{2}(1+e^{-2r}) > \frac{1}{2}, \tag{2.109}$$

and

$$c_r(c_r + s_r) = \frac{1}{2}(1+e^{2r}) > \frac{1}{2}, \tag{2.110}$$

so with Eq. (2.108) we obtain

$$Q_{\hat{G}_m|C\rangle}(z) \leq \frac{1}{\pi c_{r_m}} e^{-\frac{1}{2}|\alpha_m(z)|^2} \sum_{n=0}^{K} \left| t_{r_m}^{n/2} He_n\left(\frac{e^{-i\theta_m}\alpha_m(z)}{\sqrt{t_{r_m}}}\right)\right|^2. \tag{2.111}$$

Finally, we obtain the following bound for all $n \in \{0,\dots,K\}$:

$$\begin{aligned} \left| t_{r_m}^{n/2} He_n\left(\frac{e^{-i\theta_m}\alpha_m(z)}{\sqrt{t_{r_m}}}\right)\right| &= \left| t_{r_m}^{n/2} \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^k n!}{2^k k!(n-2k)!} \left(\frac{e^{-i\theta_m}\alpha_m(z)}{\sqrt{t_{r_m}}}\right)^{n-2k} \right| \\ &\leq \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{n!}{2^k k!(n-2k)!} t_{r_m}^k |\alpha_m(z)|^{n-2k} \\ &\leq \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{n!}{2^k k!(n-2k)!} |\alpha_m(z)|^{n-2k}, \end{aligned} \tag{2.112}$$

for all $m \in \mathbb{N}$ and all $z \in \mathbb{C}$. Let us define for brevity the polynomial

$$T(X) := \sum_{n=0}^{K} \left(\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{n!}{2^k k!(n-2k)!} X^{n-2k} \right)^2. \tag{2.113}$$

Plugging Eq. (2.112) in Eq. (2.111) yields

$$Q_{\hat{G}_m|C\rangle}(z) \leq \frac{1}{\pi c_{r_m}} e^{-\frac{1}{2}|\alpha_m(z)|^2} T(|\alpha_m(z)|), \tag{2.114}$$

for all $m \in \mathbb{N}$ and all $z \in \mathbb{C}$.

With this bound on the Q function obtained, we may now prove that the sequences $\{\xi_m\}_{m\in\mathbb{N}} = \{r_m e^{i\theta_m}\}_{m\in\mathbb{N}}$ and $\{\gamma_m\}_{m\in\mathbb{N}}$ are bounded.

Assuming that $\{r_m\}_{m\in\mathbb{N}}$ is unbounded implies that it has a subsequence $\{r_{m_k}\}_{k\in\mathbb{N}}$ going to infinity. Since the function $x \mapsto e^{-\frac{1}{2}x^2}T(x)$ is bounded, $Q_{\hat{G}_{m_k}|C\rangle}(z) \to 0$ for all $z \in \mathbb{C}$ when $k \to +\infty$ by Eq. (2.114). But $Q_{\hat{G}_{m_k}|C\rangle}(z) \to Q_\psi(z)$ for all $z \in \mathbb{C}$ when $k \to +\infty$, by property of the convergence in trace norm. This would imply $Q_\psi(z) = 0$ for all $z \in \mathbb{C}$, which is impossible since $|\psi\rangle$ is normalised. Hence $\{r_m\}_{m\in\mathbb{N}}$ is a bounded sequence, and so is $\{\xi_m\}_{m\in\mathbb{N}}$.

With the same reasoning, if $\{|\alpha_m(z)|\}_{m\in\mathbb{N}}$ was unbounded for all $z \in \mathbb{C}$, this would imply by Eq. (2.114) that $Q_\psi(z) = 0$ for all $z \in \mathbb{C}$, giving the same contradiction. Hence, there exists $z_0 \in \mathbb{C}$ such that the sequence $\{|\alpha_m(z_0)|\}_{m\in\mathbb{N}}$ is bounded. By Eq. (2.107), this implies that the sequence $\{\gamma_m\}_{m\in\mathbb{N}}$ is also bounded, since the sequence $\{r_m\}_{m\in\mathbb{N}}$ is bounded.

□

The sequences $\{\xi_m\}_{m\in\mathbb{N}}$ and $\{\gamma_m\}_{m\in\mathbb{N}}$ being bounded, one can consider simultaneously converging subsequences $\{\xi_{m_k}\}_{k\in\mathbb{N}}$ and $\{\gamma_{m_k}\}_{k\in\mathbb{N}}$. We write $\xi = re^{i\theta} = \lim_{k\to\infty}\xi_{m_k}$ and $\gamma = \lim_{k\to\infty}\gamma_{m_k}$. On one hand, we have

$$\begin{aligned} F^\star_{\hat{G}_{m_k}|C\rangle}(z) &= F^\star_{\hat{D}(\gamma_{m_k})\hat{S}(\xi_{m_k})|C\rangle}(z) \\ &= e^{\gamma_{m_k} z - \frac{1}{2}|\gamma_{m_k}|^2} F^\star_{\hat{S}(\xi_{m_k})|C\rangle}(z-\gamma^*_{m_k}) \\ &= \frac{1}{\sqrt{c_{r_{m_k}}}} \sum_{n=0}^{K} \frac{p_n \lambda_{m_k}^{n/2}}{\sqrt{n!}} He_n\left(\frac{z-\gamma^*_{m_k}}{c_{r_{m_k}}\sqrt{\lambda_{m_k}}}\right) e^{-\frac{1}{2}t_{r_{m_k}} e^{-i\theta_{m_k}}(z-\gamma^*_{m_k})^2 + \gamma_{m_k} z - \frac{1}{2}|\gamma_{m_k}|^2}, \end{aligned} \tag{2.115}$$

for all $k \in \mathbb{N}$ and all $z \in \mathbb{C}$, where we have used Eq. (2.59) in the second line, where $\lambda_{m_k} = -e^{i\theta_{m_k}} t_{r_{m_k}}$, and where we have used Eq. (2.105) in the last line. Setting $\lambda = -e^{i\theta} t_r$, we obtain

$$\begin{aligned} \lim_{k\to\infty} F^\star_{\hat{G}_{m_k}|C\rangle}(z) &= \frac{1}{\sqrt{c_r}} \sum_{n=0}^{K} \frac{p_n \lambda^{n/2}}{\sqrt{n!}} He_n\left(\frac{z-\gamma^*}{c_r\sqrt{\lambda}}\right) e^{-\frac{1}{2}t_r e^{-i\theta}(z-\gamma^*)^2 + \gamma z - \frac{1}{2}|\gamma|^2} \\ &= F^\star_{\hat{G}|C\rangle}(z), \end{aligned} \tag{2.116}$$

for all $z \in \mathbb{C}$, where $\hat{G} = \hat{D}(\gamma)\hat{S}(\xi)$, and where the second line comes from reversing the calculations of Eq. (2.115). On the other hand, for all $z \in \mathbb{C}$,

$$\begin{aligned} \lim_{k\to\infty} F^\star_{\hat{G}_{m_k}|C\rangle}(z) &= e^{\frac{1}{2}|z|^2} \lim_{k\to\infty} \langle z^*|\hat{G}_{m_k}|C\rangle \\ &= e^{\frac{1}{2}|z|^2} \langle z^*|\psi\rangle \\ &= F^\star_\psi(z), \end{aligned} \tag{2.117}$$

by property of the convergence in trace norm. Combining Eq. (2.116) and Eq. (2.117) yields

$$F^\star_\psi(z) = F^\star_{\hat{G}|C\rangle}(z), \tag{2.118}$$

for all $z \in \mathbb{C}$. By Lemma 1, this implies that $|\psi\rangle = \hat{G}|C\rangle \in R_K$. This shows that $\overline{\bigcup_{K=0}^{N} R_K} = \bigcup_{K=0}^{N} R_K$, so $\overline{R_N} \subset \bigcup_{K=0}^{N} R_K$, which concludes the proof.

■

This result implies that the set $\bigcup_{0\leq K\leq N} R_K$, containing the states of stellar rank smaller or equal to N, is a closed set in $\mathscr{H}$ for the trace norm, for all $N \in \mathbb{N}$. In particular, since all ranks of the stellar hierarchy are disjoint, for any state of finite rank N, there is no sequence of states of strictly lower stellar rank converging to it. Each state of a given finite stellar rank is thus isolated from all the lower stellar ranks, i.e., there is a ball around it in trace norm which only contains states of equal or higher stellar rank.

Moreover, each state of infinite stellar rank is isolated from states of finite stellar rank lower than N, for all $N \in \mathbb{N}^*$, i.e., there is a ball around it in trace norm which only contains states of stellar rank higher than N.

On the other hand, with the other inclusion, no state of a given finite stellar rank is isolated from any equal or higher stellar rank, i.e., one can always find a sequence of states of any higher rank converging to this state in trace norm.

We also prove the following density result:

Lemma 2.6. *The set of states of finite stellar rank is dense for the trace norm in the set of normalised pure single-mode states:*

$$\overline{\bigcup_{N\in\mathbb{N}} R_N} = \mathscr{H}, \tag{2.119}$$

where $\overline{X}$ denotes the closure of X for the trace norm in the set of normalised states in $\mathscr{H}$.

Proof. Recall that the set of normalised pure single-mode states is closed for the trace norm in the whole Hilbert space, since it is the reciprocal image of $\{1\}$ by the trace norm, which is Lipschitz continuous—with Lipschitz constant 1—hence continuous.

Let $|\psi\rangle \in \mathscr{H}$ be a normalised state. We consider the sequence of normalised cut-off states

$$|\psi_m\rangle = \frac{1}{\sqrt{\mathscr{N}_m}} \sum_{n=0}^{m} \psi_n |n\rangle, \tag{2.120}$$

where $\mathscr{N}_m = \sum_{n=0}^{m} |\psi_n|^2$ is a normalising factor (non-zero for m large enough). All the states $|\psi_m\rangle$ have a finite support over the Fock basis, so their stellar function is a polynomial. Hence $\{|\psi_m\rangle\}_{m\in\mathbb{N}} \in \bigcup_{N\in\mathbb{N}} R_N$.

Moreover, for all $m \in \mathbb{N}$,

$$\begin{aligned} D(\psi_m, \psi) &= \sqrt{1 - |\langle\psi_m|\psi\rangle|^2} \\ &= \sqrt{1 - \sum_{n=0}^{m} |\psi_n|^2} \end{aligned} \tag{2.121}$$

$$= \sqrt{\sum_{n \geq m+1} |\psi_n|^2},$$

where we used that $|\psi\rangle$ and $|\psi_m\rangle$ are pure states in the first line, and the fact that $|\psi\rangle$ is normalised in the third line. Furthermore, $\sum_{n \geq m+1} |\psi_n|^2 \to 0$ when $m \to +\infty$, because $|\psi\rangle$ is normalised. Hence, $\{|\psi_m\rangle\}_{m \in \mathbb{N}}$ converges in trace norm to $|\psi\rangle$, which concludes the proof.

■

This result implies that states of infinite stellar rank are not isolated from lower stellar ranks, unlike states of finite stellar rank. Given a state of infinite stellar rank, there always exists a sequence of states of finite stellar ranks converging to it. However, the ranks of the states in this sequence have to go to infinity, since by Theorem 2.5 states of infinite stellar rank are isolated from states of finite stellar rank lower than N, for all $N \in \mathbb{N}^*$.

The consequences of Theorem 2.5 and Lemma 2.6 for the robustness are summarised with the following result:

Corollary 2.1. *For all $|\psi\rangle \in \bigcup_{N \in \mathbb{N}} R_N$ and for all $k \in \overline{\mathbb{N}}^*$,*

$$\begin{cases} R_k^\star(\psi) > 0 & \text{for } k \leq r^\star(\psi), \quad (2.122a) \\ R_k^\star(\psi) = 0 & \text{for } k > r^\star(\psi). \quad (2.122b) \end{cases}$$

In particular, states of finite stellar rank are robust: for all states $|\psi\rangle \in \bigcup_{N \in \mathbb{N}} R_N$, we have $R^\star(\psi) = R^\star_{r^\star(\psi)}(\psi) > 0$.

For all $|\psi\rangle \in R_\infty$ and for all $k \in \overline{\mathbb{N}}^*$,

$$\begin{cases} R_k^\star(\psi) > 0 & \text{for } k \in \mathbb{N}, \quad (2.123a) \\ R_k^\star(\psi) = 0 & \text{for } k = \infty. \quad (2.123b) \end{cases}$$

In particular, states of infinite stellar rank are not robust: for all states $|\psi\rangle \in R_\infty$, we have $R^\star(\psi) = R^\star_\infty(\psi) = 0$.

Eqs. (2.122a), (2.122b) and (2.123a) are deduced from Theorem 2.5, and Eq. (2.123b) is deduced from Lemma 2.6.

This result implies that the robust states (i.e., $R^\star > 0$) are exactly the non-Gaussian states of finite stellar rank. When considering imperfect single-mode non-Gaussian state engineering, one may thus restrict to states of finite stellar rank, which by Theorem 2.2 are obtained uniquely by a finite number of single-photon additions to a Gaussian state. Alternatively, one may also describe such states using Theorem 2.4 as finite superpositions of equally displaced and squeezed number states, or equivalently as Gaussian-convertible to core states. Engineering of such states has recently been considered in [SMS19], by photon detection of Gaussian states.

Moreover, for $k \in \mathbb{N}^*$, the states that are robust with respect to states of stellar rank lower than k (i.e., $R_k^\star > 0$) thus are the states $|\psi\rangle$ such that $r^\star(\psi) \geq k$.

2.3.3 Computing the robustness

Importantly, the k-robustness is state-dependent and the following result gives a simple expression. Let us define, for all $n \in \mathbb{N}$,

$$\Pi_n = \sum_{m=0}^{n} |m\rangle\langle m| \tag{2.124}$$

the projector onto the subspace spanned by the Fock states $|0\rangle, \ldots, |n\rangle$.

Theorem 2.6. *Let $k \in \mathbb{N}^*$ and let $|\psi\rangle \in \mathscr{H}$. Then,*

$$R_k^\star(\psi) = \sqrt{1 - \sup_{\hat{G}\in\mathscr{G}} \mathrm{Tr}\left[\Pi_{k-1}\hat{G}|\psi\rangle\langle\psi|\hat{G}^\dagger\right]}, \tag{2.125}$$

where the supremum is over Gaussian unitary operations. Moreover, assuming the optimisation yields a Gaussian operation $\hat{G}_0$, an optimal approximating state is

$$\hat{G}_0^\dagger\left(\frac{\Pi_{k-1}\hat{G}_0|\psi\rangle}{\left\|\Pi_{k-1}\hat{G}_0|\psi\rangle\right\|}\right). \tag{2.126}$$

Proof. From Lemma 2.4 and in particular Eq. (2.88) we have

$$R_k^\star(\psi) = \sqrt{1 - \sup_{r^\star(\phi)<k} |\langle\phi|\psi\rangle|^2}. \tag{2.127}$$

By Theorem 2.4, for any pure state $|\phi\rangle$ such that $r^\star(\phi) < k$, there exist a normalised core state $|C_\phi\rangle$ of stellar rank lower than k and a Gaussian operation $\hat{G}_\phi$ such that

$$|\phi\rangle = \hat{G}_\phi|C_\phi\rangle. \tag{2.128}$$

We obtain

$$\begin{aligned} |\langle\phi|\psi\rangle|^2 &= |\langle C_\phi|\hat{G}_\phi^\dagger|\psi\rangle|^2 \\ &= |\langle C_\phi|\Pi_{k-1}\hat{G}_\phi^\dagger|\psi\rangle|^2 \\ &\leq |\langle C_\phi|C_\phi\rangle|^2 |\langle\psi|\hat{G}_\phi\Pi_{k-1}\hat{G}_\phi^\dagger|\psi\rangle|^2 \\ &= \mathrm{Tr}\left[\Pi_{k-1}\hat{G}_\phi^\dagger|\psi\rangle\langle\psi|\hat{G}_\phi\right], \end{aligned} \tag{2.129}$$

where we used $|C_\phi\rangle = \Pi_{k-1}|C_\phi\rangle$ in the second line, since $|C_\phi\rangle$ is a core state of stellar rank lower than k (hence its support is contained in the support of Π_{k-1}), Cauchy-Schwarz inequality in the third line and $|\langle C_\phi|C_\phi\rangle|^2 = 1$ in the last line. This upperbound is attained if

$$|C_\phi\rangle = \frac{\Pi_{k-1}\hat{G}_\phi^\dagger|\psi\rangle}{\sqrt{\mathrm{Tr}\left[\Pi_{k-1}\hat{G}_\phi^\dagger|\psi\rangle\langle\psi|\hat{G}_\phi\right]}}, \tag{2.130}$$

which is indeed a normalised core state of stellar rank lower than k.

With Eqs. (2.127) and (2.129), the robustness of the state $|\psi\rangle$ is then given by

$$\begin{aligned} R_k^\star(\psi) &= \sqrt{1 - \sup_{\hat{G}_\phi \in \mathcal{G}} \mathrm{Tr}\left[\Pi_{k-1}\hat{G}_\phi^\dagger |\psi\rangle\langle\psi| \hat{G}_\phi\right]} \\ &= \sqrt{1 - \sup_{\hat{G} \in \mathcal{G}} \mathrm{Tr}\left[\Pi_{k-1}\hat{G} |\psi\rangle\langle\psi| \hat{G}^\dagger\right]}, \end{aligned} \tag{2.131}$$

where the supremum is over Gaussian unitary operations and where we used the fact that the set of Gaussian unitary operations is invariant under adjoint in the second line. With Eq. (2.128), assuming the optimisation yields an optimal Gaussian unitary $\hat{G}_0$, an optimal approximating state is $|\phi\rangle = \hat{G}_\phi |C_\phi\rangle$, where $\hat{G}_\phi = \hat{G}_0^\dagger$ and where $|C_\phi\rangle$ is given by Eq. (2.130). Namely,

$$|\phi\rangle = \hat{G}_0^\dagger \left(\frac{\Pi_{k-1}\hat{G}_0 |\psi\rangle}{\left\| \Pi_{k-1}\hat{G}_0 |\psi\rangle \right\|} \right), \tag{2.132}$$

i.e., $\hat{G}_0 |\phi\rangle$ is the renormalised truncation of $\hat{G}_0 |\psi\rangle$ at photon number $k-1$.

■

From Theorem 2.6, the robustness profile $\mathscr{R}(\psi) = \left(R_k^\star\right)_{k\in\mathbb{N}^*}$ is a non-increasing sequence for any state $|\psi\rangle$, and each term in the sequence may be obtained with an optimisation over two complex parameters.

In particular, with the Hermite polynomials

$$\begin{aligned} He_m(z) &= (-1)^m e^{\frac{1}{2}z^2} \partial_z^m e^{-\frac{1}{2}z^2} \\ &= \sum_{p=0}^{\lfloor \frac{m}{2} \rfloor} \frac{m!(-1)^p}{2^p p!(m-2p)!} z^{m-2p}, \end{aligned} \tag{2.133}$$

for all $m \in \mathbb{N}$ and all $z \in \mathbb{C}$, the robustness of cat states has the following expression:

Corollary 2.2. *Let $k \in \mathbb{N}^*$ and let $\alpha \in \mathbb{C}$. Then, writing $c_x = \cosh x$, $s_x = \sinh x$, and $t_x = \tanh x$ for brevity,*

$$R_k^\star(cat_\alpha^+) = \sqrt{1 - \sup_{\xi = re^{i\theta}, \beta \in \mathbb{C}} \frac{e^{-|\beta|^2}}{4c_r c_{|\alpha|^2}} \sum_{m=0}^{k-1} \frac{t_r^m}{m!} \left| u_m^+(|\alpha|, \xi, \beta) \right|^2}, \tag{2.134}$$

and

$$R_k^\star(cat_\alpha^-) = \sqrt{1 - \sup_{\xi = re^{i\theta}, \beta \in \mathbb{C}} \frac{e^{-|\beta|^2}}{4c_r s_{|\alpha|^2}} \sum_{m=0}^{k-1} \frac{t_r^m}{m!} \left| u_m^-(|\alpha|, \xi, \beta) \right|^2}, \tag{2.135}$$

where

$$u_m^\pm(|\alpha|, \xi, \beta) := e^{-|\alpha|\beta^* + \frac{1}{2}t_r e^{i\theta}(|\alpha|+\beta)^2} He_m\left(\frac{|\alpha|+\beta}{\sqrt{c_r s_r}} e^{i\theta/2} \right) \pm e^{|\alpha|\beta^* + \frac{1}{2}t_r e^{i\theta}(\beta-|\alpha|)^2} He_m\left(\frac{\beta-|\alpha|}{\sqrt{c_r s_r}} e^{i\theta/2} \right). \tag{2.136}$$

Proof. Let $\alpha \in \mathbb{C}$. We have

$$|\text{cat}_\alpha^\pm\rangle = \frac{1}{\sqrt{\mathcal{N}_\alpha^\pm}}(|\alpha\rangle \pm |-\alpha\rangle), \tag{2.137}$$

where $\mathcal{N}_\alpha^\pm = 2(1 \pm e^{-2|\alpha|^2})$. By Theorem 2.6,

$$R_k^\star(\text{cat}_\alpha^\pm) = \sqrt{1 - \sup_{\hat{G}\in\mathcal{G}} \text{Tr}\left[\Pi_{k-1}\hat{G}\,|\text{cat}_\alpha^\pm\rangle\langle\text{cat}_\alpha^\pm|\,\hat{G}^\dagger\right]}. \tag{2.138}$$

for all $k \in \mathbb{N}^*$, where Π_{k-1} in the projector onto the Fock basis with less the $k-1$ photons. We have

$$\text{Tr}\left[\Pi_{k-1}\hat{G}\,|\text{cat}_\alpha^\pm\rangle\langle\text{cat}_\alpha^\pm|\,\hat{G}^\dagger\right] = \sum_{m=0}^{k-1} \left|\langle m|\hat{G}|\text{cat}_\alpha^\pm\rangle\right|^2. \tag{2.139}$$

Setting $\hat{G} = \hat{S}(\xi)\hat{D}(\beta)$, for $\xi = re^{i\theta}, \beta \in \mathbb{C}$, we obtain

$$\begin{aligned}\langle m|\hat{G}|\text{cat}_\alpha^\pm\rangle &= \frac{1}{\sqrt{\mathcal{N}_\alpha^\pm}}\left(\langle m|\hat{S}(\xi)\hat{D}(\beta)|\alpha\rangle \pm \langle m|\hat{S}(\xi)\hat{D}(\beta)|-\alpha\rangle\right)\\ &= \frac{1}{\sqrt{\mathcal{N}_\alpha^\pm m!}}\left(e^{\frac{1}{2}(\alpha^*\beta-\alpha\beta^*)}\langle 0|\hat{a}^m\hat{S}(\xi)\hat{D}(\alpha+\beta)|0\rangle \pm e^{\frac{1}{2}(\alpha\beta^*-\alpha^*\beta)}\langle 0|\hat{a}^m\hat{S}(\xi)\hat{D}(\beta-\alpha)|0\rangle\right).\end{aligned} \tag{2.140}$$

Switching to the stellar representation we obtain

$$\begin{aligned}\langle m|\hat{G}|\text{cat}_\alpha^\pm\rangle &= \frac{1}{\sqrt{\mathcal{N}_\alpha^\pm m!}}\left[\partial_z^m e^{\frac{1}{2}(\alpha^*\beta-\alpha\beta^*)}G^\star_{\xi,\alpha+\beta}(z) \pm \partial_z^m e^{\frac{1}{2}(\alpha\beta^*-\alpha^*\beta)}G^\star_{\xi,-\alpha+\beta}(z)\right]_{z=0}\\ &= \frac{e^{-\frac{1}{2}(|\alpha|^2+|\beta|^2)}}{\sqrt{c_r\mathcal{N}_\alpha^\pm m!}}\Bigg[e^{-\alpha\beta^*+\frac{1}{2}t_re^{i\theta}(\alpha+\beta)^2}\partial_z^m e^{-\frac{1}{2}e^{-i\theta}t_rz^2+\frac{\alpha+\beta}{c_r}z}\\ &\qquad \pm e^{\alpha\beta^*+\frac{1}{2}t_re^{i\theta}(\beta-\alpha)^2}\partial_z^m e^{-\frac{1}{2}e^{-i\theta}t_rz^2+\frac{\beta-\alpha}{c_r}z}\Bigg]_{z=0},\end{aligned} \tag{2.141}$$

where we used $\langle 0|\psi\rangle = F_\psi^\star(0)$ and Eq. (2.7). With Eq. (2.133) we have

$$\left[\partial_z^m e^{-\frac{1}{2}az^2+bz}\right]_{z=0} = a^{m/2}He_m\left(\frac{b}{\sqrt{a}}\right), \tag{2.142}$$

where He_m is the m^{th} Hermite polynomial.
With Eq. (2.141) we obtain

$$\begin{aligned}\left|\langle m|\hat{G}|\text{cat}_\alpha^\pm\rangle\right|^2 = \frac{e^{-(|\alpha|^2+|\beta|^2)}t_r^m}{c_r\mathcal{N}_\alpha^\pm m!}\Bigg|&e^{-\alpha\beta^*+\frac{1}{2}t_re^{i\theta}(\alpha+\beta)^2}He_m\left(\frac{\alpha+\beta}{\sqrt{c_rs_r}}e^{i\theta/2}\right)\\ &\pm e^{\alpha\beta^*+\frac{1}{2}t_re^{i\theta}(\beta-\alpha)^2}He_m\left(\frac{\beta-\alpha}{\sqrt{c_rs_r}}e^{i\theta/2}\right)\Bigg|^2.\end{aligned} \tag{2.143}$$

Combining Eqs. (2.138), (2.139) and (2.143) yields

$$R_k^\star(\mathrm{cat}_\alpha^\pm) = \sqrt{1 - \sup_{\xi=re^{i\theta},\beta\in\mathbb{C}} \frac{e^{-(|\alpha|^2+|\beta|^2)}}{c_r \mathcal{N}_\alpha^\pm} \sum_{m=0}^{k-1} \frac{t_r^m}{m!} \left|u_m^\pm(\alpha,\xi,\beta)\right|^2}, \tag{2.144}$$

where we have set

$$u_m^\pm(\alpha,\xi,\beta) := e^{-\alpha\beta^* + \frac{1}{2}t_r e^{i\theta}(\alpha+\beta)^2} He_m\left(\frac{\alpha+\beta}{\sqrt{c_r s_r}} e^{i\theta/2}\right) \pm e^{\alpha\beta^* + \frac{1}{2}t_r e^{i\theta}(\beta-\alpha)^2} He_m\left(\frac{\beta-\alpha}{\sqrt{c_r s_r}} e^{i\theta/2}\right). \tag{2.145}$$

Since the robustness is invariant under Gaussian operations, $R_k^\star(\mathrm{cat}_\alpha^\pm)$ does not depend on the phase of α, since one can map a cat state of amplitude α to a cat state of amplitude $e^{i\phi}\alpha$ through a Gaussian rotation (this corresponds to mapping β to $e^{i\phi}\beta$ and θ to $\theta - 2\phi$ in the previous expressions). Hence, we may assume without loss of generality that $\alpha \in \mathbb{R}$ and replace α by $|\alpha|$. With

$$c_{|\alpha|^2} = \frac{\mathcal{N}_\alpha^+}{4e^{-|\alpha|^2}} \quad \text{and} \quad s_{|\alpha|^2} = \frac{\mathcal{N}_\alpha^-}{4e^{-|\alpha|^2}}, \tag{2.146}$$

this concludes the proof.

■

By Lemma 2.4, the sequence of maximum achievable fidelities for each rank $k-1$ with a given target state is obtained from the robustness profile with $F = 1 - R_k^{\star 2}$. We have computed numerically the values of $R_k^\star(\mathrm{cat}_\alpha^+)$ and $R_k^\star(\mathrm{cat}_\alpha^-)$ for different values of k and α and the corresponding achievable fidelities are depicted in Fig. 2.4 and Fig. 2.5. For each rank, if ρ denotes a state for which the maximum fidelity is achieved, then any lower fidelity may be obtain by considering the states $\rho_p = p\,|\perp\rangle\langle\perp| + (1-p)\rho$, for $0 \le p \le 1$, where $|\perp\rangle$ is a coherent state orthogonal to the target state (which exists by Theorem 2.1, since the target state is non-Gaussian). From the numerics and the obtained profiles of the cat states, we make various observations:

- The main difference between low amplitude cat^+ and cat^- states is that the former are easier to approximate by Gaussian states than the latter: at low amplitude, cat^+ states are closer to the vacuum while cat^- states are closer to the single photon Fock state.

- High amplitude cat states are 'more non-Gaussian' than low amplitude cat states, in the sense that one needs more photon additions to approximate them to the same precision.

- The maximum achievable fidelity increases more from odd to even ranks (resp. even to odd ranks) than from even to odd ranks (resp. odd to even ranks) for cat^+ states (resp. cat^- states). This is due to cat^+ states (resp. cat^- states) having support only on even (resp. odd) Fock states.

- For each given amplitude, there is a critical stellar rank after which good approximation of the cat state becomes possible. Before that stellar rank, the best Gaussian operation in the

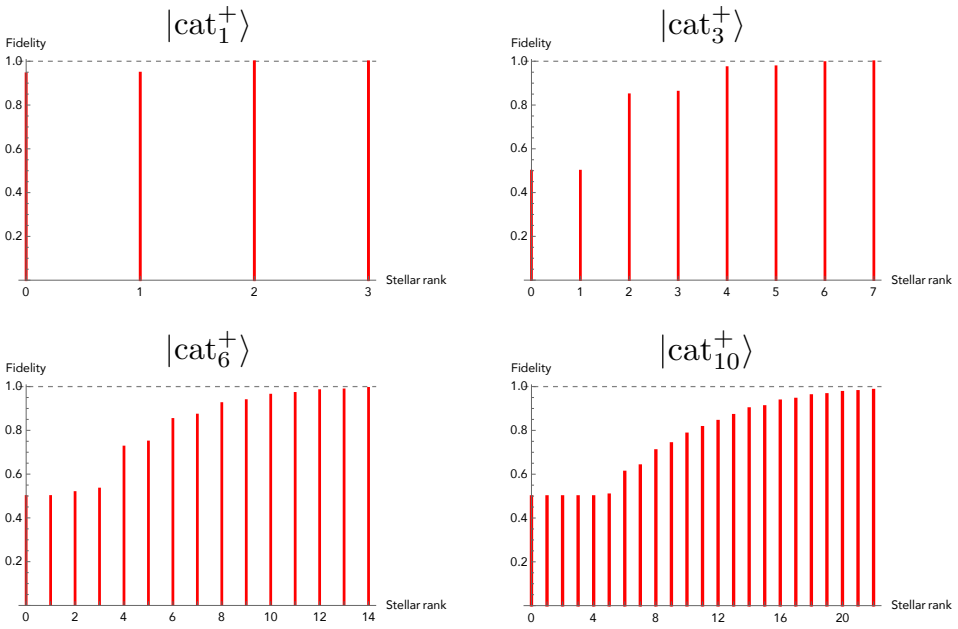

Figure 2.4: Achievable fidelities for target cat^+ states of amplitudes 1, 3, 6 and 10. For each rank $k \in \mathbb{N}^*$, the vertical line depicts the achievable fidelities between the target state and states of rank k.

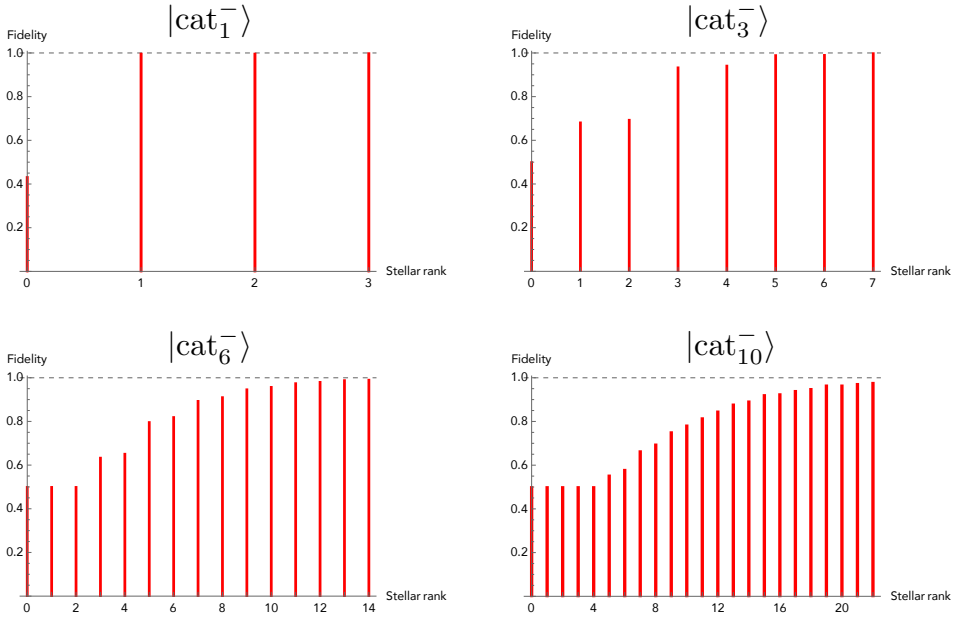

Figure 2.5: Achievable fidelities for target cat^- states of amplitudes 1, 3, 6 and 10. For each rank $k \in \mathbb{N}^*$, the vertical line depicts the achievable fidelities between the target state and states of rank k.

optimisation of Corollary 2.2 is roughly a displacement of the amplitude of the cat. Past that rank, it is a smaller displacement combined with a squeezing.

If the state $|\psi\rangle$ is of finite rank, the k^{th} term of the sequence $\mathscr{R}(\psi)$ is zero for all $k > r^\star(\psi)$ by Corollary 2.1. For $k \leq r^\star(\psi)$, an expression depending on the core state of $|\psi\rangle$ may be obtained for $R_k^\star(\psi)$:

Corollary 2.3. *Let $k \in \mathbb{N}^*$ and let $|\psi\rangle \in \mathscr{H}$ be a non-Gaussian pure state of finite stellar rank $r^\star(\psi) \geq k$, with core state $|C_\psi\rangle = \sum_{n=0}^{r_\psi^\star} C_n |n\rangle$. Then,*

$$R_k^\star(\psi) = \sqrt{1 - \sup_{\xi,\alpha\in\mathbb{C}} \sum_{m=0}^{k-1} |u_m(\xi,\alpha)|^2}, \tag{2.147}$$

where for all $m \in \{0,\dots,k-1\}$ and all $\xi = re^{i\theta}, \alpha \in \mathbb{C}$,

$$u_m(\xi,\alpha) = \frac{1}{\sqrt{m! c_r}} \sum_{n=0}^{r_\psi^\star} \frac{C_n^*}{\sqrt{n!}} \left[\partial_z^n (c_r z + s_r e^{i\theta}\partial_z - \alpha^*)^m e^{-\frac{1}{2}e^{-i\theta}t_r z^2 + \frac{\alpha}{c_r}z + \frac{1}{2}e^{i\theta}t_r\alpha^2 - \frac{1}{2}|\alpha|^2}\right]_{z=0}, \tag{2.148}$$

with $c_r = \cosh r$, $s_r = \sinh r$ and $t_r = \tanh r$. Moreover, assuming the optimisation yields values $\xi_0, \alpha_0 \in \mathbb{C}$, an optimal approximating state is

$$\hat{D}^\dagger(\alpha_0)\hat{S}^\dagger(\xi_0)\left(\frac{\Pi_{k-1}\hat{S}(\xi_0)\hat{D}(\alpha_0)|C_\psi\rangle}{\left\|\Pi_{k-1}\hat{S}(\xi_0)\hat{D}(\alpha_0)|C_\psi\rangle\right\|}\right). \tag{2.149}$$

Proof. Let $|\psi\rangle \in \mathscr{H}$ be a pure state of finite stellar rank $r^\star(\psi) \in \mathbb{N}^*$ with core state $|C_\psi\rangle = \sum_{n=0}^{r_\psi^\star} C_n |n\rangle$. By Theorem 2.4, there exist a Gaussian operation $\hat{G}_\psi$ such that $|\psi\rangle = \hat{G}_\psi |C_\psi\rangle$. From Theorem 2.6 we have

$$\begin{aligned} R_k^\star(\psi) &= \sqrt{1 - \sup_{\hat{G}\in\mathscr{G}} \mathrm{Tr}\left[\Pi_{k-1}\hat{G}|\psi\rangle\langle\psi|\hat{G}^\dagger\right]} \\ &= \sqrt{1 - \sup_{\hat{G}'\in\mathscr{G}} \mathrm{Tr}\left[\Pi_{k-1}\hat{G}'|C_\psi\rangle\langle C_\psi|\hat{G}'^\dagger\right]} \\ &= \sqrt{1 - \sup_{\xi,\alpha\in\mathbb{C}} \mathrm{Tr}\left[\Pi_{k-1}\hat{D}^\dagger(\alpha)\hat{S}^\dagger(\xi)|C_\psi\rangle\langle C_\psi|\hat{S}(\xi)\hat{D}(\alpha)\right]} \\ &= \sqrt{1 - \sup_{\xi,\alpha\in\mathbb{C}} \sum_{m=0}^{k-1} \left|\langle C_\psi|\hat{S}(\xi)\hat{D}(\alpha)|m\rangle\right|^2}, \end{aligned} \tag{2.150}$$

where we used the group structure of the Gaussian unitary operations in the second line and the fact that any single-mode Gaussian unitary operation may be decomposed as a squeezing and a displacement in the third line. Assuming the optimisation yields values $\xi_0, \alpha_0 \in \mathbb{C}$, the optimal core state used in the approximation is

$$|C\rangle = \frac{\Pi_{k-1}\hat{S}(\xi_0)\hat{D}(\alpha_0)|C_\psi\rangle}{\left\|\Pi_{k-1}\hat{S}(\xi_0)\hat{D}(\alpha_0)|C_\psi\rangle\right\|}. \tag{2.151}$$

Now for all $m \in \{0,\dots,k-1\}$,

$$\langle C_\psi|\hat{S}(\xi)\hat{D}(\alpha)|m\rangle = \sum_{n=0}^{r_\psi^\star} C_n^* \langle n|\hat{S}(\xi)\hat{D}(\alpha)|m\rangle, \tag{2.152}$$

and for all $n \in \{0,\dots,r_\psi^\star\}$,

$$\begin{aligned}\langle n|\hat{S}(\xi)\hat{D}(\alpha)|m\rangle &= \frac{1}{\sqrt{m!n!}}\langle 0|\hat{a}^n\hat{S}(\xi)\hat{D}(\alpha)(\hat{a}^\dagger)^m|0\rangle \\ &= \frac{1}{\sqrt{m!n!}}\langle 0|\hat{a}^n(c_r\hat{a}^\dagger + s_r e^{i\theta}\hat{a} - \alpha^*)^m\hat{S}(\xi)\hat{D}(\alpha)|0\rangle\,,\end{aligned} \tag{2.153}$$

where we used Eq. (2.37) in the second line, with $c_r = \cosh r$, $s_r = \sinh r$. Hereafter we also set $t_r = \tanh r$. We have $\langle 0|\chi\rangle = F_\chi^\star(0)$ for all states χ, hence switching to the stellar representation Eq. (2.153) rewrites

$$\langle n|\hat{S}(\xi)\hat{D}(\alpha)|m\rangle = \frac{1}{\sqrt{m!n!c_r}}\left[\partial_z^n\left(c_r z + s_r e^{i\theta}\partial_z - \alpha^*\right)^m e^{-\frac{1}{2}e^{-i\theta}t_r z^2 + \frac{\alpha}{c_r}z + \frac{1}{2}e^{i\theta}t_r\alpha^2 - \frac{1}{2}|\alpha|^2}\right]_{z=0}, \tag{2.154}$$

where we used Eq. (2.7). Hence,

$$\langle C_\psi|\hat{S}(\xi)\hat{D}(\alpha)|m\rangle = \sum_{n=0}^{r_\psi^\star}\frac{C_n^*}{\sqrt{m!n!c_r}}\left[\partial_z^n\left(c_r z + s_r e^{i\theta}\partial_z - \alpha^*\right)^m e^{-\frac{1}{2}e^{-i\theta}t_r z^2 + \frac{\alpha}{c_r}z + \frac{1}{2}e^{i\theta}t_r\alpha^2 - \frac{1}{2}|\alpha|^2}\right]_{z=0}. \tag{2.155}$$

Setting, for all $m \in \{0,\dots,k-1\}$, $u_m(\xi,\alpha) = \langle C_\psi|\hat{S}(\xi)\hat{D}(\alpha)|m\rangle$, thus omitting the dependency in ψ, we finally obtain with Eq. (2.150),

$$R_k^\star(\psi) = \sqrt{1 - \sup_{\xi,\alpha\in\mathbb{C}}\sum_{m=0}^{k-1}|u_m(\xi,\alpha)|^2}, \tag{2.156}$$

where for all $m \in \{0,\dots,k-1\}$ and all $\xi = re^{i\theta}, \alpha \in \mathbb{C}$,

$$u_m(\xi,\alpha) = \frac{1}{\sqrt{m!c_r}}\sum_{n=0}^{r_\psi^\star}\frac{C_n^*}{\sqrt{n!}}\left[\partial_z^n(c_r z + s_r e^{i\theta}\partial_z - \alpha^*)^m e^{-\frac{1}{2}e^{-i\theta}t_r z^2 + \frac{\alpha}{c_r}z + \frac{1}{2}e^{i\theta}t_r\alpha^2 - \frac{1}{2}|\alpha|^2}\right]_{z=0}, \tag{2.157}$$

with $c_r = \cosh r$, $s_r = \sinh r$ and $t_r = \tanh r$. Moreover, assuming the optimisation yields values $\xi_0, \alpha_0 \in \mathbb{C}$, an optimal approximating state is $|\phi\rangle = \hat{D}^\dagger(\alpha_0)\hat{S}^\dagger(\xi_0)|C\rangle$, where $|C\rangle$ is defined in Eq. (2.151). Namely,

$$|\phi\rangle = \hat{D}^\dagger(\alpha_0)\hat{S}^\dagger(\xi_0)\left(\frac{\Pi_{k-1}\hat{S}(\xi_0)\hat{D}(\alpha_0)|C_\psi\rangle}{\left\|\Pi_{k-1}\hat{S}(\xi_0)\hat{D}(\alpha_0)|C_\psi\rangle\right\|}\right), \tag{2.158}$$

i.e., $\hat{S}(\xi_0)\hat{D}(\alpha_0)|\phi\rangle$ is the renormalised truncation of $\hat{S}(\xi_0)\hat{D}(\alpha_0)|C_\psi\rangle$ at photon number $k-1$.

■

From this result, the value of the stellar robustness may be obtained analytically for low stellar rank states and numerically for all finite stellar rank states. In particular, we obtain:

Lemma 2.7. *For the single photon Fock state* $|1\rangle$ *of stellar rank* 1 *we have*

$$R^\star(1) = \sqrt{1 - \frac{3\sqrt{3}}{4e}}. \tag{2.159}$$

The corresponding maximum achievable fidelity is given by

$$\frac{3\sqrt{3}}{4e} \approx 0.478. \tag{2.160}$$

Proof. The single-photon Fock state $|1\rangle$ is a core state of stellar rank 1 (its stellar function is given by $F_1^\star(z) = z$ for all $z \in \mathbb{C}$). Hence, by Corollary 2.3,

$$\begin{aligned} R_1^\star(1) &= R^\star(1) \\ &= \sqrt{1 - \sup_{\xi = re^{i\theta}, \alpha \in \mathbb{C}} \frac{1}{c_r} \left| \left[\partial_z e^{-\frac{t_r}{2} e^{-i\theta} z^2 + \frac{\alpha}{c_r} z + \frac{t_r}{2} e^{i\theta} \alpha^2 - \frac{1}{2}|\alpha|^2} \right]_{z=0} \right|^2} \\ &= \sqrt{1 - \sup_{\xi = re^{i\theta}, \alpha \in \mathbb{C}} \frac{|\alpha|^2}{c_r^3} \left| e^{\frac{t_r}{2} e^{i\theta} \alpha^2 - \frac{1}{2}|\alpha|^2} \right|^2} \\ &= \sqrt{1 - \sup_{\xi = re^{i\theta}, \alpha \in \mathbb{C}} \frac{|\alpha|^2}{c_r^3} e^{\frac{t_r}{2}(\alpha^2 e^{i\theta} + \alpha^{*2} e^{-i\theta}) - |\alpha|^2}}. \end{aligned} \tag{2.161}$$

Setting $\gamma = x + iy = i\alpha e^{i\theta/2}$ we obtain

$$\begin{aligned} \frac{|\alpha|^2}{c_r^3} e^{\frac{t_r}{2}(\alpha^2 e^{i\theta} + \alpha^{*2} e^{-i\theta}) - |\alpha|^2} &= \frac{|\gamma|^2}{c_r^3} e^{-|\gamma|^2 - \frac{t_r}{2}(\gamma^2 + \gamma^{*2})} \\ &= \frac{x^2 + y^2}{c_r^3} e^{-(1+t_r)x^2} e^{-(1-t_r)y^2} \\ &= (1 - t_r^2)^{3/2} (x^2 + y^2) e^{-(1+t_r)x^2} e^{-(1-t_r)y^2} \\ &= (1 - t_r^2)^{3/2} (x^2 + y^2) e^{-(1-t_r)(x^2+y^2)} e^{-2t_r x^2} \\ &\leq (1 - t_r^2)^{3/2} (x^2 + y^2) e^{-(1-t_r)(x^2+y^2)} \\ &\leq \frac{(1 - t_r^2)^{3/2}}{e(1 - t_r)} \\ &= \frac{1}{e} \sqrt{(1 - t_r)(1 + t_r)^3}, \end{aligned} \tag{2.162}$$

and this upperbound is attained for $x = 0$ and $y = \frac{1}{\sqrt{1-t_r}}$. Finally, we have $\max_{u \in [0,1]} (1-u)(1+u)^3 =$

$\frac{27}{16}$, attained for $u = \frac{1}{2}$, so we obtain the stellar robustness of a single photon Fock state:

$$R^\star(1) = \sqrt{1 - \frac{3\sqrt{3}}{4e}}. \tag{2.163}$$

■

Since the stellar robustness inherits the property of invariance under Gaussian unitary operations of the stellar rank, Corollary 2.1 implies the same robustness value for states obtained from a single photon Fock state by unitary Gaussian operations, such as photon-added or photon-subtracted squeezed states, by Eq. (2.81).

We have computed numerically the stellar robustness for the states $\cos\phi|0\rangle + e^{i\chi}\sin\phi|1\rangle$, for all $\phi, \chi \in [0, 2\pi]$, which is independent of χ (Fig. 2.6). Setting $\phi = \frac{\pi}{2}$ yields the single photon Fock state, which is thus the most robust state of stellar rank 1, up to Gaussian unitary operations.

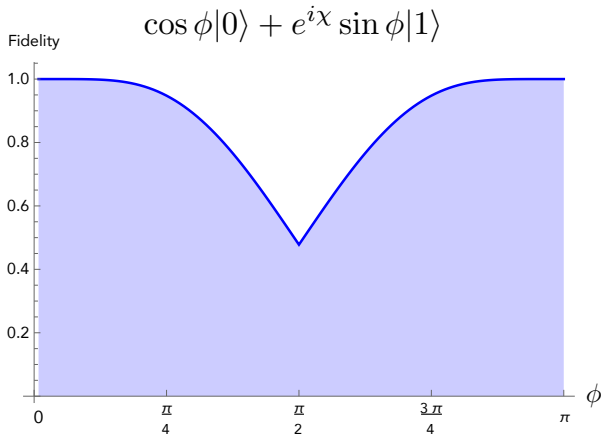

Figure 2.6: Achievable fidelities with Gaussian states for target core states $\cos\phi|0\rangle + e^{i\chi}\sin\phi|1\rangle$, for all $\phi, \chi \in [0, 2\pi]$, as a function of ϕ. The maximum fidelities are independent of χ, and yield the stellar robustness through Lemma 2.4.

We have also obtained numerically the achievable fidelities with target states $|2\rangle$, $|3\rangle$, $|4\rangle$ and $|5\rangle$, depicted in Fig. 2.7.

As previously mentioned, with Lemma 2.4, preparing a target pure state of finite stellar rank $|\psi\rangle$ with fidelity better than $1 - [R_k^\star(\psi)]^2$, which may be computed using Corollary 2.3, ensures that the obtained state has a stellar rank equal to or greater than k.

One may obtain the ϵ-smoothed non-Gaussianity of formation of a given target pure state—i.e., the minimal stellar rank of ϵ-close states—from its profile of achievable fidelities as follows: for a given $\epsilon > 0$, it is the x-coordinate of the leftmost intersection point of the horizontal line of height $1 - \epsilon^2$ with the vertical lines of the profile. For example, let $\epsilon = 0.7$, so that $1 - \epsilon^2 = 0.51$. With

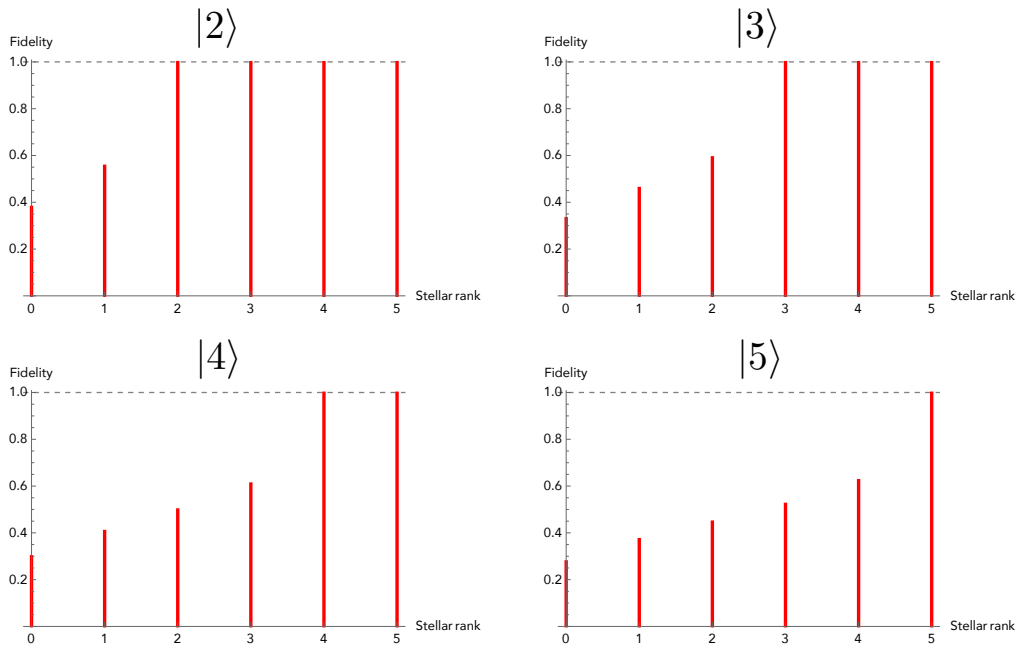

Figure 2.7: Achievable fidelities for target Fock states $|2\rangle$, $|3\rangle$, $|4\rangle$ and $|5\rangle$. For each rank $k \in \mathbb{N}^*$, the vertical line depicts the achievable fidelities between the target state and states of rank k. The sequence of maximum fidelities for each rank yields the robustness profile through Lemma 2.4.

Fig. 2.7, the 0.7-smoothed non-Gaussianity of formation of the Fock state $|3\rangle$ is equal to 2. Hence, an experimental (mixed) quantum state which has fidelity greater than 0.51 (corresponding to $\epsilon = 0.7$) with the Fock state $|3\rangle$ has a stellar rank greater or equal to 2.

The stellar hierarchy may thus be certified experimentally by direct fidelity estimation with a target pure state. In chapter 4, we make use of this result and discuss in particular the certification of the stellar rank using Gaussian measurements and heterodyne detection.

2.4 Discussion and open problems

Based on the stellar representation of single-mode continuous variable quantum states, we have defined the stellar rank as the number of zeros of the stellar function, or equivalently of the Husimi Q function. Using the analytic properties of the stellar function, we have shown that this rank is invariant under Gaussian operations and induces a hierarchy over the space of single-mode normalised states. We have characterized the states of finite stellar rank as the states obtained by successive single-photon additions to a Gaussian state, or equivalently as finite superpositions of (equally) displaced and squeezed number states. Additionally, we have given the stellar rank an operational meaning, as the minimal non-Gaussian cost for engineering a state, in terms of single-photon additions and subtractions. We have derived the equivalence

classes for Gaussian convertibility using the notion of core states, and we have studied in detail the robustness of the ranks of the stellar hierarchy, showing that finite stellar rank states are robust, while infinite stellar rank states are not. In particular, we have shown how to compute the robustness.

While the stellar representation unveils the structure of single-mode non-Gaussian states, various open questions remain:

How to extend the stellar formalism to the case of multimode states? The stellar function for multimode states is a multivariate analytic function, which prevents the use of the factorisation theorem, crucial in the derivation of the results. However, one can consider in the multimode case both the set of states that can be obtained from the vacuum by multimode Gaussian operations and a finite number of photon additions and the set of states that are obtained by multimode Gaussian operations acting on an input with a multivariate polynomial stellar function. We consider both sets of states in the next chapter and investigates their computational power.

Can we interpret geometrically the non-Gaussian properties of quantum states? Can we view the stellar representation as a limit of the Majorana representation? This representation provides a beautiful interpretation of non-entangling operations for symmetric states as a class of transformations of the sphere [RM11, Aul11]. Can we derive an analogous statement to characterize how Gaussian operations affect the roots of the stellar function? The displacement operator simply displaces the sphere on the complex plane, but the action of the squeezing seems nontrivial.

The stellar rank has an interpretation as a non-Gaussianity of formation, i.e., as a cost for quantum state engineering in terms of elementary non-Gaussian operations. Can we identify a computational task for which this rank quantifies how resourceful a state is?

Even though they are not robust, infinite stellar rank states are interesting from a conceptual point of vue, given their use for for error-correction [CMM99, GKP01]. Can we classify these states, for example based on the density of the zeros of their stellar function? To that end, what is the precise location of the zeros for GKP states?

CHAPTER 3

BEYOND-CLASSICAL QUANTUM CONTINUOUS VARIABLE MODELS

Different approaches are possible to probe the quantum computational advantage regime and to study the boundary between the quantum systems which are efficiently simulable classically and those universal for quantum computing. On the one hand, the regime of classical simulability can be explored [Fey82]: being able to efficiently simulate a quantum computational model with a classical computer up to a certain regime may suggest a quantum advantage beyond this regime [Val02, TD02, BSBN02]. On the other hand, subuniversal models of quantum computing can be defined: these models lie somewhere in-between classical and universal quantum computing, in the sense that, although not possessing the full computational power of a universal quantum computer, they may outperform classical computational capabilities with respect to specific problems [BJS10, AA13, MFF14, BMS16, FH16, DMK+17, BIS+18]. In both cases, one can aim for minimal extensions beyond the classically simulable models which are more likely to be implementable in the near term than universal quantum computers. For concrete applications, however, it is not the quantum computational model but rather the task at hand whose classical simulability matters. Studying application-specific classical simulation regimes is therefore also of great importance.

We investigate these approaches for different continuous variable quantum computational models. After introducing classical simulation notions, we consider linear optics with input single photons and adaptive photon-number measurement, and study the classial simulation regime for probability estimation and overlap estimation, two computational tasks that are central to machine learning applications [BGM19]. Next, we turn to Gaussian quantum circuits with non-Gaussian input states and derive sufficient conditions for an efficient classical strong simulation. Finally, we focus on a specific subclass of Gaussian quantum circuits with non-Gaussian input states, the CVS circuits, which relates to Boson Sampling with continuous variable measurements.

We identify the regime for which an efficient classical weak simulation of circuits in this subclass would imply a collapse of the polynomial hierarchy of complexity classes.

This chapter is based on [CMS20, CMG20a, CDM+17].

3.1 Classical simulation of quantum computations

Depending on the approach used for simulating classically the functioning of quantum devices, several notions of simulability are commonly used. In what follows, we review the ones we will be considering in this chapter.

3.1.1 Strong simulation

To each quantum computation is associated a probability distribution from which classical outcomes are sampled. In the case of continuous variable quantum computations with continuous variable outcomes, the output probability distribution is replaced by an output probability density. This motivates the following (informal) definition [TD02, PBG20].

Definition 3.1 (Strong simulation)**.** A quantum computation is *strongly simulable* if there exists a classical algorithm which evaluates its output probability distribution (density) or any of its marginals for any outcome in time polynomial in the size of the quantum computation.

Various relaxations of this definition are possible, allowing the classical evaluation to be approximate rather than exact, or to abort with a small probability. Hereafter we only consider the definition above. When there exists no efficient classical algorithm for strong simulation, we say that *strong simulation is hard*.

This notion of simulability is referred to as strong because it asks more from the classical simulation algorithm than from the quantum computation. Indeed, the quantum computation is merely sampling from a probability distribution (density), while the classical algorithm has to compute efficiently probabilities.

3.1.2 Weak simulation

A sampling counterpart to the notion of strong simulation is to ask the classical simulation algorithm to mimic the output of the quantum computation [TD02, PBG20]. Informally:

Definition 3.2 (Weak simulation)**.** A quantum computation is *weakly simulable* if there exists a classical algorithm which outputs samples from its output probability distribution (density) in time polynomial in the size of the quantum computation.

Akin to strong simulation, various relaxations of this definition are possible, allowing the classical sampling to be approximate rather than exact, or to abort with a small probability. Hereafter we

only consider the definition above. When there exists no efficient classical algorithm for weak simulation, we say that *weak simulation is hard*.

In the case of continuous variable quantum computations with continuous variable outcomes, a weaker requirement is to ask the classical simulation not to sample from the output probability density, but rather from a discretised probability distribution obtained from the probability density by performing an efficient binning of the sample space. Indeed, samples from the output probability density yield samples of such a discretised probability distribution with efficient classical post-processing.

Consider a quantum computation of size m yielding discrete classical outcomes from a probability distribution $P(X_1,\ldots,X_m)$, where X_i may take at most $M = \text{poly}\, m$ values for all $i \in \{1,\ldots,m\}$ (the sample space has size M^m). Then, weak simulation is weaker than strong simulation, with the following result [TD02, PBG20]:

Lemma 3.1. *An efficient classical algorithm for strong simulation provides an efficient classical algorithm for weak simulation (assuming one can efficiently sample from efficiently computable univariate probability distributions over a polynomial number of samples).*

We reproduce the proof below for completeness.

Proof. Assuming the existence of an efficient classical algorithm for strong simulation of a quantum computation of size m yielding classical outcomes from a discrete probability distribution $P(X_1,\ldots,X_m)$, where X_i may take at most $M = \text{poly}\, m$ values for all $i \in \{1,\ldots,m\}$, one first computes the marginal probabilities $P(X_1)$ for all M possible values of X_1. Then, one samples the value x_1 from $P(X_1)$ (which is an efficiently computable univariate probability distribution over a polynomial number of samples). With that sample x_1, one computes the conditional probability distribution

$$P(X_2|x_1) = \frac{P(x_1,X_2)}{P(x_1)}, \tag{3.1}$$

for all M possible values of X_2. Then, one samples the value x_2 from $P(X_2|x_1)$ (which is also an efficiently computable univariate probability distribution over a polynomial number of samples). Repeating the same procedure up to

$$P(X_m|x_1,\ldots,x_{m-1}) = \frac{P(x_1,\ldots,x_{m-1},X_m)}{P(x_1,\ldots,x_{m-1})}, \tag{3.2}$$

one obtains a sample $(x_1,\ldots,x_m)$ from $P(X_1,\ldots,X_m)$ efficiently.

■

For quantum computations yielding continuous variable classical outcomes, the result still holds with the same proof for binned discretised probability distributions rather than the corresponding

probability density, as long as the discretised probabilities can be computed efficiently from the probability density and have support on a polynomial number of bins for each mode.

3.1.3 Probability and overlap estimation

While the previous two notions of simulation of quantum computations are the most commonly used, other type of simulation may be useful: if the output samples of a quantum computation are used to compute a quantity which may be computed efficiently classically by other means, it is no longer necessary to simulate the whole quantum device. We consider two concrete examples which are prominent for variational quantum algorithms in quantum machine learning: probability estimation and overlap estimation [HCT+19, SK19].

Definition 3.3 (Probability estimation)**.** Let P be a probability distribution over m outcomes. Given any outcome $\boldsymbol{x}$ in the sample space of P, *probability estimation* refers to the computational task of outputting an estimate $\tilde{P}[\boldsymbol{x}]$ such that

$$P[\boldsymbol{x}] - \frac{1}{\text{poly}\, m} \leq \tilde{P}[\boldsymbol{x}] \leq P[\boldsymbol{x}] + \frac{1}{\text{poly}\, m}, \tag{3.3}$$

with probability greater than $1 - \frac{1}{\exp m}$.

Probability estimation amounts to outputting a polynomially precise additive estimate of the probability with exponentially small probability of failure. One may use the samples from a quantum computation in order to perform probability estimation for any given outcome: given a quantum device of size m which outputs samples from some probability distribution and a fixed outcome $\boldsymbol{x}$ in the sample space, one may run the device $M = \text{poly}\, m$ times, recording the value 1 whenever the outcome $\boldsymbol{x}$ is obtained and the value 0 otherwise. Then, summing and dividing by M, one obtains the frequency of the outcome $\boldsymbol{x}$ over the M uses of the quantum device, which is a polynomially precise additive estimate of the probability of the outcome $\boldsymbol{x}$ with exponentially small probability of failure, by virtue of Hoeffding inequality [Hoe63].

Weak simulation is at least as hard as probability estimation, since by the previous reasoning one may obtain polynomially precise additive estimates of probabilities from samples of the probability distribution. Moreover, they are some quantum computations for which weak simulation is hard (assuming widely believed conjectures from complexity theory), but probability estimation can be done efficiently classically. This is the case for IQP circuits [BJS10, HCT+19], Boson Sampling [AA13] and even the period-finding subroutine of Shor's algorithm [Sho94]. Let us detail the latter case: if N is an n bits integer to factor, the period-finding subroutine measures the output state

$$\frac{1}{N} \sum_x \sum_y e^{\frac{2i\pi xy}{N}} |y\rangle |f(x)\rangle \tag{3.4}$$

in the computational basis, where f is a periodic function over $\{0,\dots,N-1\}$ which can be evaluated efficiently. The probability of obtaining an outcome $y_0, f(x_0)$ is given by

$$\Pr[y_0, f(x_0)] = \left| \frac{1}{N} \sum_{f(x)=f(x_0)} e^{\frac{2i\pi x y_0}{N}} \right|^2 . \tag{3.5}$$

Now let

$$g_{x_0,y_0} : x \mapsto \begin{cases} e^{\frac{2i\pi x y_0}{N}} \text{ if } f(x) = f(x_0), \\ 0 \text{ otherwise.} \end{cases} \tag{3.6}$$

The function g_{x_0,y_0} can be evaluated efficiently and we have

$$\Pr[y_0, f(x_0)] = \left| \mathop{\mathbb{E}}_{x \leftarrow N} [g_{x_0,y_0}(x)] \right|^2 , \tag{3.7}$$

where $\mathop{\mathbb{E}}_{x \leftarrow N}$ denotes the expected value for x drawn uniformly randomly from $\{0,\dots,N-1\}$. By virtue of Hoeffding inequality, this quantity may be estimated efficiently (in n the number of bits of N) classically by sampling uniformly a polynomial number of values in $\{0,\dots,N-1\}$ and computing the modulus squared of the mean of g_{x_0,y_0} for these values.

However, note that probability estimation of quantum circuits is a BQP-complete computational task almost by definition, since given a polynomially precise estimate of the probability of acceptance of an input x to a quantum circuit, one may determine whether it is accepted or rejected by the circuit. In particular, unless factoring is in P, probability estimation for the quantum circuit corresponding to Shor's algorithm *as a whole* is hard and weak simulation of the period-finding subroutine is also hard, since in Shor's algorithm the output samples from the period-finding subroutine are used for a different classical computation than probability estimation (essentially obtaining promising candidates for the period).

A more general computational task than probability estimation in the context of quantum computing is the following:

Definition 3.4 (Overlap estimation)**.** Let $|\phi\rangle$ and $|\psi\rangle$ be quantum output states of two quantum computations of size m. *Overlap estimation* refers to the computational task of outputting an estimate $\tilde{O}$ such that

$$|\langle\phi|\psi\rangle|^2 - \frac{1}{\text{poly } m} \leq \tilde{O} \leq |\langle\phi|\psi\rangle|^2 + \frac{1}{\text{poly } m}, \tag{3.8}$$

with probability greater than $1 - 1/\exp(m)$.

The overlap between two quantum states is a measure of their distinguishability [Die88] and overlap estimation thus is related to quantum state discrimination. Several techniques exist to perform quantumly the overlap estimation of two states $|\phi\rangle$ and $|\psi\rangle$ [FRS+20]. One of them is to perform the swap test (see chapter 5 or [BCWDW01]) with various copies of both states.

Overlap estimation can be done efficiently classically for IQP circuits [HCT+19]. We prove in the next section that it is also the case for Boson Sampling and consider the more general setting of passive linear optical quantum computing with input single-photons and adaptive measurements.

3.2 Adaptive linear optics

The complexity of probability estimation and overlap estimation of quantum computations has been well studied in the circuit model [PWB15, BGM19]. In what follows, we consider the case of passive linear optical quantum computing with adaptive measurements, which we refer to as adaptive linear optics (Fig. 3.1). We use multi-index notations (see section 1.1.1).

Formally, we consider unitary interferometers of size m, described by $m \times m$ unitary matrices (see section 1.4). We identify the multimode Fock states with n photons over m modes with the elements of $\Phi_{m,n} = \{\boldsymbol{s} \in \mathbb{N}^m,\ |\boldsymbol{s}| = n\}$, for all $n \in \mathbb{N}$. We fix the input state $|\boldsymbol{t}\rangle = |\mathbf{1}^n \mathbf{0}^{m-n}\rangle$, with single photons in the n first modes, where the superscript indicates the size of the string $(0,\dots,0)$ or $(1,\dots,1)$ when there is a possible ambiguity. For $p \in \mathbb{N}$ and $\boldsymbol{p} \in \Phi_{k,p}$, let us define

$$U^{\boldsymbol{p}} := [\mathbb{1}_k \oplus U_k(p_1,\dots,p_k)][\mathbb{1}_{k-1} \oplus U_{k-1}(p_1,\dots,p_{k-1})]\dots[\mathbb{1}_1 \oplus U_1(p_1)]U_0, \tag{3.9}$$

where $\mathbb{1}_j$ is the identity matrix of size j. The matrices U_j depend on the measurement outcomes $p_1,\dots,p_j$ for all $j \in \{1,\dots,k\}$. The output state where the adaptive measurement outcome $\boldsymbol{p}$ has been obtained reads

$$\mathrm{Tr}_k \left[(|\boldsymbol{p}\rangle\langle\boldsymbol{p}| \otimes \mathbb{1}_{m-k}) U^{\boldsymbol{p}} |\boldsymbol{t}\rangle\langle\boldsymbol{t}| U^{\boldsymbol{p}\dagger} \right], \tag{3.10}$$

where the partial trace is over the first k modes and where $|\boldsymbol{p}\rangle$ denotes the k-mode Fock state $|p_1 \dots p_k\rangle$. The matrix $U^{\boldsymbol{p}}$ describes the interferometer in Fig. 3.1, where the adaptive measurement outcome $\boldsymbol{p} = (p_1,\dots,p_k)$ and the final outcome $\boldsymbol{s} = (s_1,\dots,s_{m-k})$ have been obtained.

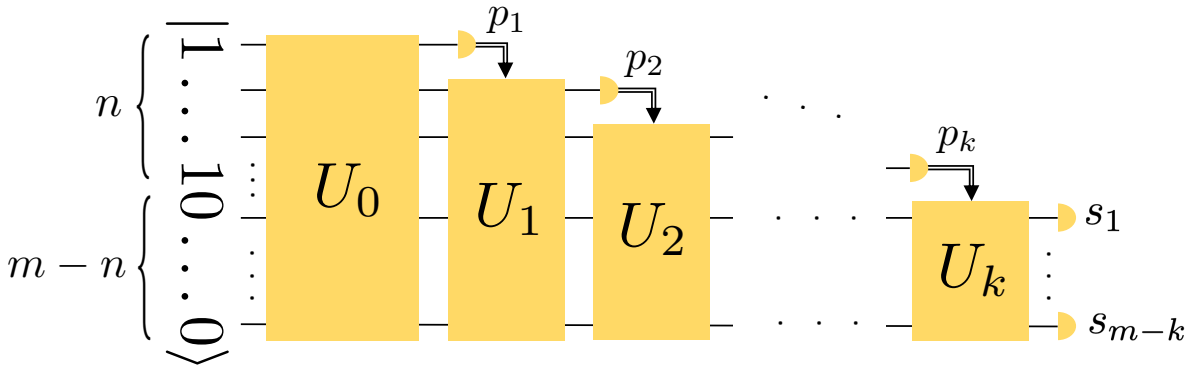

Figure 3.1: Passive linear optical computing with k adaptive measurements and input state $|1\dots10\dots0\rangle$ with n photons over m modes. The output modes are measured using photon counters. For all $j \in \{1,\dots,k\}$, the unitary interferometer U_j, acting on $m-j$ modes, may depend on the measurement outcomes $p_1,\dots,p_j$. The adaptive measurement outcomes $p_1,\dots,p_k$ are used to drive the computation, whose final outcome is $s_1,\dots,s_{m-k}$.

Boson Sampling [AA13] corresponds to the case $k = 0$ and the Knill–Laflamme–Milburn scheme for universal quantum computing [KLM01] to the case $k = O(m)$. We investigate the transition

between these two cases by giving classical algorithms for probability estimation and overlap estimation and identifying various complexity regimes for different numbers of photons n and adaptive measurements k.

3.2.1 Quantum probability and overlap estimation

For doing probability estimation with a quantum circuit, one samples the circuit $O(\text{poly } m)$ times, obtaining outcomes, for which the frequency gives a polynomially precise additive estimate of the probability which can be computed efficiently. In the case of a circuit with adaptive measurements, one only looks at the final measurement outcomes and the same holds for adaptive linear optical computations.

For doing overlap estimation with unitary quantum circuits, one may run two circuits U and V in parallel and compare their quantum output states, for example with the swap test. Doing so a polynomial number of times provides a polynomially precise estimate of the overlap. Alternatively, one may build the circuit $UV^\dagger$ and project the output quantum state onto the input state.

In the case of circuits with adaptive measurements, the overlaps are between all possible output states for all possible adaptive measurement results. In particular, if the number of possible adaptive measurement outcomes is exponential, then the probability distribution for these outcomes has to be concentrated on a polynomial number of events for the quantum overlap estimation to be efficient. This is because in order to compute a polynomially precise estimate of the overlap, say, $|\langle\phi|\psi\rangle|^2$, the states $|\phi\rangle$ and $|\psi\rangle$, both corresponding to specific adaptive measurement results, have to be obtained a polynomial number of times.

For adaptive linear optics over m modes with n input photons and k adaptive measurements, the number of possible adaptive measurement outcomes is given by

$$\begin{aligned}\sum_{r=0}^{n} |\Phi_{k,r}| &= \sum_{r=0}^{n} \binom{k+r-1}{r} \\ &= \binom{k+n}{n},\end{aligned} \tag{3.11}$$

where the sum is over the total number of photons detected at the stage of the adaptive measurements. Hence, either the probability distribution for the adaptive measurements outcomes is concentrated on a polynomial number of outcomes, or $\binom{n+k}{n} = O(\text{poly } m)$, which is the case for example when $n = O(1)$ and $k = O(m)$, $n = O(\log m)$ and $k = O(\log m)$, or $n = O(m)$ and $k = O(1)$. In what follows, we do not assume concentration of the adaptive measurement outcome probability distribution and consider general interferometers with adaptive measurements. The quantum efficient regime for overlap estimation thus corresponds to $\binom{n+k}{n} = O(\text{poly } m)$.

Let $|\phi\rangle$ and $|\psi\rangle$ be output states of two adaptive linear interferometers over m modes with n input photons and k adaptive measurements. Let $\boldsymbol{p}$ and $\boldsymbol{q}$ denote the outcomes of the adaptive measurements for $|\phi\rangle$ and $|\psi\rangle$, respectively. Let $U^{\boldsymbol{p}}$ in Eq. (3.9) be the interferometer for $|\phi\rangle$, with

input Fock state $|\boldsymbol{t}\rangle$. We have

$$\begin{aligned} |\langle\phi|\psi\rangle|^2 &= \mathrm{Tr}\left[\mathrm{Tr}_k[(|\boldsymbol{p}\rangle\langle\boldsymbol{p}|\otimes\mathbb{1}_{m-k})U^{\boldsymbol{p}}|\boldsymbol{t}\rangle\langle\boldsymbol{t}|U^{\boldsymbol{p}\dagger}]|\psi\rangle\langle\psi|\right] \\ &= \mathrm{Tr}\left[(|\boldsymbol{p}\rangle\langle\boldsymbol{p}|\otimes\mathbb{1}_{m-k})U^{\boldsymbol{p}}|\boldsymbol{t}\rangle\langle\boldsymbol{t}|U^{\boldsymbol{p}\dagger}(\mathbb{1}_k\otimes|\psi\rangle\langle\psi|)\right] \\ &= \mathrm{Tr}\left[U^{\boldsymbol{p}}|\boldsymbol{t}\rangle\langle\boldsymbol{t}|U^{\boldsymbol{p}\dagger}(|\boldsymbol{p}\rangle\langle\boldsymbol{p}|\otimes|\psi\rangle\langle\psi|)\right] \\ &= \mathrm{Tr}\left[|\boldsymbol{t}\rangle\langle\boldsymbol{t}|U^{\boldsymbol{p}\dagger}\left(|\boldsymbol{p}\rangle\langle\boldsymbol{p}|\otimes|\psi\rangle\langle\psi|\right)U^{\boldsymbol{p}}\right], \end{aligned} \tag{3.12}$$

where we used Eq. (3.10) in the first line. Because of the conservation of the total number of photons, the overlap between the states $|\phi\rangle$ and $|\psi\rangle$ is zero if $|\boldsymbol{p}| \neq |\boldsymbol{q}|$. Otherwise, it can be estimated using a polynomial number of copies of the state $|\psi\rangle$ as follows: send the input $|\boldsymbol{p}\rangle\otimes|\psi\rangle$ into the interferometer with unitary matrix $U^{\boldsymbol{p}\dagger}$ and mesure the photon number in each output mode. Record the value 1 if the measurement pattern matches the Fock state $\boldsymbol{t}$ and the value 0 otherwise. Then, the mean of the obtained values yields a polynomially precise estimate of the overlap $|\langle\phi|\psi\rangle|^2$ by Eq. (3.12) and Hoeffding inequality. Note that this overlap estimation requires the preparation of the Fock state $\boldsymbol{p}$. By symmetry, one could estimate the overlap alternatively using a polynomial number of copies of the state $|\phi\rangle$ and preparing the Fock state $|\boldsymbol{q}\rangle$.

3.2.2 Classical probability estimation

In this section, we obtain a classical algorithm for probability estimation of adaptive linear optics over m modes with n input photons and k adaptive measurements.

We first consider the case $k = 0$, i.e., Boson Sampling. The probability of the outcome $\boldsymbol{s} \in \Phi_{m,n}$ for the interferometer U given the input $\boldsymbol{t} = (\mathbf{1}^n, \mathbf{0}^{m-n}) \in \Phi_{m,n}$ is given by (see section 1.4.5 and [AA13])

$$\Pr_{m,n}[\boldsymbol{s}] = \frac{1}{\boldsymbol{s}!}|\mathrm{Per}(U_{\boldsymbol{s},\boldsymbol{t}})|^2, \tag{3.13}$$

where $U_{\boldsymbol{s},\boldsymbol{t}}$ is the $n\times n$ matrix obtained from U by repeating s_i times its i^{th} row for $i \in \{1,\dots,m\}$ and removing its j^{th} column for $j = \{n+1,\dots,m\}$, and where the permanent of an $n\times n$ square matrix $A = (a_{ij}))_{1\leq i,j\leq n}$ is given by

$$\mathrm{Per}(A) = \sum_{\sigma\in S_n}\prod_{i=1}^{n} a_{i\sigma(i)}, \tag{3.14}$$

where the sum is over the permutations of the set $\{1,\dots,n\}$. When $|\boldsymbol{s}| \neq n$ however, the probability is 0, since $\boldsymbol{t}$ has n photons and the linear interferometer does not change the total number of photons. The permanent of a square matrix of size n can be computed exactly in time $O(n2^n)$, thanks to Ryser's formula [AWH78]. However, polynomially precise estimates of the permanent can be obtained in polynomial time [Gur05], so the probability estimation can be done classically efficiently, which was already noted in [AA13].

We now turn to the case $k > 0$, using notations of Eq. (3.9) and Fig. 3.1. This case is a direct extension of the case $k = 0$. For $p \in \mathbb{N}$, $\boldsymbol{p} \in \Phi_{k,p}$ and $\boldsymbol{s} \in \Phi_{m-k,n-p}$, the probability of an total

outcome $(\boldsymbol{p},\boldsymbol{s}) \in \Phi_{m,n}$ (adaptive measurement and final outcome) is given by

$$\Pr^{\text{total}}_{m,n}[\boldsymbol{p},\boldsymbol{s}] = \frac{1}{\boldsymbol{p}!\boldsymbol{s}!}\left|\text{Per}\left(U^{\boldsymbol{p}}_{(\boldsymbol{p},\boldsymbol{s}),\boldsymbol{t}}\right)\right|^2. \tag{3.15}$$

Let $p \in \{0,\dots,n\}$ and let $\boldsymbol{s} \in \Phi_{m-k,n-p}$. Then, the probability of obtaining the final outcome $\boldsymbol{s}$ after the adaptive measurements reads

$$\begin{aligned}\Pr^{\text{final}}_{m,n}[\boldsymbol{s}] &= \sum_{\boldsymbol{p}\in\Phi_{k,p}} \Pr^{\text{total}}_{m,n}[\boldsymbol{p},\boldsymbol{s}] \\ &= \frac{1}{\boldsymbol{s}!}\sum_{\boldsymbol{p}\in\Phi_{k,p}} \frac{1}{\boldsymbol{p}!}\left|\text{Per}\left(U^{\boldsymbol{p}}_{(\boldsymbol{p},\boldsymbol{s}),\boldsymbol{t}}\right)\right|^2.\end{aligned} \tag{3.16}$$

The sum is taken over the elements of $\Phi_{k,p}$, which has $\binom{k+p-1}{p} \leq \binom{k+n-1}{n}$ elements. This last quantity is $O(\text{poly } m)$ when the number of input photons n and the number of adaptive measurements k are small enough compared to m.

n \ k	$O(1)$	$O(\log m)$	$O(m)$
$O(1)$	$O(m)$	$O(\text{poly } m)$	$O(\text{poly } m)$
$O(\log m)$	$O(\text{poly } m)$	$O(\text{poly } m)$	$O(2^{\log^2 m})$
$O(m)$	$O(\text{poly } m)$	$O(2^{\log^2 m})$	$O(4^m)$

Table 3.1: Simulability regimes for probability estimation. The scalings indicated are upper bounds on the running time of the classical algorithm, up to a factor $O(\text{poly } m)$. In blue is the parameter region for which the quantum algorithm is efficient.

The simulability regimes are summarised in Table 3.1, where the scalings are obtained using Stirling equivalent $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$. In particular, as long as both k and n are $O(\log m)$, the output probability can be estimated efficiently (and even computed exactly efficiently).

The universal quantum computing regime corresponds to $n = O(m)$ and $k = O(m)$. The time complexity of the classical algorithm is $O\left(\binom{k+n-1}{n}\text{poly } m\right)$, so there is a possibility of subuniversal quantum advantage for probability estimation for $n = O(\log m)$ and $k = O(m)$, or $n = O(m)$ and $k = O(\log m)$. However, the runtime of the classical algorithm is subexponential in these cases, namely $O(2^{\log^2 m})$.

3.2.3 Classical overlap estimation

In this section, we obtain a classical algorithm for overlap estimation of adaptive linear optics over m modes with n input photons and k adaptive measurements.

Once again, we start with $k = 0$. The output state of an m-mode interferometer U with input state $\boldsymbol{t} \in \Phi_{m,n}$ reads

$$\begin{aligned} |\phi\rangle &= \sum_{\boldsymbol{s}\in\Phi_{m,n}} \langle \boldsymbol{s}|\hat{U}|\boldsymbol{t}\rangle\, |\boldsymbol{s}\rangle \\ &= \sum_{\boldsymbol{s}\in\Phi_{m,n}} \frac{\mathrm{Per}(U_{\boldsymbol{s},\boldsymbol{t}})}{\sqrt{\boldsymbol{s}!\boldsymbol{t}!}}\, |\boldsymbol{s}\rangle, \end{aligned} \tag{3.17}$$

where $U_{\boldsymbol{s},\boldsymbol{t}}$ is the $n \times n$ matrix obtained from U by repeating s_i times its i^{th} row for $i \in \{1,\dots,m\}$ and repeating t_j times its j^{th} row for $j \in \{1,\dots,m\}$. The composition of two interferometers is another interferometer which unitary representation is the product of the unitary representations of the composed interferometers. Hence, the inner product of the output states $|\phi\rangle$ and $|\psi\rangle$ of two m-mode interferometers U and V with the same input state $\boldsymbol{t} \in \Phi_{m,n}$, is equal to the matrix element $\boldsymbol{t},\boldsymbol{t}$ of $\hat{U}^\dagger\hat{V}$:

$$\begin{aligned} \langle\phi|\psi\rangle &= \sum_{\boldsymbol{u},\boldsymbol{v}\in\Phi_{m,n}} \langle \boldsymbol{t}|\hat{U}^\dagger|\boldsymbol{u}\rangle \langle \boldsymbol{v}|\hat{V}|\boldsymbol{t}\rangle \langle \boldsymbol{u}|\boldsymbol{v}\rangle \\ &= \sum_{\boldsymbol{s}\in\Phi_{m,n}} \langle \boldsymbol{t}|\hat{U}^\dagger|\boldsymbol{s}\rangle \langle \boldsymbol{s}|\hat{V}|\boldsymbol{t}\rangle \\ &= \langle \boldsymbol{t}|\hat{U}^\dagger\hat{V}|\boldsymbol{t}\rangle \\ &= \frac{\mathrm{Per}\left[(U^\dagger V)_{\boldsymbol{t},\boldsymbol{t}}\right]}{\boldsymbol{t}!}, \end{aligned} \tag{3.18}$$

where we used in the third line $\boldsymbol{t} \in \Phi_{m,n}$ and the fact that $\hat{U}^\dagger\hat{V}$ conserves the space $\Phi_{m,n}$. With the input $\boldsymbol{t} = (\mathbf{1}^n, \mathbf{0}^{m-n})$ with n photons in m modes, this reduces to

$$\langle\phi|\psi\rangle = \mathrm{Per}\left[(U^\dagger V)_n\right], \tag{3.19}$$

where $(U^\dagger V)_n$ is the $n \times n$ top left submatrix of $U^\dagger V$. Hence, the inner product and the overlap may be approximated to a polynomial precision efficiently, since this is the case for the permanent [Gur05].

We now consider the case $k > 0$. Let $p \in \mathbb{N}$ and let $\boldsymbol{p} \in \Phi_{k,p}$. Writing $\mathrm{Pr}^{\mathrm{adap}}_{m,n}[\boldsymbol{p}]$ the probability of the adaptive measurement outcome $\boldsymbol{p}$, the output state of the interferometer $U^{\boldsymbol{p}}$ with k adaptive measurements with input $\boldsymbol{t} = (\mathbf{1}^n, \mathbf{0}^{m-n})$ in Fig. 3.1, when the adaptive measurement outcome $\boldsymbol{p}$ is obtained, reads

$$\frac{1}{\sqrt{\mathrm{Pr}^{\mathrm{adap}}_{m,n}[\boldsymbol{p}]}}\, |\psi_{\boldsymbol{p}}\rangle, \tag{3.20}$$

where

$$|\psi_{\boldsymbol{p}}\rangle := \sum_{\boldsymbol{s}\in\Phi_{m-k,n-p}} \frac{\mathrm{Per}\left(U^{\boldsymbol{p}}_{(\boldsymbol{p},\boldsymbol{s}),\boldsymbol{t}}\right)}{\sqrt{\boldsymbol{p}!\boldsymbol{s}!}}\, |\boldsymbol{s}\rangle \tag{3.21}$$

and where $\mathrm{Pr}^{\mathrm{adap}}_{m,n}[\boldsymbol{p}] = \langle\psi_{\boldsymbol{p}}|\psi_{\boldsymbol{p}}\rangle$. More generally, the inner product of two (not normalised) output states $|\psi_{\boldsymbol{p}}\rangle$ and $|\psi_{\boldsymbol{q}}\rangle$ of m-mode interferometers $U^{\boldsymbol{p}}$ and $V^{\boldsymbol{q}}$ with k adaptive measurements thus

is zero if $|\boldsymbol{p}| \neq |\boldsymbol{q}|$. If $r := |\boldsymbol{p}| = |\boldsymbol{q}|$, it is given by

$$\begin{aligned}\langle \psi_{\boldsymbol{p}} | \psi_{\boldsymbol{q}} \rangle &= \frac{1}{\sqrt{\boldsymbol{p}!\boldsymbol{q}!}} \sum_{\boldsymbol{s}\in\Phi_{m-k,n-r}} \frac{1}{\boldsymbol{s}!} \operatorname{Per}\left(U^{\boldsymbol{p}}_{(\boldsymbol{p},\boldsymbol{s}),\boldsymbol{t}}\right)^* \operatorname{Per}\left(V^{\boldsymbol{q}}_{(\boldsymbol{q},\boldsymbol{s}),\boldsymbol{t}}\right) \\ &= \frac{1}{\sqrt{\boldsymbol{p}!\boldsymbol{q}!}} \sum_{\boldsymbol{s}\in\Phi_{m-k,n-r}} \frac{1}{\boldsymbol{s}!} \operatorname{Per}\left(U^{\boldsymbol{p}^\dagger}_{\boldsymbol{t},(\boldsymbol{p},\boldsymbol{s})}\right) \operatorname{Per}\left(V^{\boldsymbol{q}}_{(\boldsymbol{q},\boldsymbol{s}),\boldsymbol{t}}\right).\end{aligned} \tag{3.22}$$

This expression is a sum of $|\Phi_{m-k,n-r}|$ terms, which is generally exponential in m whenever n is not constant. It is reminiscent of the permanent composition formula [Per12, Bar16]: for all $m,n,c \in \mathbb{N}^*$, all $s \in \mathbb{N}$, all $\boldsymbol{u} \in \Phi_{m,s}$ and all $\boldsymbol{v} \in \Phi_{n,s}$,

$$\operatorname{Per}\left[(MN)_{\boldsymbol{u},\boldsymbol{v}}\right] = \sum_{\boldsymbol{s}\in\Phi_{c,s}} \frac{1}{\boldsymbol{s}!} \operatorname{Per}\left(M_{\boldsymbol{u},\boldsymbol{s}}\right) \operatorname{Per}\left(N_{\boldsymbol{s},\boldsymbol{v}}\right) \tag{3.23}$$

where M is a $m \times c$ matrix and N is a $n \times c$ matrix. In what follows, we prove that this expression in Eq. (3.22) may be rewritten as a sum over fewer terms using the permanent composition formula in Eq. (3.23). However, this formula is not directly applicable to the expression in Eq. (3.22). In order to obtain a suitable expression, we first make use of the Laplace formula for the permanent: we expand the permanent of $U^{\boldsymbol{p}^\dagger}_{\boldsymbol{t},(\boldsymbol{p},\boldsymbol{s})}$ along the columns that are repeated according to $\boldsymbol{p}$ and we expand the permanent of $V^{\boldsymbol{q}}_{(\boldsymbol{q},\boldsymbol{s}),\boldsymbol{t}}$ along the rows that are repeated according to $\boldsymbol{q}$. The general Laplace column expansion formula for the permanent reads: let $n \in \mathbb{N}^*$, let W be an $n \times n$ matrix, and let $\boldsymbol{j} \in \{0,1\}^n$. Then,

$$\operatorname{Per}(W) = \sum_{\substack{\boldsymbol{i}\in\{0,1\}^n \\ |\boldsymbol{i}|=|\boldsymbol{j}|}} \operatorname{Per}\left(W_{\boldsymbol{i},\boldsymbol{j}}\right) \operatorname{Per}\left(W_{\mathbf{1}^n-\boldsymbol{i},\mathbf{1}^n-\boldsymbol{j}}\right), \tag{3.24}$$

where $W_{\boldsymbol{i},\boldsymbol{j}}$ is the matrix obtained from W by keeping only the k^{th} rows and l^{th} columns such that $i_k = 1$ and $j_l = 1$, respectively, and $W_{\mathbf{1}^n-\boldsymbol{i},\mathbf{1}^n-\boldsymbol{j}}$ is the matrix obtained from W by keeping only the k^{th} rows and l^{th} columns such that $i_k = 0$ and $j_l = 0$, respectively. This formula is obtained by applying the Laplace expansion formula for one column various times, for each column with index l such that $j_l = 1$, and the same formula holds for rows.

Lemma 3.2. *Let $r \in \mathbb{N}$. The inner product of two (not normalised) output states $|\psi_{\boldsymbol{p}}\rangle$ and $|\psi_{\boldsymbol{q}}\rangle$ of m-mode interferometers $U^{\boldsymbol{p}}$ and $V^{\boldsymbol{q}}$ with adaptive measurements outcome $\boldsymbol{p},\boldsymbol{q} \in \Phi_{k,r}$ is given by*

$$\langle \psi_{\boldsymbol{p}} | \psi_{\boldsymbol{q}} \rangle = \frac{1}{\sqrt{\boldsymbol{p}!\boldsymbol{q}!}} \sum_{\substack{\boldsymbol{i},\boldsymbol{j}\in\{0,1\}^n \\ |\boldsymbol{i}|=|\boldsymbol{j}|=r}} \operatorname{Per}\left(A^{\boldsymbol{i}}\right) \operatorname{Per}\left(B^{\boldsymbol{j}}\right) \operatorname{Per}\left(C^{\boldsymbol{i},\boldsymbol{j}}\right), \tag{3.25}$$

where for all $\boldsymbol{i},\boldsymbol{j} \in \{0,1\}^n$ such that $|\boldsymbol{i}| = |\boldsymbol{j}| = r$,

$$A^{\boldsymbol{i}} = U^{\boldsymbol{p}^\dagger}_{(\boldsymbol{i},\mathbf{0}^{m-n}),(\boldsymbol{p},\mathbf{0}^{m-k})} \tag{3.26}$$

is an $r \times r$ matrix which can be obtained efficiently from $U^{\boldsymbol{p}}$,

$$B^{\boldsymbol{j}} = V^{\boldsymbol{q}}_{(\boldsymbol{q},\mathbf{0}^{m-k}),(\boldsymbol{j},\mathbf{0}^{m-n})} \tag{3.27}$$

is an $r \times r$ matrix which can be obtained efficiently from $V^{\boldsymbol{q}}$, and

$$C^{\boldsymbol{i},\boldsymbol{j}} = U^{\boldsymbol{p}\dagger}_{(\mathbf{1}^n-\boldsymbol{i},\mathbf{0}^{m-n}),(\mathbf{0}^k,\mathbf{1}^{m-k})} V^{\boldsymbol{q}}_{(\mathbf{0}^k,\mathbf{1}^{m-k}),(\mathbf{1}^n-\boldsymbol{j},\mathbf{0}^{m-n})} \tag{3.28}$$

is an $(n-r)\times(n-r)$ matrix which can be obtained efficiently from $U^{\boldsymbol{p}}$ and $V^{\boldsymbol{q}}$.

Proof. We consider the expression for the inner product obtained in Eq. (3.22):

$$\langle \psi_{\boldsymbol{p}} | \psi_{\boldsymbol{q}} \rangle = \frac{1}{\sqrt{\boldsymbol{p}!\boldsymbol{q}!}} \sum_{\boldsymbol{s}\in\Phi_{m-k,n-r}} \frac{1}{\boldsymbol{s}!} \mathrm{Per}\left(U^{\boldsymbol{p}\dagger}_{\boldsymbol{t},(\boldsymbol{p},\boldsymbol{s})}\right) \mathrm{Per}\left(V^{\boldsymbol{q}}_{(\boldsymbol{q},\boldsymbol{s}),\boldsymbol{t}}\right). \tag{3.29}$$

We first apply the general column expansion formula in Eq. (3.24) to the matrix $U^{\boldsymbol{p}\dagger}_{\boldsymbol{t},(\boldsymbol{p},\boldsymbol{s})}$ with $\boldsymbol{j} = (\mathbf{1}^r, \mathbf{0}^{n-r}) \in \{0,1\}^n$, obtaining

$$\mathrm{Per}\left(U^{\boldsymbol{p}\dagger}_{\boldsymbol{t},(\boldsymbol{p},\boldsymbol{s})}\right) = \sum_{\substack{\boldsymbol{i}\in\{0,1\}^n \\ |\boldsymbol{i}|=r}} \mathrm{Per}\left[\left(U^{\boldsymbol{p}\dagger}_{\boldsymbol{t},(\boldsymbol{p},\boldsymbol{s})}\right)_{\boldsymbol{i},\boldsymbol{j}}\right] \mathrm{Per}\left[\left(U^{\boldsymbol{p}\dagger}_{\boldsymbol{t},(\boldsymbol{p},\boldsymbol{s})}\right)_{\mathbf{1}^n-\boldsymbol{i},\mathbf{1}^n-\boldsymbol{j}}\right]. \tag{3.30}$$

Let us consider the matrix $\left(U^{\boldsymbol{p}\dagger}_{\boldsymbol{t},(\boldsymbol{p},\boldsymbol{s})}\right)_{\boldsymbol{i},\boldsymbol{j}}$ appearing in this last expression, for $\boldsymbol{i} \in \{0,1\}^n$. Its rows are obtained by keeping the first n lines of $U^{\boldsymbol{p}\dagger}$ since $\boldsymbol{t} = (\mathbf{1}^n, \mathbf{0}^{m-n})$, then by keeping only the l^{th} rows such that $i_l = 1$. Its columns are obtained by repeating p_l times the l^{th} column for $l \in \{1,\ldots,k\}$ and s_l times for $l \in \{k+1,\ldots,m\}$, then by only keeping the first r columns since $\boldsymbol{j} = (\mathbf{1}^r, \mathbf{0}^{n-r})$. However, since $|\boldsymbol{p}| = |\boldsymbol{j}| = r$, these are the columes repeated according to $\boldsymbol{p}$. Hence,

$$\left(U^{\boldsymbol{p}\dagger}_{\boldsymbol{t},(\boldsymbol{p},\boldsymbol{s})}\right)_{\boldsymbol{i},\boldsymbol{j}} = U^{\boldsymbol{p}\dagger}_{(\boldsymbol{i},\mathbf{0}^{m-n}),(\boldsymbol{p},\mathbf{0}^{m-k})}, \tag{3.31}$$

where $U^{\boldsymbol{p}\dagger}_{(\boldsymbol{i},\mathbf{0}^{m-n}),(\boldsymbol{p},\mathbf{0}^{m-k})}$ is the matrix obtained from $U^{\boldsymbol{p}\dagger}$ by keeping only the l^{th} rows such that $i_l = 1$ and removing the others, and by repeating p_l times the l^{th} column for $l \in \{1,\ldots,k\}$ and removing the others. Similarly, with $|\boldsymbol{s}| = |\mathbf{1}^n - \boldsymbol{j}| = n-r$,

$$\left(U^{\boldsymbol{p}\dagger}_{\boldsymbol{t},(\boldsymbol{p},\boldsymbol{s})}\right)_{\mathbf{1}^n-\boldsymbol{i},\mathbf{1}^n-\boldsymbol{j}} = U^{\boldsymbol{p}\dagger}_{(\mathbf{1}^n-\boldsymbol{i},\mathbf{0}^{m-n}),(\mathbf{0}^k,\boldsymbol{s})}, \tag{3.32}$$

where $U^{\boldsymbol{p}\dagger}_{(\mathbf{1}^n-\boldsymbol{i},\mathbf{0}^{m-n}),(\mathbf{0}^k,\boldsymbol{s})}$ is the matrix obtained from $U^{\boldsymbol{p}\dagger}$ by keeping only the l^{th} rows such that $i_l = 0$ and removing the others, and by repeating s_l times the l^{th} column for $l \in \{k+1,\ldots,m\}$ and removing the others. With Eqs. (3.30), (3.31) and (3.32) we obtain

$$\begin{aligned} \mathrm{Per}\left(U^{\boldsymbol{p}\dagger}_{\boldsymbol{t},(\boldsymbol{p},\boldsymbol{s})}\right) &= \sum_{\substack{\boldsymbol{i}\in\{0,1\}^n \\ |\boldsymbol{i}|=r}} \mathrm{Per}\left(U^{\boldsymbol{p}\dagger}_{(\boldsymbol{i},\mathbf{0}^{m-n}),(\boldsymbol{p},\mathbf{0}^{m-k})}\right) \mathrm{Per}\left(U^{\boldsymbol{p}\dagger}_{(\mathbf{1}^n-\boldsymbol{i},\mathbf{0}^{m-n}),(\mathbf{0}^k,\boldsymbol{s})}\right) \\ &= \sum_{\substack{\boldsymbol{i}\in\{0,1\}^n \\ |\boldsymbol{i}|=r}} \mathrm{Per}\left(A^{\boldsymbol{i}}\right) \mathrm{Per}\left(U^{\boldsymbol{p}\dagger}_{(\mathbf{1}^n-\boldsymbol{i},\mathbf{0}^{m-n}),(\mathbf{0}^k,\boldsymbol{s})}\right), \end{aligned} \tag{3.33}$$

where we have defined, for all $\boldsymbol{i} \in \{0,1\}^n$ such that $|\boldsymbol{i}| = r$,

$$A^{\boldsymbol{i}} := U^{\boldsymbol{p}\dagger}_{(\boldsymbol{i},\mathbf{0}^{m-n}),(\boldsymbol{p},\mathbf{0}^{m-k})}, \tag{3.34}$$

which is an $r \times r$ matrix independent of $\boldsymbol{s}$ that can be obtained efficiently from $U^{\boldsymbol{p}}$.

The same reasoning with the general row expansion formula for the matrix $V^{\boldsymbol{q}}_{(\boldsymbol{q},\boldsymbol{s}),\boldsymbol{t}}$ and the rows $\boldsymbol{i} = (\mathbf{1}^r, \mathbf{0}^{n-r})$ gives

$$\begin{aligned} \mathrm{Per}\left(V^{\boldsymbol{q}}_{(\boldsymbol{q},\boldsymbol{s}),\boldsymbol{t}}\right) &= \sum_{\substack{\boldsymbol{j}\in\{0,1\}^n \\ |\boldsymbol{j}|=r}} \mathrm{Per}\left[\left(V^{\boldsymbol{q}}_{(\boldsymbol{q},\boldsymbol{s}),\boldsymbol{t}}\right)_{\boldsymbol{i},\boldsymbol{j}}\right] \mathrm{Per}\left[\left(V^{\boldsymbol{q}}_{(\boldsymbol{q},\boldsymbol{s}),\boldsymbol{t}}\right)_{\mathbf{1}^n-\boldsymbol{i},\mathbf{1}^n-\boldsymbol{j}}\right] \\ &= \sum_{\substack{\boldsymbol{j}\in\{0,1\}^n \\ |\boldsymbol{j}|=r}} \mathrm{Per}\left(V^{\boldsymbol{q}}_{(\boldsymbol{q},\mathbf{0}^{m-k}),(\boldsymbol{j},\mathbf{0}^{m-n})}\right) \mathrm{Per}\left(V^{\boldsymbol{q}}_{(\mathbf{0}^k,\boldsymbol{s}),(\mathbf{1}^n-\boldsymbol{j},\mathbf{0}^{m-n})}\right), \end{aligned} \tag{3.35}$$

where $V^{\boldsymbol{q}}_{(\boldsymbol{q},\mathbf{0}^{m-k}),(\boldsymbol{j},\mathbf{0}^{m-n})}$ is the matrix obtained from $V^{\boldsymbol{q}}$ by repeating q_l times the l^{th} row for $l \in \{1,\ldots,k\}$ and removing the others and by keeping only the l^{th} columns such that $j_l = 1$, and where $V^{\boldsymbol{q}}_{(\mathbf{0}^k,\boldsymbol{s}),(\mathbf{1}^n-\boldsymbol{j},\mathbf{0}^{m-n})}$ is the matrix obtained from $V^{\boldsymbol{q}}$ by repeating s_l times the l^{th} row for $l \in \{k+1,\ldots,m\}$ and removing the others and by keeping only the l^{th} columns such that $j_l = 0$. Defining, for all $\boldsymbol{j} \in \{0,1\}^n$ such that $|\boldsymbol{j}| = r$,

$$B^{\boldsymbol{j}} := V^{\boldsymbol{q}}_{(\boldsymbol{q},\mathbf{0}^{m-k}),(\boldsymbol{j},\mathbf{0}^{m-n})}, \tag{3.36}$$

the expression in Eq. (3.35) rewrites

$$\mathrm{Per}\left(V^{\boldsymbol{q}}_{(\boldsymbol{q},\boldsymbol{s}),\boldsymbol{t}}\right) = \sum_{\substack{\boldsymbol{j}\in\{0,1\}^n \\ |\boldsymbol{j}|=r}} \mathrm{Per}\left(B^{\boldsymbol{j}}\right) \mathrm{Per}\left(V^{\boldsymbol{q}}_{(\mathbf{0}^k,\boldsymbol{s}),(\mathbf{1}^n-\boldsymbol{j},\mathbf{0}^{m-n})}\right), \tag{3.37}$$

where $B^{\boldsymbol{j}}$ are $r \times r$ matrices independent of $\boldsymbol{s}$ and can be obtained efficiently from $V^{\boldsymbol{q}}$.

Plugging Eqs. (3.33) and (3.37) in Eq. (3.29) we obtain

$$\begin{aligned} \langle \psi_{\boldsymbol{p}} | \psi_{\boldsymbol{q}} \rangle = \frac{1}{\sqrt{\boldsymbol{p}!\boldsymbol{q}!}} \sum_{\substack{\boldsymbol{i},\boldsymbol{j}\in\{0,1\}^n \\ |\boldsymbol{i}|=|\boldsymbol{j}|=r}} \Bigg[&\mathrm{Per}\left(A^{\boldsymbol{i}}\right) \mathrm{Per}\left(B^{\boldsymbol{j}}\right) \\ &\times \sum_{\boldsymbol{s}\in\Phi_{m-k,n-r}} \frac{1}{\boldsymbol{s}!} \mathrm{Per}\left(U^{\boldsymbol{p}\dagger}_{(\mathbf{1}^n-\boldsymbol{i},\mathbf{0}^{m-n}),(\mathbf{0}^k,\boldsymbol{s})}\right) \mathrm{Per}\left(V^{\boldsymbol{q}}_{(\mathbf{0}^k,\boldsymbol{s}),(\mathbf{1}^n-\boldsymbol{j},\mathbf{0}^{m-n})}\right) \Bigg]. \end{aligned} \tag{3.38}$$

The sum appearing in the second line may now be expressed as a single permanent using the permanent composition formula: for all $\boldsymbol{i},\boldsymbol{j} \in \{0,1\}^n$ such that $|\boldsymbol{i}| = |\boldsymbol{j}| = r$, let us define the $(n-r) \times (m-k)$ matrix

$$\tilde{U}^{\boldsymbol{p},\boldsymbol{i}} := U^{\boldsymbol{p}\dagger}_{(\mathbf{1}^n-\boldsymbol{i},\mathbf{0}^{m-n}),(\mathbf{0}^k,\mathbf{1}^{m-k})}, \tag{3.39}$$

and the $(m-k)\times(n-r)$ matrix

$$\tilde{V}^{\boldsymbol{q},\boldsymbol{j}} := V^{\boldsymbol{q}}_{(\mathbf{0}^k,\mathbf{1}^{m-k}),(\mathbf{1}^n-\boldsymbol{j},\mathbf{0}^{m-n})}, \tag{3.40}$$

so that

$$U^{\boldsymbol{p}\dagger}_{(\mathbf{1}^n-\boldsymbol{i},\mathbf{0}^{m-n}),(\mathbf{0}^k,\boldsymbol{s})} = \tilde{U}^{\boldsymbol{p},\boldsymbol{i}}_{\mathbf{1}^{n-r},\boldsymbol{s}} \quad \text{and} \quad V^{\boldsymbol{q}}_{(\mathbf{0}^k,\boldsymbol{s}),(\mathbf{1}^n-\boldsymbol{j},\mathbf{0}^{m-n})} = \tilde{V}^{\boldsymbol{q},\boldsymbol{j}}_{\boldsymbol{s},\mathbf{1}^{n-r}}. \tag{3.41}$$

With the permanent composition formula in Eq. (3.23) we obtain

$$\sum_{\boldsymbol{s}\in\Phi_{m-k,n-r}} \frac{1}{\boldsymbol{s}!}\operatorname{Per}\left(U^{\boldsymbol{p}\dagger}_{(\mathbf{1}^n-\boldsymbol{i},\mathbf{0}^{m-n}),(\mathbf{0}^k,\boldsymbol{s})}\right)\operatorname{Per}\left(V^{\boldsymbol{q}}_{(\mathbf{0}^k,\boldsymbol{s}),(\mathbf{1}^n-\boldsymbol{j},\mathbf{0}^{m-n})}\right) = \operatorname{Per}\left[\left(\tilde{U}^{\boldsymbol{p},\boldsymbol{i}}\tilde{V}^{\boldsymbol{q},\boldsymbol{j}}\right)_{\mathbf{1}^{n-r},\mathbf{1}^{n-r}}\right]. \tag{3.42}$$

Since $\tilde{U}^{\boldsymbol{p},\boldsymbol{i}}\tilde{V}^{\boldsymbol{q},\boldsymbol{j}}$ is an $(n-r)\times(n-r)$ matrix we thus have

$$\sum_{\boldsymbol{s}\in\Phi_{m-k,n-r}} \frac{1}{\boldsymbol{s}!}\operatorname{Per}\left(U^{\boldsymbol{p}\dagger}_{(\mathbf{1}^n-\boldsymbol{i},\mathbf{0}^{m-n}),(\mathbf{0}^k,\boldsymbol{s})}\right)\operatorname{Per}\left(V^{\boldsymbol{q}}_{(\mathbf{0}^k,\boldsymbol{s}),(\mathbf{1}^n-\boldsymbol{j},\mathbf{0}^{m-n})}\right) = \operatorname{Per}\left(\tilde{U}^{\boldsymbol{p},\boldsymbol{i}}\tilde{V}^{\boldsymbol{q},\boldsymbol{j}}\right). \tag{3.43}$$

Then, Eq. (3.38) rewrites

$$\langle\psi_{\boldsymbol{p}}|\psi_{\boldsymbol{q}}\rangle = \frac{1}{\sqrt{\boldsymbol{p}!\boldsymbol{q}!}} \sum_{\substack{\boldsymbol{i},\boldsymbol{j}\in\{0,1\}^n \\ |\boldsymbol{i}|=|\boldsymbol{j}|=r}} \operatorname{Per}\left(A^{\boldsymbol{i}}\right)\operatorname{Per}\left(B^{\boldsymbol{j}}\right)\operatorname{Per}\left(C^{\boldsymbol{i},\boldsymbol{j}}\right), \tag{3.44}$$

where we have defined

$$\begin{aligned} C^{\boldsymbol{i},\boldsymbol{j}} &:= \tilde{U}^{\boldsymbol{p},\boldsymbol{i}}\tilde{V}^{\boldsymbol{q},\boldsymbol{j}} \\ &= U^{\boldsymbol{p}\dagger}_{(\mathbf{1}^n-\boldsymbol{i},\mathbf{0}^{m-n}),(\mathbf{0}^k,\mathbf{1}^{m-k})} V^{\boldsymbol{q}}_{(\mathbf{0}^k,\mathbf{1}^{m-k}),(\mathbf{1}^n-\boldsymbol{j},\mathbf{0}^{m-n})}, \end{aligned} \tag{3.45}$$

is an $(n-r)\times(n-r)$ matrix which can be obtained efficiently from $U^{\boldsymbol{p}}$ and $V^{\boldsymbol{q}}$.

■

By Lemma 3.2, the overlap is expressed as the modulus squared of a sum over $\binom{n}{r}^2$ products of three permanents, of square matrices of sizes $|\boldsymbol{p}| = r$, $|\boldsymbol{q}| = r$ and $(n-r)$, respectively. In the worst case, when $r = n/2$, the sum has at most $O(4^n)$ terms, up to a polynomial factor in n. In particular, when $n = O(\log m)$, the overlap reduces to a sum of a polynomial number of terms, which can all be computed in time $O(\text{poly } m)$. Moreover, the cost of computing the overlap is independent of the number k of adaptive measurements, up to the cost of constructing the matrices with repeated lines and columns (which is $O(\text{poly } m)$). The overlap of normalised ouput states is given by

$$\frac{|\langle\psi_{\boldsymbol{p}}|\psi_{\boldsymbol{q}}\rangle|^2}{\langle\psi_{\boldsymbol{p}}|\psi_{\boldsymbol{p}}\rangle\langle\psi_{\boldsymbol{q}}|\psi_{\boldsymbol{q}}\rangle}, \tag{3.46}$$

which may also be computed efficiently when $n = O(\log m)$. The efficiency of the classical algorithm is summarised as a function of n and k in Table 3.2 and as a function of n and the number of photons r detected during the adaptive measurements in Table 3.3, where the scalings are obtained using Stirling equivalent $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$.

n \ k	$O(1)$	$O(\log m)$	$O(m)$
$O(1)$	$O(\text{poly } m)$	$O(\text{poly } m)$	$O(\text{poly } m)$
$O(\log m)$	$O(\text{poly } m)$	$O(\text{poly } m)$	$O(\text{poly } m)$
$O(m)$	$O(4^m)$	$O(4^m)$	$O(4^m)$

Table 3.2: Simulability regimes for overlap estimation as a function of n and k. The scalings indicated are upper bounds on the running time of the classical algorithm, up to a factor $O(\text{poly } m)$. Since the running time is independent of k, the columns are the same. In blue is the parameter region for which the quantum algorithm is efficient.

n \ r	$O(1)$	$O(\log n)$	$O(n)$
$O(1)$	$O(\text{poly } m)$	$O(\text{poly } m)$	$O(\text{poly } m)$
$O(\log m)$	$O(\text{poly } m)$	$O(\text{poly } m)$	$O(\text{poly } m)$
$O(m)$	$O(\text{poly } m)$	$O(4^{\log^2 m})$	$O(4^m)$

Table 3.3: Simulability regimes for overlap estimation as a function of n and r. The scalings indicated are upper bounds on the running time of the classical algorithm, up to a factor $O(\text{poly } m)$.

Since the quantum efficient regime corresponds to $\binom{k+n}{n} = O(\text{poly } m)$ and the time complexity of the classical algorithm is $O(4^n)$, up to $O(\text{poly } m)$ factors, there is a possibility of quantum advantage for overlap estimation when $k = O(1)$ and $n = O(m)$.
In the case of probability estimation, the possible regimes for quantum advantage do not correspond to near-term implementations: k and n must be both greater than $\log m$. However, for overlap estimation, there is a possiblity of near-term beyond-classical computing with adaptive linear optics using one adaptive measurement, which requires the preparation of photon number states. Note that the interferometer should be concentrating many photons r onto the adaptive measurement in order to obtain possibly hard to estimate overlaps. Using more adaptive measurements does not increase the complexity (apart from polynomial factors in m).

Having characterised these specific simulation regimes, we consider in what follows stronger notions of simulation. In particular, we give classical algorithms for strong simulation of a large class of continuous variable quantum computational models.

3.3 The computational power of non-Gaussian states

Continuous variable systems are being recognized as a promising alternative to the use of qubits, as they allow for the deterministic generation of unprecedented large entangled quantum

states, of up to one-million elementary systems [YUA+13a, YYK+16] and also offer detection techniques, such as homodyne and heterodyne, with high efficiency and reliability (see section 1.4.2). Any given continuous variable quantum circuit is defined by (i) an input state lying in an infinite-dimensional Hilbert space, (ii) an evolution and (iii) measurements (see scetion 1.1.3). An important theorem [BSBN02, ME12] states that if all these elements are described by positive Wigner functions, then there exists a classical algorithm able to efficiently simulate this circuit. Hence, including a negative Wigner function element is mandatory in order to design a continuous variable subuniversal quantum circuit that cannot be efficiently simulated by a classical device. Since Gaussian states and processes have positive Wigner functions, this necessarily corresponds to the use of non-Gaussian resources.

Therefore, if one aims at minimal extensions of Gaussian models, three different families of non trivial quantum circuits can be defined, depending on whether the element yielding the Wigner function negativity is provided by the input state, the unitary evolution, or the measurement.

In what follows, we analyse the computational power of non-Gaussian states and thus focus on the case where Gaussian circuits and measurements are supplemented with non-Gaussian input states as a computational resource. The results obtained have consequences for all three families of circuits, since non-Gaussian gates and non-Gaussian measurements can be implemented by Gaussian operations together with non-Gaussian ancillary states [GKP01, GS07, SW18].

3.3.1 Gaussian circuits with non-Gaussian inputs

We first extend a few definitions from the previous chapter to the multimode case, using multi-index notations (see section 1.1.1). First, the stellar function, which provides a representation of multimode pure states as multivariate holomorphic functions:

Definition 3.5 (Multimode stellar function)**.** Let $m \in \mathbb{N}^*$ and let $|\psi\rangle = \sum_{\boldsymbol{n}\geq \boldsymbol{0}} \psi_{\boldsymbol{n}} |\boldsymbol{n}\rangle \in \mathscr{H}^{\otimes m}$ be a normalised pure state over m modes. The *stellar function* of the state $|\psi\rangle$ is defined as

$$F^{\star}_{\psi}(\boldsymbol{z}) = e^{\frac{1}{2}\|\boldsymbol{z}\|^2} \langle \boldsymbol{z}^* | \psi \rangle = \sum_{\boldsymbol{n}\geq \boldsymbol{0}} \frac{\psi_{\boldsymbol{n}}}{\sqrt{\boldsymbol{n}!}} \boldsymbol{z}^{\boldsymbol{n}}, \tag{3.47}$$

for all $\boldsymbol{z} \in \mathbb{C}^m$, where $|\boldsymbol{z}\rangle = e^{-\frac{1}{2}\|\boldsymbol{z}\|^2} \sum_{\boldsymbol{n}\geq \boldsymbol{0}} \frac{\boldsymbol{z}^{\boldsymbol{n}}}{\sqrt{\boldsymbol{n}!}} |\boldsymbol{n}\rangle \in \mathscr{H}^{\otimes m}$ is the coherent state of amplitude $\boldsymbol{z}$.

The following definition also extends naturally from the single-mode case:

Definition 3.6 (Multimode core state)**.** *Multimode core states* are defined as the normalised pure quantum states which have a (multivariate) polynomial stellar function.

Like in the single-mode case, these are the states with a finite support over the (multimode) Fock basis. For any $m \in \mathbb{N}^*$, the set of multimode core states over m modes is dense in the set of normalised states for the trace norm (by considering renormalised cutoff states). We also introduce the following definitions:

Definition 3.7 (Degree of a multimode core state)**.** The *degree* of a multimode core state is defined as the degree-sum of its stellar function.

Definition 3.8 (Support of a multimode core state)**.** The *support* of a multimode core state is the set of Fock basis elements which have nonzero overlap with the core state.

For example, the 3-mode core state $\frac{1}{\sqrt{2}}(|210\rangle + |001\rangle)$ is of degree 3 and has a support of size 2, and its stellar function is given by $z_1^2 z_2/2 + z_3/\sqrt{2}$, for all $(z_1, z_2, z_3) \in \mathbb{C}^3$.

We consider Gaussian circuits with Gaussian measurements, supplemented by non-Gaussian multimode core states in input, which we refer to as G_{core} circuits. Without loss of generality, a Gaussian measurement may be written as a tensor product of single-mode balanced heterodyne detections preceded by a Gaussian unitary (see section 1.4.2). G_{core} circuits are thus described by two (multidimensional) parameters: a multimode core state $|\boldsymbol{C}\rangle$ in the input and a Gaussian unitary evolution $\hat{G}$ (Fig. 3.2).

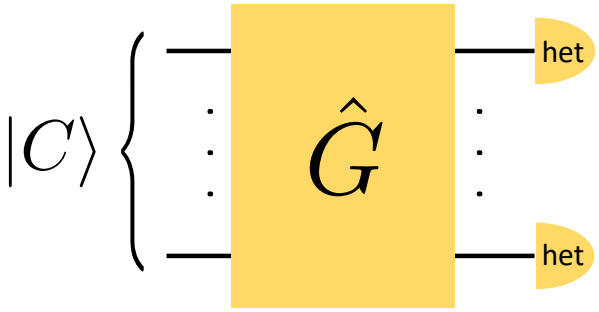

Figure 3.2: Representation of a G_{core} circuit with multimode core state input $|\boldsymbol{C}\rangle$. The unitary $\hat{G}$ is Gaussian and the measurement is performed by heterodyne detection.

In what follows, we derive a general expression for the output probability density of these circuits. Then, we study the classical simulability of G_{core} circuits and of various subclasses of circuits.

We first recall a few combinatorial functions related to the permanent, which appear in the expressions of the output probability densities. The hafnian of a square matrix $A = (a_{ij})_{1 \leq i,j \leq 2m}$ of size $2m$ is defined as [Cai53]

$$\text{Haf}(A) := \sum_{M \in \text{PMP}(2m)} \prod_{\{i,j\} \in M} a_{ij}, \tag{3.48}$$

where the sum is over the perfect matchings of the set $\{1, \ldots, 2m\}$, i.e., the partitions of $\{1, \ldots, 2m\}$ in subsets of size 2. The hafnian of a matrix of odd size is 0. The hafnian is related to the permanent by

$$\text{Haf}\begin{pmatrix} \mathbb{0}_m & B \\ B^T & \mathbb{0}_m \end{pmatrix} = \text{Per}(B), \tag{3.49}$$

for any $m \times m$ matrix B. By convention we set $\mathrm{Haf}(\emptyset) = 1$, where $\emptyset$ is a square matrix of size 0.

The loop hafnian of a square matrix $R = (r_{ij})_{1 \leq i,j \leq r}$ of size r is defined as [BGQ19]

$$\mathrm{lHaf}(R) := \sum_{M \in \mathrm{SMP}(r)} \prod_{\{i,j\} \in M} r_{ij}, \tag{3.50}$$

where the sum is over the single pair matchings of the set $\{1,\dots,r\}$, defined as the set of perfect matchings of a complete graph with loops with r vertices. This set is isomorphic to the set $\Pi_{1,2}(\{1,\dots,r\})$ of partitions of $\{1,\dots,r\}$ in subsets of size 1 and 2 (by mapping a block $\{k\}$ of size 1 of a partition to the matching $\{k,k\}$ and a block $\{i,j\}$ of size 2 to the matching $\{i,j\}$). In particular, when R is a matrix whose diagonal entries are all 0, we have $\mathrm{lHaf}(R) = \mathrm{Haf}(R)$.

We obtain a closed expression for the output probability density of Gaussian circuits with multimode core states input in Theorem 3.1, by adapting proof techniques from [HKS+16, KHS+19, Que19]. We first state an intermediate technical result.

Lemma 3.3. *Let $m \in \mathbb{N}^*$, let V be a $2m \times 2m$ symmetric matrix and let D be a column vector of size $2m$. For all $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{N}^m$, there exists a square matrix $A_{\boldsymbol{p},\boldsymbol{q}}(V,D)$ of size $|\boldsymbol{p}| + |\boldsymbol{q}|$ such that*

$$\begin{aligned} T_{\boldsymbol{p},\boldsymbol{q}}(V,D) &:= \int_{\boldsymbol{\beta} \in \mathbb{C}^m} \exp\left[\frac{1}{2}\tilde{\boldsymbol{\beta}}^T V \tilde{\boldsymbol{\beta}} + D^T \tilde{\boldsymbol{\beta}}\right] \left(\frac{\partial}{\partial \boldsymbol{\beta}}\right)^{\boldsymbol{p}} \left(\frac{\partial}{\partial \boldsymbol{\beta}^*}\right)^{\boldsymbol{q}} \delta^{2m}(\boldsymbol{\beta}, \boldsymbol{\beta}^*) d^m\boldsymbol{\beta}\, d^m\boldsymbol{\beta}^* \\ &= (-1)^{|\boldsymbol{p}|+|\boldsymbol{q}|}\, \mathrm{lHaf}\left[A_{\boldsymbol{p},\boldsymbol{q}}(V,D)\right], \end{aligned} \tag{3.51}$$

assuming the integral is well defined. The matrix $A_{\boldsymbol{p},\boldsymbol{q}}(V,D)$ is obtained by repeating the entries of V according to $\boldsymbol{p}$ and $\boldsymbol{q}$ and replacing the diagonal of the matrix obtained by the corresponding elements of D (a detailed example follows the proof).

Proof. Writing $\boldsymbol{p} = (p_1,\dots,p_m)$ and $\boldsymbol{q} = (q_1,\dots,q_m)$, we first get rid of the integral by successive integration by parts:

$$\begin{aligned} T_{\boldsymbol{p},\boldsymbol{q}}(V,D) &= (-1)^{|\boldsymbol{p}|+|\boldsymbol{q}|} \left(\frac{\partial}{\partial \boldsymbol{\beta}}\right)^{\boldsymbol{p}} \left(\frac{\partial}{\partial \boldsymbol{\beta}^*}\right)^{\boldsymbol{q}} \exp\left[\frac{1}{2}\tilde{\boldsymbol{\beta}}^T V \tilde{\boldsymbol{\beta}} + D^T \tilde{\boldsymbol{\beta}}\right]\Big|_{\tilde{\boldsymbol{\beta}}=\boldsymbol{0}} \\ &= (-1)^{|\boldsymbol{p}|+|\boldsymbol{q}|} \prod_{j=1}^{m} \left(\frac{\partial}{\partial \beta_j}\right)^{p_j} \left(\frac{\partial}{\partial \beta_j^*}\right)^{q_j} \exp\left[\frac{1}{2}\tilde{\boldsymbol{\beta}}^T V \tilde{\boldsymbol{\beta}} + D^T \tilde{\boldsymbol{\beta}}\right]\Big|_{\tilde{\boldsymbol{\beta}}=\boldsymbol{0}} \\ &= (-1)^{|\boldsymbol{p}|+|\boldsymbol{q}|} \prod_{j \in \mathscr{E}_{\boldsymbol{p},\boldsymbol{q}}} \left(\frac{\partial}{\partial \tilde{\beta}_j}\right) \exp\left[\frac{1}{2}\tilde{\boldsymbol{\beta}}^T V \tilde{\boldsymbol{\beta}} + D^T \tilde{\boldsymbol{\beta}}\right]\Big|_{\tilde{\boldsymbol{\beta}}=\boldsymbol{0}}, \end{aligned} \tag{3.52}$$

where the multiset $\mathscr{E}_{\boldsymbol{p},\boldsymbol{q}}$ is defined as the set of size $|\boldsymbol{p}| + |\boldsymbol{q}|$ obtained from $\{1,\dots,2m\}$ by repeating p_k times the index k and q_k times the index $m+k$, for all $k \in \{1,\dots,m\}$.

We make use of Faà di Bruno's formula [Har06] in order to expand the product of partial derivatives and we obtain

$$T_{\boldsymbol{p},\boldsymbol{q}}(V,D) = (-1)^{|\boldsymbol{p}|+|\boldsymbol{q}|} \sum_{\pi \in \Pi(\mathscr{E}_{\boldsymbol{p},\boldsymbol{q}})} \prod_{B \in \pi} \left(\frac{\partial^{|B|}}{\prod_{j \in B} \partial \tilde{\beta}_j}\right) \left[\frac{1}{2}\tilde{\boldsymbol{\beta}}^T V \tilde{\boldsymbol{\beta}} + D^T \tilde{\boldsymbol{\beta}}\right]\Big|_{\tilde{\boldsymbol{\beta}}=\boldsymbol{0}}, \tag{3.53}$$

where $\Pi(\mathscr{E}_{\boldsymbol{p},\boldsymbol{q}})$ denotes the set of all partitions of the multiset $\mathscr{E}_{\boldsymbol{p},\boldsymbol{q}}$, and where the product runs over the blocks B of the partition $\pi \in \Pi(\mathscr{E}_{\boldsymbol{p},\boldsymbol{q}})$, with $|B|$ the size of the block. The function $\tilde{\boldsymbol{\beta}}^\dagger V \tilde{\boldsymbol{\beta}} + D^\dagger \tilde{\boldsymbol{\beta}}$ is a sum of a quadratic and a linear functions, so all derivatives of order greater than 2 in the sum vanish. We thus have

$$\begin{aligned} T_{\boldsymbol{p},\boldsymbol{q}}(V,D) &= (-1)^{|\boldsymbol{p}|+|\boldsymbol{q}|} \sum_{\pi \in \Pi_{1,2}(\mathscr{E}_{\boldsymbol{p},\boldsymbol{q}})} \prod_{B \in \pi} \left(\frac{\partial^{|B|}}{\prod_{j \in B} \partial \tilde{\beta}_j} \right) \left[\frac{1}{2} \tilde{\boldsymbol{\beta}}^T V \tilde{\boldsymbol{\beta}} + D^T \tilde{\boldsymbol{\beta}} \right] \Big|_{\tilde{\boldsymbol{\beta}}=\mathbf{0}} \\ &= (-1)^{|\boldsymbol{p}|+|\boldsymbol{q}|} \sum_{\pi \in \Pi_{1,2}(\mathscr{E}_{\boldsymbol{p},\boldsymbol{q}})} \prod_{\{i,j\} \in \pi} \left(\frac{\partial^2}{\partial \tilde{\beta}_i \partial \tilde{\beta}_j} \right) \left[\frac{1}{2} \tilde{\boldsymbol{\beta}}^T V \tilde{\boldsymbol{\beta}} + D^T \tilde{\boldsymbol{\beta}} \right] \Big|_{\tilde{\boldsymbol{\beta}}=\mathbf{0}} \\ &\qquad \times \prod_{\{k\} \in \pi} \left(\frac{\partial}{\partial \tilde{\beta}_k} \right) \left[\frac{1}{2} \tilde{\boldsymbol{\beta}}^T V \tilde{\boldsymbol{\beta}} + D^T \tilde{\boldsymbol{\beta}} \right] \Big|_{\tilde{\boldsymbol{\beta}}=\mathbf{0}}, \end{aligned} \tag{3.54}$$

where $\Pi_{1,2}(\mathscr{E}_{\boldsymbol{p},\boldsymbol{q}})$ denotes the set of all partitions of the multiset $\mathscr{E}_{\boldsymbol{p},\boldsymbol{q}}$ in subsets of size 1 and 2. All derivatives of order 2 of the linear term vanish, and all derivatives of order 1 of the quadratic term vanish when evaluated at $\tilde{\boldsymbol{\beta}} = \mathbf{0}$. We thus obtain

$$T_{\boldsymbol{p},\boldsymbol{q}}(V,D) = (-1)^{|\boldsymbol{p}|+|\boldsymbol{q}|} \sum_{\pi \in \Pi_{1,2}(\mathscr{E}_{\boldsymbol{p},\boldsymbol{q}})} \prod_{\{i,j\} \in \pi} \left(\frac{\partial^2}{\partial \tilde{\beta}_i \partial \tilde{\beta}_j} \right) \left[\frac{1}{2} \tilde{\boldsymbol{\beta}}^T V \tilde{\boldsymbol{\beta}} \right] \Big|_{\tilde{\boldsymbol{\beta}}=\mathbf{0}} \prod_{\{k\} \in \pi} \left(\frac{\partial}{\partial \tilde{\beta}_k} \right) \left[D^T \tilde{\boldsymbol{\beta}} \right] \Big|_{\tilde{\boldsymbol{\beta}}=\mathbf{0}}. \tag{3.55}$$

Writing $V = (v_{ij})_{1 \le i,j \le 2m}$, with $V = V^T$, and $D = (d_k)_{1 \le k \le 2m}$ we obtain

$$T_{\boldsymbol{p},\boldsymbol{q}}(V,D) = (-1)^{|\boldsymbol{p}|+|\boldsymbol{q}|} \sum_{\pi \in \Pi_{1,2}(\mathscr{E}_{\boldsymbol{p},\boldsymbol{q}})} \prod_{\{i,j\} \in \pi} v_{ij} \prod_{\{k\} \in \pi} d_k. \tag{3.56}$$

We now show that this expression may be rewritten as the loop hafnian of a matrix of size $|\boldsymbol{p}| + |\boldsymbol{q}|$. Define $V_{\boldsymbol{p},\boldsymbol{q}}$ the $(|\boldsymbol{p}| + |\boldsymbol{q}|) \times (|\boldsymbol{p}| + |\boldsymbol{q}|)$ matrix obtained from V by repeating p_k times its k^{th} rows and columns and q_k times its $(m+k)^{th}$ rows and columns, for $k \in \{1, \dots, m\}$. Similarly, define $D_{\boldsymbol{p},\boldsymbol{q}}$ the column vector of size $|\boldsymbol{p}| + |\boldsymbol{q}|$ obtained from D by repeating p_k times its k^{th} element and q_k times its $(m+k)^{th}$ element, for $k \in \{1, \dots, m\}$. Finally, let $A_{\boldsymbol{p},\boldsymbol{q}}(V,D) = (a_{ij})_{1 \le i,j \le |\boldsymbol{p}|+|\boldsymbol{q}|}$ be the $(|\boldsymbol{p}| + |\boldsymbol{q}|) \times (|\boldsymbol{p}| + |\boldsymbol{q}|)$ matrix obtained from $V_{\boldsymbol{p},\boldsymbol{q}}$ by replacing its diagonal with the vector $D_{\boldsymbol{p},\boldsymbol{q}}$. Then, Eq. (3.56) rewrites

$$\begin{aligned} T_{\boldsymbol{p},\boldsymbol{q}}(V,D) &= (-1)^{|\boldsymbol{p}|+|\boldsymbol{q}|} \sum_{\pi \in \Pi_{1,2}(\{1,\dots,|\boldsymbol{p}|+|\boldsymbol{q}|\})} \prod_{\{i,j\} \in \pi} a_{ij} \prod_{\{k\} \in \pi} a_{kk} \\ &= (-1)^{|\boldsymbol{p}|+|\boldsymbol{q}|} \sum_{M \in \mathrm{SMP}(|\boldsymbol{p}|+|\boldsymbol{q}|)} \prod_{\{i,j\} \in M} a_{ij} \\ &= (-1)^{|\boldsymbol{p}|+|\boldsymbol{q}|} \mathrm{lHaf}\left[A_{\boldsymbol{p},\boldsymbol{q}}(V,D) \right], \end{aligned} \tag{3.57}$$

where the sum in the first line is over the partitions of $\{1, \dots, |\boldsymbol{p}| + |\boldsymbol{q}|\}$ in subsets of size 1 and 2, where the sum in the second line is over the single pair matchings of the set $\{1, \dots, |\boldsymbol{p}| + |\boldsymbol{q}|\}$ and where the third line comes from the definition of the loop hafnian in Eq. (3.50).

■

Let us illustrate with an example how the matrix $A_{\boldsymbol{p},\boldsymbol{q}}(V,D)$ appearing in Lemma 3.3 is constructed from the matrix V and the vector D. Let us set $m = 2$, $\boldsymbol{p} = (2,0)$ and $\boldsymbol{q} = (1,0)$. We write

$$V = \begin{pmatrix} v_{11} & v_{12} & v_{13} & v_{14} \\ v_{21} & v_{22} & v_{23} & v_{24} \\ v_{31} & v_{32} & v_{33} & v_{34} \\ v_{41} & v_{42} & v_{43} & v_{44} \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{pmatrix}. \tag{3.58}$$

We first build the matrix $V_{\boldsymbol{p},\boldsymbol{q}}$ by repeating p_k times the k^{th} rows and columns of V and q_k times the $(m+k)^{th}$ rows and columns. In that case, $\boldsymbol{p} = (p_1,p_2) = (2,0)$, so we repeat 2 times the first row and column and discard the second row and column, and $\boldsymbol{q} = (q_1,q_2) = (1,0)$, so we keep the third row and column and discard the fourth row and column, obtaining the 3×3 matrix

$$V_{\boldsymbol{p},\boldsymbol{q}} = \begin{pmatrix} v_{11} & v_{11} & v_{13} \\ v_{11} & v_{11} & v_{13} \\ v_{31} & v_{31} & v_{33} \end{pmatrix}. \tag{3.59}$$

Similarly, we obtain the vector $D_{\boldsymbol{p},\boldsymbol{q}}$ by repeating p_k times the k^{th} element of D and q_k times the $(m+k)^{th}$ element, as

$$D_{\boldsymbol{p},\boldsymbol{q}} = \begin{pmatrix} d_1 \\ d_1 \\ d_3 \end{pmatrix}. \tag{3.60}$$

Finally, we replace the diagonal of $V_{\boldsymbol{p},\boldsymbol{q}}$ by $D_{\boldsymbol{p},\boldsymbol{q}}$:

$$A_{\boldsymbol{p},\boldsymbol{q}}(V,D) = \begin{pmatrix} d_1 & v_{11} & v_{13} \\ v_{11} & d_1 & v_{13} \\ v_{31} & v_{31} & d_3 \end{pmatrix}. \tag{3.61}$$

Note that this construction by repeating rows and columns differ from the one encountered in the previous section when dealing with the permanent of matrices, for which the first index denotes which rows are repeated and the second which columns. Here, we are dealing with hafnians of matrices of double size, where the first index denotes which rows and columns are repeated for indices in $\{1,\dots,m\}$, while the second index denotes which rows and columns are repeated for indices in $\{m+1,\dots,2m\}$. However, the two constructions coincide when looking at matrices of the form

$$\begin{pmatrix} \mathbb{0}_m & B \\ B^T & \mathbb{0}_m \end{pmatrix}, \tag{3.62}$$

through the relation in Eq. (3.49):

$$\mathrm{Haf}\begin{pmatrix} \mathbb{0}_m & B \\ B^T & \mathbb{0}_m \end{pmatrix} = \mathrm{Per}(B), \tag{3.63}$$

for any $m\times m$ square matrix B.

Combining Lemma 3.3 with phase space formalism (see section 1.2) and properties of Gaussian states (see section 1.3), we obtain the following result:

Theorem 3.1. *Let $m,n \in \mathbb{N}^*$ and let*

$$|\boldsymbol{C}\rangle = \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^m \\ |\boldsymbol{p}|\leq n}} c_{\boldsymbol{p}} |\boldsymbol{p}\rangle, \tag{3.64}$$

be an m-mode core state of degree n. Let $\hat{G}$ be a Gaussian unitary over m modes. For all $\boldsymbol{\alpha} \in \mathbb{C}^m$, let us write $\boldsymbol{V}$ and $\tilde{\boldsymbol{d}} = (\boldsymbol{d}, \boldsymbol{d}^*)$ the covariance matrix and the displacement vector of the Gaussian state $\hat{G}^\dagger |\boldsymbol{\alpha}\rangle$. Then, the output probability density for the G_{core} circuit $\hat{G}$ with input $|\boldsymbol{C}\rangle$ and heterodyne detection, evaluated at $\boldsymbol{\alpha}$, is given by

$$Pr_{core}[\boldsymbol{\alpha}] = \kappa(\boldsymbol{\alpha}, \hat{G}) \sum_{\substack{\boldsymbol{p},\boldsymbol{q}\in\mathbb{N}^m \\ |\boldsymbol{p}|\leq n, |\boldsymbol{q}|\leq n}} \frac{(-1)^{|\boldsymbol{p}|+|\boldsymbol{q}|}}{\sqrt{\boldsymbol{p}!\boldsymbol{q}!}} c_{\boldsymbol{p}} c_{\boldsymbol{q}}^* \, \mathrm{lHaf}\left(A_{\boldsymbol{p},\boldsymbol{q}}\right), \tag{3.65}$$

where $A_{\boldsymbol{p},\boldsymbol{q}}$ is the square matrix of size $|\boldsymbol{p}|+|\boldsymbol{q}|$ obtained with Lemma 3.3 from

$$V = \begin{pmatrix} \mathbb{0}_m & \mathbb{1}_m \\ \mathbb{1}_m & \mathbb{0}_m \end{pmatrix} \left[\mathbb{1}_{2m} - (\boldsymbol{V} + \mathbb{1}_{2m}/2)^{-1}\right] \quad \textit{and} \quad D = \left[\tilde{\boldsymbol{d}}^\dagger (\boldsymbol{V} + \mathbb{1}_{2m}/2)^{-1}\right]^T, \tag{3.66}$$

and where

$$\kappa(\boldsymbol{\alpha}, \hat{G}) = \frac{\exp\left[-\frac{1}{2}\tilde{\boldsymbol{d}}^\dagger (\boldsymbol{V} + \mathbb{1}_{2m}/2)^{-1} \tilde{\boldsymbol{d}}\right]}{\pi^m \sqrt{\mathrm{Det}(\boldsymbol{V} + \mathbb{1}_{2m}/2)}} \tag{3.67}$$

is a Gaussian prefactor.

Proof. The Gaussian circuit is composed of a Gaussian unitary $\hat{G}$ and balanced heterodyne detection. The output probability density reads, for all $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m) \in \mathbb{C}^m$,

$$\begin{aligned} \mathrm{Pr}_{\mathrm{core}}[\boldsymbol{\alpha}] &= \mathrm{Tr}\left[\hat{G}|\boldsymbol{C}\rangle\langle\boldsymbol{C}|\hat{G}^\dagger \Pi_{\boldsymbol{\alpha}}\right] \\ &= \frac{1}{\pi^m} \mathrm{Tr}\left[\hat{G}^\dagger |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}| \hat{G} |\boldsymbol{C}\rangle\langle\boldsymbol{C}|\right] \\ &= \int_{\boldsymbol{\beta}\in\mathbb{C}^m} Q_{\hat{G}^\dagger|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|\hat{G}}(\boldsymbol{\beta}) P_{|\boldsymbol{C}\rangle\langle\boldsymbol{C}|}(\boldsymbol{\beta}) d^m\boldsymbol{\beta}\, d^m\boldsymbol{\beta}^*, \end{aligned} \tag{3.68}$$

where $\Pi_{\boldsymbol{\alpha}} = \frac{1}{\pi^m} |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|$ is the POVM element corresponding to the heterodyne detection of $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)$. The state $\hat{G}^\dagger |\boldsymbol{\alpha}\rangle$ is a Gaussian state: let $\boldsymbol{V}$ be its covariance matrix and $\boldsymbol{d}$ its displacement vector. For all $\boldsymbol{\gamma} \in \mathbb{C}^m$, we write $\tilde{\boldsymbol{\gamma}} = (\gamma_1, \ldots, \gamma_m, \gamma_1^*, \ldots, \gamma_m^*)$. Then, for all $\boldsymbol{\beta} \in \mathbb{C}^m$,

$$\begin{aligned} Q_{\hat{G}^\dagger|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|\hat{G}}(\boldsymbol{\beta}) &= \frac{1}{\pi^m \sqrt{\mathrm{Det}(\boldsymbol{V} + \mathbb{1}_{2m}/2)}} \exp\left[-\frac{1}{2}(\tilde{\boldsymbol{\beta}} - \tilde{\boldsymbol{d}})^\dagger (\boldsymbol{V} + \mathbb{1}_{2m}/2)^{-1} (\tilde{\boldsymbol{\beta}} - \tilde{\boldsymbol{d}})\right] \\ &= \frac{\exp\left[-\frac{1}{2}\tilde{\boldsymbol{d}}^\dagger (\boldsymbol{V} + \mathbb{1}_{2m}/2)^{-1} \tilde{\boldsymbol{d}}\right]}{\pi^m \sqrt{\mathrm{Det}(\boldsymbol{V} + \mathbb{1}_{2m}/2)}} \exp\left[-\frac{1}{2}\tilde{\boldsymbol{\beta}}^\dagger (\boldsymbol{V} + \mathbb{1}_{2m}/2)^{-1} \tilde{\boldsymbol{\beta}} + \tilde{\boldsymbol{d}}^\dagger (\boldsymbol{V} + \mathbb{1}_{2m}/2)^{-1} \tilde{\boldsymbol{\beta}}\right], \end{aligned} \tag{3.69}$$

i.e., it is a Gaussian function which can be computed efficiently. On the other hand, we have

$$|\boldsymbol{C}\rangle\langle\boldsymbol{C}| = \sum_{\substack{\boldsymbol{p},\boldsymbol{q}\in\mathbb{N}^m \\ |\boldsymbol{p}|\leq n, |\boldsymbol{q}|\leq n}} c_{\boldsymbol{p}} c_{\boldsymbol{q}}^* \, |\boldsymbol{p}\rangle\langle\boldsymbol{q}|, \tag{3.70}$$

so that

$$P_{|\boldsymbol{C}\rangle\langle\boldsymbol{C}|}(\boldsymbol{\beta}) = \sum_{\substack{\boldsymbol{p},\boldsymbol{q}\in\mathbb{N}^m \\ |\boldsymbol{p}|\leq n, |\boldsymbol{q}|\leq n}} c_{\boldsymbol{p}} c_{\boldsymbol{q}}^* P_{|\boldsymbol{p}\rangle\langle\boldsymbol{q}|}(\boldsymbol{\beta}), \tag{3.71}$$

for all $\boldsymbol{\beta} \in \mathbb{C}^m$. Moreover we have, for all $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{N}^m$ and all $\boldsymbol{\beta} \in \mathbb{C}^m$,

$$\begin{aligned} P_{|\boldsymbol{p}\rangle\langle\boldsymbol{q}|}(\boldsymbol{\beta}) &= \frac{e^{\|\boldsymbol{\beta}\|^2}}{\sqrt{\boldsymbol{p}!\boldsymbol{q}!}} \left(\frac{\partial}{\partial \boldsymbol{\beta}}\right)^{\boldsymbol{p}} \left(\frac{\partial}{\partial \boldsymbol{\beta}^*}\right)^{\boldsymbol{q}} \delta^{2m}(\boldsymbol{\beta}, \boldsymbol{\beta}^*) \\ &= \frac{e^{\frac{1}{2}\tilde{\boldsymbol{\beta}}^\dagger \tilde{\boldsymbol{\beta}}}}{\sqrt{\boldsymbol{p}!\boldsymbol{q}!}} \left(\frac{\partial}{\partial \boldsymbol{\beta}}\right)^{\boldsymbol{p}} \left(\frac{\partial}{\partial \boldsymbol{\beta}^*}\right)^{\boldsymbol{q}} \delta^{2m}(\boldsymbol{\beta}, \boldsymbol{\beta}^*), \end{aligned} \tag{3.72}$$

where $\delta^{2m}(\boldsymbol{\beta}, \boldsymbol{\beta}^*) = \delta(\beta_1)\cdots\delta(\beta_m)\delta(\beta_1^*)\cdots\delta(\beta_m^*)$. Combining Eqs. (3.69), (3.71) and (3.72) with Eq. (3.68) we obtain

$$\begin{aligned} \Pr_{\text{core}}[\boldsymbol{\alpha}] = \kappa(\boldsymbol{\alpha}, \hat{G}) \sum_{\substack{\boldsymbol{p},\boldsymbol{q}\in\mathbb{N}^m \\ |\boldsymbol{p}|\leq n, |\boldsymbol{q}|\leq n}} \frac{c_{\boldsymbol{p}} c_{\boldsymbol{q}}^*}{\sqrt{\boldsymbol{p}!\boldsymbol{q}!}} \int_{\boldsymbol{\beta}\in\mathbb{C}^m} \Bigg\{ \exp\left[-\frac{1}{2}\tilde{\boldsymbol{\beta}}^\dagger (\boldsymbol{V} + \mathbb{1}_{2m}/2)^{-1} \tilde{\boldsymbol{\beta}}\right] \\ \times \exp\left[\tilde{\boldsymbol{d}}^\dagger (\boldsymbol{V} + \mathbb{1}_{2m}/2)^{-1} \tilde{\boldsymbol{\beta}}\right] e^{\frac{1}{2}\tilde{\boldsymbol{\beta}}^\dagger \tilde{\boldsymbol{\beta}}} \left(\frac{\partial}{\partial \boldsymbol{\beta}}\right)^{\boldsymbol{p}} \left(\frac{\partial}{\partial \boldsymbol{\beta}^*}\right)^{\boldsymbol{q}} \delta^{2m}(\boldsymbol{\beta}, \boldsymbol{\beta}^*) \Bigg\} d^m\boldsymbol{\beta} \, d^m\boldsymbol{\beta}^*, \end{aligned} \tag{3.73}$$

where we have set

$$\kappa(\boldsymbol{\alpha}, \hat{G}) = \frac{\exp\left[-\frac{1}{2}\tilde{\boldsymbol{d}}^\dagger (\boldsymbol{V} + \mathbb{1}_{2m}/2)^{-1} \tilde{\boldsymbol{d}}\right]}{\pi^m \sqrt{\text{Det}(\boldsymbol{V} + \mathbb{1}_{2m}/2)}}. \tag{3.74}$$

Given that

$$\tilde{\boldsymbol{\beta}}^\dagger = \tilde{\boldsymbol{\beta}}^T \begin{pmatrix} \mathbb{0}_m & \mathbb{1}_m \\ \mathbb{1}_m & \mathbb{0}_m \end{pmatrix}, \tag{3.75}$$

for all $\boldsymbol{\beta} \in \mathbb{C}^m$, the integral terms in Eq. (3.73) rewrite as

$$\int_{\boldsymbol{\beta}\in\mathbb{C}^m} \exp\left[\frac{1}{2}\tilde{\boldsymbol{\beta}}^T V \tilde{\boldsymbol{\beta}} + D^T \tilde{\boldsymbol{\beta}}\right] \left(\frac{\partial}{\partial \boldsymbol{\beta}}\right)^{\boldsymbol{p}} \left(\frac{\partial}{\partial \boldsymbol{\beta}^*}\right)^{\boldsymbol{q}} \delta^{2m}(\boldsymbol{\beta}, \boldsymbol{\beta}^*) d^m\boldsymbol{\beta} \, d^m\boldsymbol{\beta}^*, \tag{3.76}$$

for $|\boldsymbol{p}| \leq n$ and $|\boldsymbol{q}| \leq n$, where

$$V = \begin{pmatrix} \mathbb{0}_m & \mathbb{1}_m \\ \mathbb{1}_m & \mathbb{0}_m \end{pmatrix} \left[\mathbb{1}_{2m} - (\boldsymbol{V} + \mathbb{1}_{2m}/2)^{-1}\right] \tag{3.77}$$

is a $2m \times 2m$ symmetric matrix, due to the initial structure of the covariance matrix, and where

$$D = \left[\tilde{\boldsymbol{d}}^\dagger (\boldsymbol{V} + \mathbb{1}_{2m}/2)^{-1}\right]^T \tag{3.78}$$

is a column vector of size $2m$. By Lemma 3.3, the terms in Eq. (3.76) are equal to

$$(-1)^{|\boldsymbol{p}|+|\boldsymbol{q}|}\,\mathrm{lHaf}\left(A_{\boldsymbol{p},\boldsymbol{q}}\right), \tag{3.79}$$

where the square matrices $A_{\boldsymbol{p},\boldsymbol{q}}$ of size $|\boldsymbol{p}|+|\boldsymbol{q}|$ are obtained from V by repeating its entries according to $\boldsymbol{p}$ and $\boldsymbol{q}$ and replacing the diagonal by the corresponding elements of D (see the example following Lemma 3.3 for a detailed description of the construction). With Eq. (3.73) we finally obtain

$$\mathrm{Pr}_{\mathrm{core}}[\boldsymbol{\alpha}] = \kappa(\boldsymbol{\alpha},\hat{G}) \sum_{\substack{\boldsymbol{p},\boldsymbol{q}\in\mathbb{N}^m \\ |\boldsymbol{p}|\leq n,|\boldsymbol{q}|\leq n}} \frac{(-1)^{|\boldsymbol{p}|+|\boldsymbol{q}|}}{\sqrt{\boldsymbol{p}!\boldsymbol{q}!}} c_{\boldsymbol{p}} c_{\boldsymbol{q}}^{*}\,\mathrm{lHaf}\left(A_{\boldsymbol{p},\boldsymbol{q}}\right), \tag{3.80}$$

where

$$\kappa(\boldsymbol{\alpha},\hat{G}) = \frac{\exp\left[-\frac{1}{2}\tilde{\boldsymbol{d}}^{\dagger}\left(\boldsymbol{V}+\mathbb{1}_{2m}/2\right)^{-1}\tilde{\boldsymbol{d}}\right]}{\pi^m\sqrt{\mathrm{Det}\left(\boldsymbol{V}+\mathbb{1}_{2m}/2\right)}}, \tag{3.81}$$

where $\boldsymbol{V}$ and $\boldsymbol{d}$ are the covariance matrix and the diplacement vector of the Gaussian state $\hat{G}^{\dagger}|\boldsymbol{\alpha}\rangle$, respectively.

■

When the input core state is a multimode Fock state, we refer to the corresponding subclass of G_{core} circuits as G_{Fock} circuits. In that case, the sum in Eq. (3.65) reduces to a single term and we obtain the following expression:

Corollary 3.1. *Let $m,n\in\mathbb{N}^*$ and let $\boldsymbol{p}=(p_1,\ldots,p_m)$ with $|\boldsymbol{p}|=n$. Let $\hat{G}$ be a Gaussian unitary over m modes. For all $\boldsymbol{\alpha}\in\mathbb{C}^m$, let us write $\boldsymbol{V}$ and $\tilde{\boldsymbol{d}}=(\boldsymbol{d},\boldsymbol{d}^*)$ the covariance matrix and the displacement vector of the Gaussian state $\hat{G}^{\dagger}|\boldsymbol{\alpha}\rangle$. Then, the output probability density for the G_{Fock} circuit $\hat{G}$ with Fock state input $|\boldsymbol{p}\rangle$ and heterodyne detection, evaluated at $\boldsymbol{\alpha}$, is given by*

$$Pr_{Fock}[\boldsymbol{\alpha}] = \frac{\exp\left[-\frac{1}{2}\tilde{\boldsymbol{d}}^{\dagger}\left(\boldsymbol{V}+\mathbb{1}_{2m}/2\right)^{-1}\tilde{\boldsymbol{d}}\right]}{\boldsymbol{p}!\pi^m\sqrt{\mathrm{Det}\left(\boldsymbol{V}+\mathbb{1}_{2m}/2\right)}}\,\mathrm{lHaf}(A_{\boldsymbol{p},\boldsymbol{p}}), \tag{3.82}$$

where $A_{\boldsymbol{p},\boldsymbol{p}}$ is the square matrix of size $2n$ obtained with Lemma 3.3 from

$$V=\begin{pmatrix}\mathbb{0}_m & \mathbb{1}_m\\ \mathbb{1}_m & \mathbb{0}_m\end{pmatrix}\left[\mathbb{1}_{2m}-\left(\boldsymbol{V}+\mathbb{1}_{2m}/2\right)^{-1}\right] \quad \textit{and} \quad D=\left[\tilde{\boldsymbol{d}}^{\dagger}\left(\boldsymbol{V}+\mathbb{1}_{2m}/2\right)^{-1}\right]^{T}. \tag{3.83}$$

3.3.2 Strong simulation of weakly non-Gaussian quantum circuits

In this section, we use the expression obtained in Theorem 3.1 in order to study strong simulation of Gaussian circuits with few non-Gaussian elements. The first general result deals with general G_{core} circuits, i.e., Gaussian circuits with multimode core state input.

Theorem 3.2. *Let $m \in \mathbb{N}^*$ and let $|\boldsymbol{C}\rangle$ be an m-mode core state of support size $O(\text{poly}\, m)$ and degree $n = O(\log m)$. Then, G_{core} circuits over m modes with input $|\boldsymbol{C}\rangle$ and heterodyne detection can be strongly simulated efficiently classically.*

Proof. By Theorem 3.1, up to an efficiently computable prefactor, the output probability density is a sum of a polynomial number of loop hafnians, since the support size of the input core state is polynomial. The loop hafnian of a matrix of size r may be computed in time $O(r^3 2^{r/2})$ [BGQ19]. For $|\boldsymbol{p}| \leq n$ and $\boldsymbol{q} \leq n$, the matrices $A_{\boldsymbol{p},\boldsymbol{q}}$ appearing in Eq. (3.65) are efficiently computable square matrices of size $|\boldsymbol{p}| + |\boldsymbol{q}| \leq 2n$, so for $n = O(\log m)$, all the loop hafnians may be computed in time $O(\text{poly}\, m)$. Hence, the output probability density can be evaluated in time $O(\text{poly}\, m)$.

We now consider the evaluations of the marginal probability densities. Let $k \in \{1, \dots, m-1\}$, for all $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_k) \in \mathbb{C}^k$ we have

$$\begin{aligned} \text{Pr}_{\text{core}}[\boldsymbol{\alpha}] &= \text{Tr}\left[\hat{G}\,|\boldsymbol{C}\rangle\langle\boldsymbol{C}|\hat{G}^\dagger(\Pi_{\boldsymbol{\alpha}} \otimes \mathbb{1}_{m-k})\right] \\ &= \frac{1}{\pi^k}\,\text{Tr}\left[\hat{G}^\dagger(|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}| \otimes \mathbb{1}_{m-k})\hat{G}\,|\boldsymbol{C}\rangle\langle\boldsymbol{C}|\right] \\ &= \pi^{m-k}\int_{\boldsymbol{\beta}\in\mathbb{C}^m} Q_{\hat{G}^\dagger(|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|\otimes\mathbb{1}_{m-k})\hat{G}}(\boldsymbol{\beta})\, P_{|\boldsymbol{C}\rangle\langle\boldsymbol{C}|}(\boldsymbol{\beta})\, d^m\boldsymbol{\beta}\, d^m\boldsymbol{\beta}^*, \end{aligned} \tag{3.84}$$

where $\Pi_{\boldsymbol{\alpha}} = \frac{1}{\pi^k}|\alpha_1, \dots, \alpha_k\rangle\langle\alpha_1, \dots, \alpha_k|$ is the POVM element corresponding to the heterodyne detection of $(\alpha_1, \dots, \alpha_k)$ over the first k modes. With Lemma 3.3 and the proof of Theorem 3.1, it is sufficient to show that $Q_{\hat{G}^\dagger(|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|\otimes\mathbb{1}_{m-k})\hat{G}}$ is an efficiently computable Gaussian function in order to prove that the marginal probability density can be evaluated efficiently.

For all $(\alpha_1, \dots, \alpha_k) \in \mathbb{C}^k$ and all $(\gamma_1, \dots, \gamma_{m-k}) \in \mathbb{C}^{m-k}$ we write $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_k, 0, \dots, 0) \in \mathbb{C}^m$ and $\boldsymbol{\gamma} = (0, \dots, 0, \gamma_1, \dots, \gamma_{m-k}) \in \mathbb{C}^m$ so that $\boldsymbol{\alpha} + \boldsymbol{\gamma} = (\alpha_1, \dots, \alpha_k, \gamma_1, \dots, \gamma_{m-k}) \in \mathbb{C}^m$. Using the overcompleteness of coherent states we obtain, for all $(\alpha_1, \dots, \alpha_k) \in \mathbb{C}^k$ and for all $\boldsymbol{\beta} \in \mathbb{C}^m$,

$$\pi^{m-k} Q_{\hat{G}^\dagger(|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|\otimes\mathbb{1}_{m-k})\hat{G}}(\boldsymbol{\beta}) = \int_{\boldsymbol{\gamma}=(\gamma_1,\dots,\gamma_{m-k})\in\mathbb{C}^{m-k}} Q_{\hat{G}^\dagger|\boldsymbol{\alpha}+\boldsymbol{\gamma}\rangle\langle\boldsymbol{\alpha}+\boldsymbol{\gamma}|\hat{G}}(\boldsymbol{\beta})\, d^{m-k}\boldsymbol{\gamma}\, d^{m-k}\boldsymbol{\gamma}^*. \tag{3.85}$$

Let S and $\tilde{\boldsymbol{d}} = (\boldsymbol{d}, \boldsymbol{d}^*)$ be the symplectic matrix and the displacement vector associated with the Gaussian unitary $\hat{G}^\dagger$. The Gaussian state

$$\hat{G}^\dagger\,|\alpha_1, \dots, \alpha_k, \gamma_1, \dots, \gamma_{m-k}\rangle = \hat{G}^\dagger\,|\boldsymbol{\alpha} + \boldsymbol{\gamma}\rangle \tag{3.86}$$

is described by the covariance matrix $\boldsymbol{V} = \frac{1}{2}SS^\dagger$ and the displacement vector $S(\tilde{\boldsymbol{\alpha}} + \tilde{\boldsymbol{\gamma}}) + \tilde{\boldsymbol{d}}$. Its Q function is thus given by

$$Q_{\hat{G}^\dagger|\boldsymbol{\alpha}+\boldsymbol{\gamma}\rangle\langle\boldsymbol{\alpha}+\boldsymbol{\gamma}|\hat{G}}(\boldsymbol{\beta}) = \frac{\exp\left[-\frac{1}{2}(\tilde{\boldsymbol{\beta}} - S(\tilde{\boldsymbol{\alpha}} + \tilde{\boldsymbol{\gamma}}) - \tilde{\boldsymbol{d}})^\dagger(\boldsymbol{V} + \mathbb{1}_{2m}/2)^{-1}(\tilde{\boldsymbol{\beta}} - S(\tilde{\boldsymbol{\alpha}} + \tilde{\boldsymbol{\gamma}}) - \tilde{\boldsymbol{d}})\right]}{\pi^m\sqrt{\text{Det}(\boldsymbol{V} + \mathbb{1}_{2m}/2)}}, \tag{3.87}$$

for all $(\alpha_1,\ldots,\alpha_k)\in\mathbb{C}^k$, for all $(\gamma_1,\ldots,\gamma_{m-k})\in\mathbb{C}^{m-k}$ and for all $\boldsymbol{\beta}\in\mathbb{C}^m$. Let us discard the efficiently computable denominator and expand the product in the exponential. Writing $M=(\boldsymbol{V}+\mathbb{1}_{2m}/2)^{-1}$, we are left with

$$\exp\left[-\frac{1}{2}(\tilde{\boldsymbol{\beta}}-S\tilde{\boldsymbol{\alpha}}-\tilde{\boldsymbol{d}})^\dagger M(\tilde{\boldsymbol{\beta}}-S\tilde{\boldsymbol{\alpha}}-\tilde{\boldsymbol{d}})\right]\cdot\exp\left[-\frac{1}{2}\tilde{\boldsymbol{\gamma}}^\dagger S^\dagger MS\tilde{\boldsymbol{\gamma}}+(\tilde{\boldsymbol{\beta}}-S\tilde{\boldsymbol{\alpha}}-\tilde{\boldsymbol{d}})^\dagger MS\tilde{\boldsymbol{\gamma}}\right], \tag{3.88}$$

The first exponential term is an efficiently computable Gaussian function which factors out of the integral in Eq. (3.85). Rewriting Eq. (3.85) up to this efficiently computable Gaussian function we are left with

$$\begin{aligned}\int_{\boldsymbol{\gamma}=(0,\ldots,0,\gamma_1,\ldots,\gamma_{m-k})\in\mathbb{C}^m}&\exp\left[-\frac{1}{2}\tilde{\boldsymbol{\gamma}}^\dagger S^\dagger MS\tilde{\boldsymbol{\gamma}}+(\tilde{\boldsymbol{\beta}}-S\tilde{\boldsymbol{\alpha}}-\tilde{\boldsymbol{d}})^\dagger MS\tilde{\boldsymbol{\gamma}}\right]d^{m-k}\boldsymbol{\gamma}d^{m-k}\boldsymbol{\gamma}^*\\&=\int_{\boldsymbol{\gamma}=(\gamma_1,\ldots,\gamma_{m-k})\in\mathbb{C}^{m-k}}\exp\left[-\frac{1}{2}\tilde{\boldsymbol{\gamma}}^T V\tilde{\boldsymbol{\gamma}}+D^T\tilde{\boldsymbol{\gamma}}\right]d^{2(m-k)}\tilde{\boldsymbol{\gamma}},\end{aligned} \tag{3.89}$$

where V is the $2(m-k)\times 2(m-k)$ submatrix of

$$\begin{pmatrix}\mathbb{0}_m & \mathbb{1}_m\\ \mathbb{1}_m & \mathbb{0}_m\end{pmatrix}S^\dagger MS \tag{3.90}$$

obtained by removing the rows and colums of indices l and $m+l$ for $l\in\{1,\ldots,k\}$, and where D is the column vector of size $2(m-k)$ obtained by removing the elements of

$$\left[(\tilde{\boldsymbol{\beta}}-S\tilde{\boldsymbol{\alpha}}-\tilde{\boldsymbol{d}})^\dagger MS\right]^T \tag{3.91}$$

of indices l and $m+l$ for $l\in\{1,\ldots,k\}$. The matrix V and the vector D are efficiently computable. Moreover,

$$\int_{\boldsymbol{\gamma}=(\gamma_1,\ldots,\gamma_{m-k})\in\mathbb{C}^{m-k}}\exp\left[-\frac{1}{2}\tilde{\boldsymbol{\gamma}}^T V\tilde{\boldsymbol{\gamma}}+D^T\tilde{\boldsymbol{\gamma}}\right]d^{2(m-k)}\tilde{\boldsymbol{\gamma}}=\frac{(2\pi)^{m-k}}{\sqrt{\mathrm{Det}(V)}}\exp\left[\frac{1}{2}D^T V^{-1}D\right], \tag{3.92}$$

which is an efficiently computable Gaussian function of $\boldsymbol{\beta}$.

This implies that the value of the marginal probability density $\Pr[\alpha_1,\ldots,\alpha_k]$ may be computed efficiently. Moreover, it is clear that this does not depent on the choice of $k\in\{1,\ldots,m-1\}$ and on the choice of the modes. Hence, all marginal probability densities may be evaluated in time $O(\mathrm{poly}\,m)$.

■

This result has consequences for the simulability of various continuous variable quantum computing models, in particular those based on Gaussian operations and photon additions or subtractions. We consider three examples in what follows: Interleaved Photon-Added Gaussian circuits (IPAG), Interleaved Photon-Subtracted Gaussian circuits (IPSG) and Gaussian circuits with input Fock states (G_{Fock}).

The stellar hierarchy of single-mode pure quantum states derived in the previous chapter details the engineering of a single-mode quantum state from vacuum using unitary Gaussian operations and single photon addition as a non-Gaussian operation. In particular, the states of finite stellar rank, which corresponds to the states that can be obtained from the vacuum using a finite number of single photon additions or subtractions, are shown to be exactly the states that are obtained by applying a Gaussian unitary operation to a single-mode core state (Theorem 2.4).

As we will see here, the situation is different in the multimode case: we show that the set of states that can be obtained from a multimode core state with a multimode Gaussian unitary operation is strictly larger than the set of states that can be obtained from the vacuum using a finite number of single photon additions and Gaussian unitary operations (Lemma 3.4). We also deduce strong simulability results for Gaussian sampling of the latter states. To that end, we consider the family of quantum circuits which sample from states in this set with product unbalanced heterodyne detection, which we refer to as Interleaved Photon-Added Gaussian circuits (IPAG) due to their structure (Fig. 3.3).

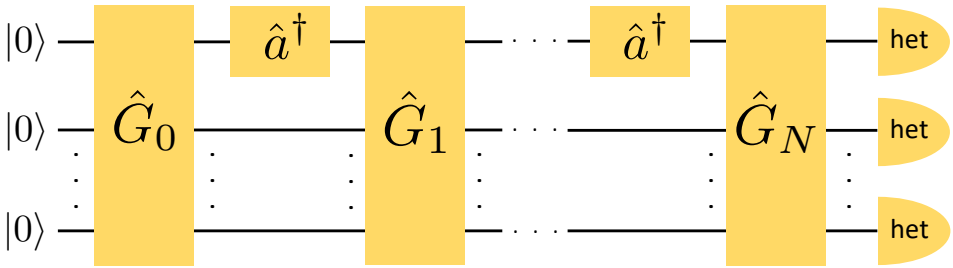

Figure 3.3: Representation of Interleaved Photon-Added Gaussian circuits with n photons additions. The unitaries $\hat{G}_0, \ldots \hat{G}_n$ are Gaussian and the measurement is performed by balanced heterodyne detection. Note that all photon additions act on the first mode without loss of generality, since swapping two modes is a Gaussian operation.

Formally, IPAG circuits with m modes and n photon additions are defined as: (i) product vacuum state over m modes in input, (ii) an evolution composed of interleaved multimode Gaussian unitaries $\hat{G}_0, \ldots, \hat{G}_n$ and n single-mode photon additions, and (iii) product unbalanced heterodyne detection (not necessarily with the same unbalancing for each mode). Without loss of generality, all the photon additions act on the first mode, since swapping two modes is a Gaussian operation. Moreover, up to an added multimode squeezing to the final Gaussian unitary $\hat{G}_n$, the measurement may be written as a product balanced heterodyne detection.

We first establish a reduction to an equivalent model where the evolution and measurement are Gaussian and only the input state is non-Gaussian. This is done by commuting the photon additions to the input of the circuit. The output state of an IPAG circuit with m modes, n photon

additions and Gaussian unitaries $\hat{G}_0,\dots,\hat{G}_n$ is given by

$$\hat{G}_n\hat{a}_1^\dagger\hat{G}_{n-1}\hat{a}_1^\dagger\dots\hat{G}_1\hat{a}_1^\dagger\hat{G}_0\,|0\rangle^{\otimes m}, \tag{3.93}$$

where we have assumed that all the photon additions act on the first mode without loss of generality. Gaussian operations act on annihilation and creation operators through their symplectic representation. They induce affine transformations of the vector of annihilation and creation operators (see section 1.3). Let us define the column vector of ladder operators

$$\boldsymbol{\lambda}^\dagger = \begin{pmatrix} \hat{a}_1^\dagger \\ \vdots \\ \hat{a}_m^\dagger \\ \hat{a}_1 \\ \vdots \\ \hat{a}_m \end{pmatrix}, \tag{3.94}$$

and let $\hat{G}$ be an m-mode Gaussian operation. Then, there exists a $2m\times 2m$ symplectic matrix $S=(s_{ij})_{1\le i,j\le 2m}$ and a complex vector $d=(d_1,\dots,d_m)$, such that for all $k\in\{1,\dots,m\}$,

$$\begin{aligned}\hat{G}\hat{a}_k^\dagger\hat{G}^\dagger &= d_k + (S\boldsymbol{\lambda}^\dagger)_k \\ &= d_k + \sum_{l=1}^{m} s_{k,l}\hat{a}_l^\dagger + s_{k,m+l}\hat{a}_l,\end{aligned} \tag{3.95}$$

where $(S\boldsymbol{\lambda}^\dagger)_k$ indicates the k^{th} element of the column vector $S\boldsymbol{\lambda}^\dagger$. Hence, commuting to the right the creation operators in Eq. (3.93), starting by the rightmost one, yields

$$\begin{aligned}\hat{G}_n\hat{a}_1^\dagger\dots\hat{G}_1\hat{a}_1^\dagger\hat{G}_0\,|0\rangle^{\otimes m} &= \hat{G}_n\hat{a}_1^\dagger\hat{G}_2\dots a_1^\dagger\hat{G}_1\hat{G}_0\left[d_1^{(0)}+(S^{(0)}\boldsymbol{\lambda})_1\right]|0\rangle^{\otimes m} \\ &= \dots \\ &= \hat{G}_n\dots\hat{G}_0\left[d_1^{(n-1)}+(S^{(n-1)}\boldsymbol{\lambda})_1\right]\dots\left[d_1^{(0)}+(S^{(0)}\boldsymbol{\lambda})_1\right]|0\rangle^{\otimes m},\end{aligned} \tag{3.96}$$

where $S^{(k)}$ and $d^{(k)}$ implement the affine transformation corresponding to the action of $(\hat{G}_k\hat{G}_{k-1}\dots\hat{G}_0)^\dagger$, for all $k\in\{0,\dots,n-1\}$. Writing $\hat{G}:=\hat{G}_n\hat{G}_{n-1}\dots\hat{G}_0$, $S^{(k)}=(s_{i,j}^{(k)})_{1\le i,j\le 2m}$, and $d^{(k)}=(d_1^{(k)},\dots,d_m^{(k)})$ for $k\in\{0,\dots,n-1\}$, we obtain the output state

$$\hat{G}\,|\boldsymbol{C}_{\mathrm{IPAG}}\rangle, \tag{3.97}$$

where the state

$$|\boldsymbol{C}_{\mathrm{IPAG}}\rangle := \left(d_1^{(n-1)}+\sum_{l=1}^{m}s_{1,l}^{(n-1)}\hat{a}_l^\dagger+s_{1,m+l}^{(n-1)}\hat{a}_l\right)\dots\left(d_1^{(0)}+\sum_{l=1}^{m}s_{1,l}^{(0)}\hat{a}_l^\dagger+s_{1,m+l}^{(0)}\hat{a}_l\right)|0\rangle^{\otimes m} \tag{3.98}$$

is a multimode core state of degree n (and not less, by property of symplectic matrices). Using this characterisation, we obtain the following result:

Lemma 3.4. *The set of output states of IPAG circuits is strictly included in the set of output states of G_{core} circuits.*

Proof. The inclusion is immediate with Eq. (3.97). Up to the Gaussian unitary, it is sufficient to consider core states. To prove the strict inclusion, we show that the m-mode core state $(|20\rangle + |01\rangle) \otimes |0\rangle^{\otimes m-2}$ (we omit normalisation), which has degree 2, is not a core state of the form of Eq. (3.98).

By Eq. (3.98), all m-mode core states of IPAG circuits of degree 2 have the form

$$\left(d^{(1)} + \sum_{k=1}^{m} s_k^{(1)} \hat{a}_k^\dagger + s_{m+k}^{(1)} \hat{a}_k\right)\left(d^{(0)} + \sum_{l=1}^{m} s_l^{(0)} \hat{a}_l^\dagger + s_{1,m+l}^{(0)} \hat{a}_l\right)|0\rangle^{\otimes m}, \tag{3.99}$$

for some complex numbers $d^{(0)}, d^{(1)}, s_1^{(0)}, \ldots, s_{2m}^{(0)}, s_1^{(1)}, \ldots, s_{2m}^{(1)}$. This expression rewrites

$$\left(d^{(1)} + \sum_{k=1}^{m} s_k^{(1)} \hat{a}_k^\dagger + s_{m+k}^{(1)} \hat{a}_k\right)\left(\sum_{l=1}^{m} s_l^{(0)} |\mathbf{1}_l\rangle + d^{(0)} |\mathbf{0}\rangle\right), \tag{3.100}$$

where for all $l \in \{1, \ldots, m\}$, we write $\mathbf{1}_l = (0, \ldots, 0, 1, 0 \ldots, 0)$, with a 1 at the l^{th} position. We finally obtain

$$\sqrt{2} \sum_{k=1}^{m} s_k^{(0)} s_k^{(1)} |\mathbf{2}_k\rangle + \sum_{\substack{k,l=1 \\ k \neq l}}^{m} s_k^{(0)} s_l^{(1)} |\mathbf{1}_k + \mathbf{1}_l\rangle + \sum_{k=1}^{m} \left(d^{(1)} s_k^{(0)} + d^{(0)} s_k^{(1)}\right) |\mathbf{1}_k\rangle + \left(d^{(0)} d^{(1)} + \sum_{k=1}^{m} s_k^{(0)} s_{m+k}^{(1)}\right) |\mathbf{0}\rangle, \tag{3.101}$$

where for all $k \in \{1, \ldots, m\}$, we write $\mathbf{2}_k = (0, \ldots, 0, 2, 0 \ldots, 0)$, with a 2 at the k^{th} position. On the other hand we have

$$(|20\rangle + |01\rangle) \otimes |0\rangle^{\otimes m-2} = |\mathbf{2}_1\rangle + |\mathbf{1}_2\rangle. \tag{3.102}$$

In order for this core state to be of the form of Eq. (3.101) we must have

$$\begin{cases} s_1^{(0)} s_1^{(1)} \neq 0 \\ s_k^{(0)} s_l^{(1)} = 0, \text{ for } k \neq l, \end{cases} \tag{3.103}$$

by considering the first and second terms of Eq. (3.101). This implies $s_k^{(0)} = s_k^{(1)} = 0$ for all $k \neq 1$. Hence, the coefficient of $|\mathbf{1}_2\rangle$ in Eq. (3.101) is equal to 0, while it is nonzero in Eq. (3.102). Therefore the core state described by Eq. (3.102) cannot be generated by an IPAG circuit.

■

In other words, the set of states that can be obtained from a multimode core state with a multimode Gaussian unitary operation is strictly larger than the set of states that can be obtained from the vacuum using a finite number of single photon additions and Gaussian unitary operations, unlike in the single mode case, where the two sets coincide.

Another consequence of Eqs. (3.97) and (3.98) is the following result:

Lemma 3.5. *IPAG circuits over m modes with $n = O(1)$ photon additions can be strongly simulated efficiently classically.*

Proof. When $n = O(1)$, the support size of the core state $|\boldsymbol{C}_{\mathrm{IPAG}}\rangle$ in Eq. (3.98) is $O(\mathrm{poly}\, m)$ and its degree is $O(1)$. Then, the result comes from a direct application of Theorem 3.2.

■

When $n = O(\log m)$ however, the support size of the core state is superpolynomial, so the classical algorithm is no longer efficient.

Similarly, we can define Interleaved Photon-Subtracted Gaussian circuits (IPSG) by replacing photon additions by subtractions in the definition of IPAG circuits. With the same reasoning we obtain the following result:

Corollary 3.2. *IPSG circuits over m modes with $n = O(1)$ photon subtractions can be strongly simulated efficiently classically.*

Proof. We use again the fact that Gaussian operations induce an affine transformation of the vector of annihilation and creation operators. Let $\hat{G}$ be an m-mode Gaussian operation with symplectic matrix S and displacement vector $\boldsymbol{d}$. Writing

$$\boldsymbol{\lambda}^\dagger = \begin{pmatrix} \hat{a}_1^\dagger \\ \vdots \\ \hat{a}_m^\dagger \\ \hat{a}_1 \\ \vdots \\ \hat{a}_m \end{pmatrix} \tag{3.104}$$

and taking this time the adjoint of Eq. (3.95) we obtain

$$\begin{aligned} \hat{G}\hat{a}_k\hat{G}^\dagger &= d_k^* + (S\boldsymbol{\lambda}^\dagger)_k^\dagger \\ &= d_k^* + \sum_{l=1}^{m} s_{k,l}^* \hat{a}_l + s_{k,m+l}^* \hat{a}_l^\dagger, \end{aligned} \tag{3.105}$$

for all $k \in \{1,\dots,m\}$. The same proof as for IPAG circuits shows that the output state of an IPSG circuit with n photon subtraction and Gaussian evolution $\hat{G}_0,\dots,\hat{G}_n$ reads

$$\hat{G}\,|\boldsymbol{C}_{\mathrm{IPSG}}\rangle\,, \tag{3.106}$$

where $\hat{G} = \hat{G}_n \dots \hat{G}_0$ and where

$$\hat{G}\,|\boldsymbol{C}_{\mathrm{IPSG}}\rangle := \left(d_1^{*(n-1)} + \sum_{l=1}^{m} s_{1,l}^{*(n-1)} \hat{a}_l + s_{1,m+l}^{*(n-1)} \hat{a}_l^\dagger\right) \dots \left(d_1^{*(0)} + \sum_{l=1}^{m} s_{1,l}^{*(0)} \hat{a}_l + s_{1,m+l}^{*(0)} \hat{a}_l^\dagger\right) |0\rangle^{\otimes m}, \tag{3.107}$$

where $S^{(k)} = (s_{i,j}^{(k)})_{1 \leq i,j \leq 2m}$ and $\boldsymbol{d}^{(k)} = (d_1^{(k)}, \dots, d_m^{(k)})$ are the symplectic matrix and the displacement vector of $(\hat{G}_k \hat{G}_{k-1} \dots \hat{G}_0)^\dagger$, for all $k \in \{0, \dots, n-1\}$. When $n = O(1)$, this core state has support size $O(\mathrm{poly}\, m)$ and degree $O(1)$, and Theorem 3.2 concludes the proof.

■

Note that the same reasoning also holds for Gaussian circuits interleaved with both photon additions and subtractions.

A particular subclass of IPAG circuits, where all the photon additions act at the beginning of the circuit, is the class of G_{Fock} circuits, i.e., Gaussian circuits with Fock state input. In that case, the input is a multimode core state of support size 1. With Corollary 3.1, we obtain the following result as an immediate consequence of Theorem 3.2:

Lemma 3.6. *Let $m \in \mathbb{N}^*$ and let $\boldsymbol{p} \in \mathbb{N}^m$, such that $|\boldsymbol{p}| = O(\log m)$. Then, G_{Fock} circuits over m modes with Fock state input $|\boldsymbol{p}\rangle$ and heterodyne detection can be strongly simulated efficiently classically.*

In other words, sampling with Gaussian measurements over m modes from $n = O(\log m)$ indistinguishable photons is strongly simulable classically. This contrasts with the case where $m = O(\mathrm{poly}\, n)$: we show in the next section that strong simulation and even weak simulation of sampling from n photons in m modes with Gaussian measurements is classically hard in that case.

3.3.3 Quantum supremacy with non-Gaussian states

In the recent years, there has been an increasing interest in quantum circuits that define subuniversal models of quantum computation [BJS10, AA13, MFF14, BMS16, FH16, DMK$^+$17, BIS$^+$18]. These models may allow for an experimental demonstration of quantum computational supremacy [HM17], i.e., the predicted dramatic speedup of quantum computers over their classical counterparts for some computational tasks [AAB$^+$19]. Subuniversal models for demonstrating quantum supremacy are associated with sampling problems for which the task is to draw random numbers according to a specific probability distribution. Some of these probability distributions are likely to be hard to sample for classical computers, assuming widely accepted conjectures in computer science, such as the fact that the polynomial hierarchy does not collapse, for example with the celebrated Boson Sampling (see section 1.4.5 and [AA13]).

For continuous variable quantum circuits, the classical hardness of circuits with Gaussian input and evolution and non-Gaussian measurement, corresponding to Gaussian Boson Sampling,

was proven in [LLRK$^+$14, HKS$^+$16]. These circuits are composed of input squeezed states, passive linear optics evolution, and photon counters. In that case, the measurement is a discrete variable measurement. Subuniversal models with Gaussian input and measurements but non-Gaussian gates are for instance related to the continuous variable implementation of Instantaneous Quantum Computing [DMK$^+$17, DMK$^+$19]. Other subuniversal continuous variable circuits that displays non-Gaussian input states together with Gaussian operations and measurements, have been recently considered [CC17, LRKR17].

In this section, we define and study a family of continuous variable quantum circuits which we refer to as CVS circuits—for Continuous Variable Sampling—that take non-Gaussian input states and have Gaussian evolution and measurements. The non-Gaussian input states are either single-photons ($\mathrm{CVS_{SP}}$ circuits), single photon-subtracted squeezed vacuum states ($\mathrm{CVS_{PS}}$ circuits), or single photon-added squeezed vacuum states ($\mathrm{CVS_{PA}}$ circuits), and the measurement is unbalanced heterodyne detection (see section 1.4.2), yielding a continuous variable outcome. These models are analog to the Boson Sampling model [AA13] and the Photon-Added or photon-Subtracted Squeezed Vacuum (PASSV) sampling model [OSM$^+$15], but with heterodyne detection replacing photon counting. We show in what follows that they allow for the demonstration of quantum computational supremacy with non-Gaussian input states and Gaussian measurements.

The family of CVS circuits is a subclass of IPAG and G_{core} circuits. Their architecture is inspired by recent experiments performed at Laboratoire Kastler Brossel (LKB), where mode-selective single photon subtraction from a collection of multimode squeezed states has been recently demonstrated [RJD$^+$17], and where simultaneous detection of all the optical modes can also be implemented by means of multipixel homodyne detection [Bec00, FGC$^+$13].

We use for brevity the notations $c_\chi = \cosh\chi$, $s_\chi = \sinh\chi$ and $t_\chi = \tanh\chi$, for all $\chi \in \mathbb{R}$. $\mathrm{CVS_{PS}}$ circuits are defined formally as follows (see Fig. 3.4, $\mathrm{CVS_{PA}}$ and $\mathrm{CVS_{SP}}$ are defined analogously by changing the non-Gaussian input states). Let m be the total number of optical modes. We recall the definition of the squeezing operator with squeezing parameter $\xi \in \mathbb{C}$: $\hat{S}(\xi) = e^{\frac{1}{2}(\xi \hat{a}^{\dagger 2} - \xi^* \hat{a}^2)}$. We restrict to real squeezing parameters in what follows. In that case, $\xi < 0$ results in $\hat{p}$-squeezing while $\xi > 0$ in $\hat{q}$-squeezing.

The first n modes are single photon-subtracted squeezed vacuum states denoted by $\hat{a}\,|\xi\rangle$, where we omit the normalisation factor. The remaining $m - n$ modes are squeezed vacuum states $|\xi\rangle$. We assume that the real squeezing parameter ξ is uniform over all the modes and does not depend on the number of modes m. We require that n is even and that $m = O(\mathrm{poly}\, n) \geq 2n$.

The input modes undergo a passive linear evolution $\hat{U}$ that is described by an $m \times m$ unitary matrix U of the form

$$U = O e^{i\phi\Sigma} \tag{3.108}$$

with $\phi \in \mathbb{R}$, $O \in \mathcal{O}(m)$ and $\Sigma \in \mathcal{O}_S(m)$, i.e., O is a real orthogonal matrix, and Σ is a real symmetric orthogonal matrix, and hence satisfies $\Sigma^2 = 1$ (this choice yields a convenient expression for the output probability distribution of CVS circuits).

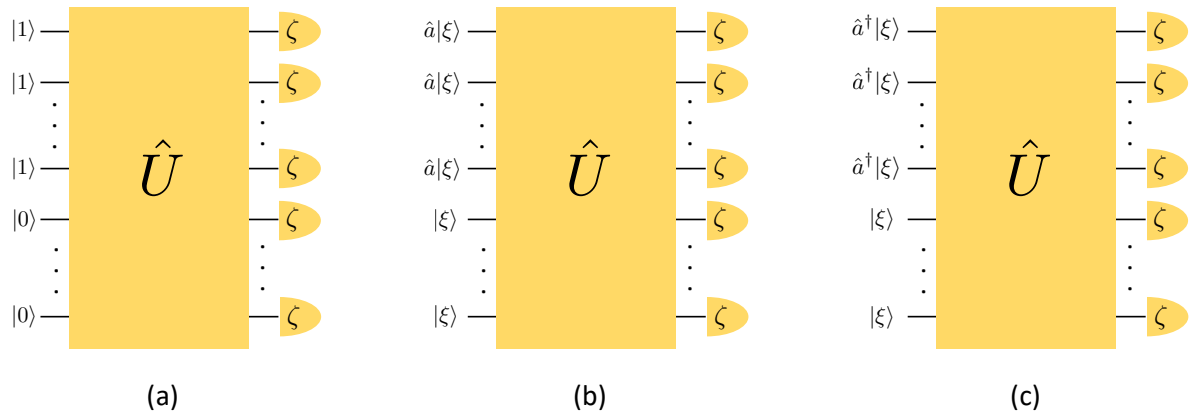

Figure 3.4: Representation of CVS circuits. The passive linear optics evolution is associated with the unitary matrix U defined in Eq. (3.108). Measurement is performed by unbalanced heterodyne detection with parameter ζ. (a) $\mathrm{CVS_{SP}}$: in input are vacuum states and single photon states. (b) $\mathrm{CVS_{PS}}$: in input are squeezed vacuum states and single photon-subtracted squeezed vacuum states. (c) $\mathrm{CVS_{PA}}$: in input are squeezed vacuum states and single photon-added squeezed vacuum states.

Finally, the mode quadratures are measured by unbalanced heterodyne detection with parameter $\zeta \in \mathbb{R}$, i.e., by projecting the output states onto displaced squeezed vacuum states $|\alpha_j,\zeta\rangle = \hat{D}(\alpha_j)\hat{S}(\zeta)|0\rangle$. The term $\alpha_j = \sqrt{c_\zeta}(e^{-\zeta/2}q_j + ie^{\zeta/2}p_j)$ corresponds to the displacement value of the j^{th} mode, where q_j and p_j are the measured outcomes at the (distinct) output modes of the j^{th}-mode heterodyne detector. $\hat{D}(\alpha)$ is the displacement operator $\hat{D}(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}$ (see also section 1.4.2). The squeezing parameter of the detection ζ is uniform over all the modes and satisfies

$$|\zeta| = \Omega\left(2^{-\operatorname{poly} m}\right). \tag{3.109}$$

From the Gaussian convertibility example in Eq. (2.81) of the previous chapter, we know that a photon-subtracted squeezed vacuum state, a photon-added squeezed vacuum state and a squeezed single-photon Fock state, all with the same real squeezing parameter $\xi \in \mathbb{R}$, are equal:

$$\hat{S}(\xi)|1\rangle = -\frac{1}{s_\xi}\hat{a}\,|\xi\rangle = \frac{1}{c_\xi}\hat{a}^\dagger|\xi\rangle\,. \tag{3.110}$$

By virtue of these identities, the architectures $\mathrm{CVS_{PS}}$ and $\mathrm{CVS_{PA}}$ are in fact identical. Moreover, the architecture $\mathrm{CVS_{SP}}$ is obtained from the first two by letting the squeezing parameter ξ go to 0. Hereafter, we therefore refer to all three configurations as CVS circuits over m modes, with n non-Gaussian input states, input squeezing $\xi \in \mathbb{R}$, evolution $U = Oe^{i\phi\Sigma}$ and unbalanced heterodyne detection $\zeta \in \mathbb{R}$ (Fig. 3.5): all CVS circuits are therefore specific G_{Fock} circuits, being also a subclass of IPAG and G_{core} circuits.

In order to obtain an output probability distribution, we introduce a finite binning of size $\eta > 0$ for the output probability density of CVS circuits. This allows for the definition of a set of indices

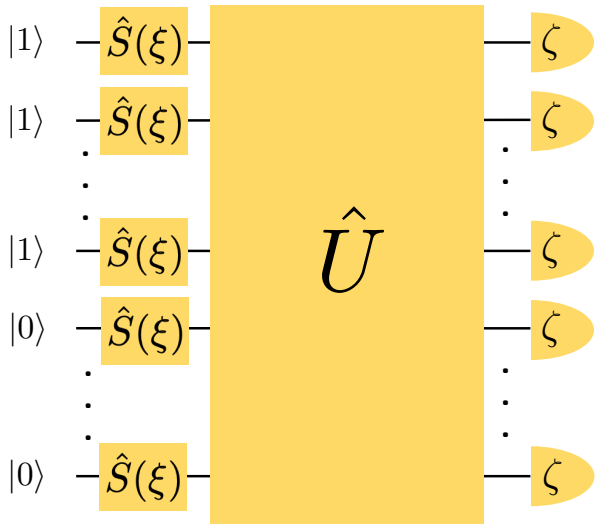

Figure 3.5: An alternative representation of CVS circuits. The input has been rewritten using the mapping of Eq. (3.110). $\hat{S}(\xi)$ is the unitary associated to a squeezing with parameter $\xi \in \mathbb{R}$, while $\hat{U}$ is a passive linear optics transformation described by a unitary matrix U defined in Eq. (3.108). The output is measured using unbalanced heterodyne detection with parameter $\zeta \in \mathbb{R}$.

$\mathbf{b} = (b_1^{(q)},\ldots,b_m^{(q)},b_1^{(p)},\ldots,b_m^{(p)}) \in \mathbb{Z}^{2m}$ that corresponds to bins for the $\hat{q}$ and $\hat{p}$ quadratures. We denote $\Pr^{\eta}_{\text{CVS}}[\mathbf{b}]$ the discrete probability that, for all $j \in \{1,\ldots,m\}$, the j^{th}-mode measurement outcome (q_j,p_j) falls into the boxes $B_j^{(q)} = \left[b_j^{(q)}\eta,(b_j^{(q)}+1)\eta\right]$, $B_j^{(p)} = \left[b_j^{(p)}\eta,(b_j^{(p)}+1)\eta\right]$. This probability distribution is related to the real-valued probability density associated with CVS circuits, $\Pr_{\text{CVS}}[q_1,p_1,\ldots,q_m,p_m]$, by

$$\Pr^{\eta}_{\text{CVS}}[\mathbf{b}] = \prod_{j=1}^{m}\int_{B_j^{(q)}}\int_{B_j^{(p)}} \Pr_{\text{CVS}}[q_1,p_1,\ldots,q_m,p_m]\,dq_j\,dp_j, \tag{3.111}$$

where $q_1,p_1,\ldots,q_m,p_m$ are the continuously distributed measurement outcomes of the product unbalanced heterodyne detection over m modes. This model of detection is equivalent to perfect heterodyne detection, followed by a binning of the outcome results performed at the stage of post-processing. We assume a resolution scaling with the number of modes as $\eta \sim 2^{-\text{poly}\,m}$.

We prove that the probability distribution $\Pr^{\eta}_{\text{CVS}}[\mathbf{b}]$ is hard to sample for a classical computer, both in the worst case scenario—i.e., weak simulation of all CVS circuits is hard—and in the average case scenario—i.e., weak simulation of a randomly chosen CVS circuit is hard—under the assumption that the polynomial hierarchy does not collapse (see section 1.4.5 for a brief review of the complexity classes appearing in this section). The argument adapts proof techniques from [AA13, HKS$^+$16, LRKR17, CC17] and follows these lines:

- We compute the expression $\Pr_{\text{CVS}}[\mathbf{0}]$ of the (continuous) probability density evaluated at $\mathbf{0} = (0,\ldots,0)$ for a given CVS circuit.

- We show that for any real matrix X, one can find a CVS circuit such that the expression $\Pr_{\text{CVS}}[\mathbf{0}]$ is related to the square of the permanent of X by a multiplicative factor.
- We show that a classical machine sampling efficiently from the (discrete) probability distribution $\Pr^{\eta}_{\text{CVS}}[\mathbf{b}]$ associated to this CVS circuit would allow us to approximate multiplicatively the square of the permanent of X in the third level of the polynomial hierarchy, yielding a contradiction with the widely believed conjecture that the polynomial hierarchy does not collapse.

Lemma 3.7. *We consider a CVS circuit over m modes with $n = 2p$ non-Gaussian input states, input squeezing $\xi \in \mathbb{R}$, evolution $U = Oe^{i\phi\Sigma}$ and unbalanced heterodyne detection $\zeta \in \mathbb{R}$. Then,*

$$Pr_{CVS}[\mathbf{0}] = \kappa(\phi,\xi,\zeta)\,\text{Haf}(\Sigma_n)^2, \tag{3.112}$$

where Σ_n is the $n \times n$ top left submatrix of Σ and where

$$\kappa(\phi,\xi,\zeta) = \frac{2^{m/2} s_{2\zeta}^n \sin^n(2\phi)}{\pi^m \left[1 + c_{2\xi}c_{2\zeta} - s_{2\xi}s_{2\zeta}\cos(2\phi)\right]^{n+m/2}}. \tag{3.113}$$

Proof. CVS circuits are G_{Fock} circuits. For a CVS circuit over m modes with $n = 2p$ non-Gaussian input states, input squeezing $\xi > 0$, evolution $U = Oe^{i\phi\Sigma}$ and unbalanced heterodyne detection $\zeta > 0$ (Fig. 3.5), the multimode Fock state input is $|1\rangle^{\otimes n} \otimes |0\rangle^{\otimes m-n}$ and the corresponding Gaussian unitary evolution is given by

$$\hat{G} = \hat{S}^{\dagger}(\zeta)^{\otimes m} \hat{U} \hat{S}(\xi)^{\otimes m}. \tag{3.114}$$

Let $\boldsymbol{V}$ be the covariance matrix of the Gaussian state $\hat{G}^{\dagger}|\mathbf{0}\rangle$, its displacement vector being $\mathbf{0}$. By Corollary 3.1, the output probability density evaluated at $(0,\ldots,0)$ is given by

$$\Pr_{\text{CVS}}[\mathbf{0}] = \frac{\text{lHaf}(A_n)}{\pi^m \sqrt{\text{Det}(\boldsymbol{V} + \mathbb{1}_{2m}/2)}}, \tag{3.115}$$

where A_n is the square matrix of size $2n$ obtained with Lemma 3.3 from

$$V = \begin{pmatrix} \mathbb{O}_m & \mathbb{1}_m \\ \mathbb{1}_m & \mathbb{O}_m \end{pmatrix} \left[\mathbb{1}_{2m} - (\boldsymbol{V} + \mathbb{1}_{2m}/2)^{-1}\right] \quad \text{and} \quad D = 0, \tag{3.116}$$

by keeping only the k^{th} and $(m+k)^{th}$ rows and columns of V for $k \in \{1,\ldots,n\}$ and by replacing its diagonal entries by the corresponding elements of D. Now $D = 0$, and for a matrix whose diagonal entries are 0, the loop hafnian is equal to the hafnian. Hence,

$$\Pr_{\text{CVS}}[\mathbf{0}] = \frac{\text{Haf}(A_n)}{\pi^m \sqrt{\text{Det}(\boldsymbol{V} + \mathbb{1}_{2m}/2)}}. \tag{3.117}$$

We now derive the expression of the matrix A_n in terms of the CVS circuit parameters:

Lemma 3.8. *Define*

$$B := \frac{-s_{2\xi}c_{2\zeta} + c_{2\xi}s_{2\zeta}\cos(2\phi)}{1 + c_{2\xi}c_{2\zeta} - s_{2\xi}s_{2\zeta}\cos(2\phi)}\mathbb{1}_m + i\frac{s_{2\zeta}\sin(2\phi)}{1 + c_{2\xi}c_{2\zeta} - s_{2\xi}s_{2\zeta}\cos(2\phi)}\Sigma. \tag{3.118}$$

Then,

$$A_n = \begin{pmatrix} B_n^* & \mathbb{0}_n \\ \mathbb{0}_n & B_n \end{pmatrix}, \tag{3.119}$$

where B_n is the $n \times n$ top left submatrix of B.

Proof. We show that $V = \begin{pmatrix} B^* & \mathbb{0}_m \\ \mathbb{0}_m & B \end{pmatrix}$. From Eq. (3.116) we have

$$V = \begin{pmatrix} \mathbb{0}_m & \mathbb{1}_m \\ \mathbb{1}_m & \mathbb{0}_m \end{pmatrix}\left[\mathbb{1}_{2m} - (\boldsymbol{V} + \mathbb{1}_{2m}/2)^{-1}\right], \tag{3.120}$$

where $\boldsymbol{V}$ is the covariance matrix of the Gaussian state

$$\hat{S}^\dagger(\xi)^{\otimes m}\hat{U}^\dagger\hat{S}(\zeta)^{\otimes m}\left|0\right\rangle\left\langle 0\right|^{\otimes m}\hat{S}^\dagger(\zeta)^{\otimes m}\hat{U}\hat{S}(\xi)^{\otimes m}. \tag{3.121}$$

This covariance matrix is given by (see section 1.3.3)

$$\boldsymbol{V} = S_{-\xi}S_{U^\dagger}S_\zeta \boldsymbol{V}_{\text{vac}} S_\zeta^\dagger S_{U^\dagger}^\dagger S_{-\xi}^\dagger, \tag{3.122}$$

where $\boldsymbol{V}_{\text{vac}} = \mathbb{1}_{2m}/2$ is the covariance matrix of the vacuum state over m modes, and $S_{-\xi}$, $S_{U^\dagger}$ and S_ζ are the symplectic matrices describing the action on the covariance matrix of the operators $\hat{S}^\dagger(\xi)^{\otimes m}$, $\hat{U}^\dagger$ and $\hat{S}(\zeta)^{\otimes m}$, respectively. Using the notation $c_\chi = \cosh\chi$ and $s_\chi = \sinh\chi$ for all $\chi \in \mathbb{R}$, we have

$$S_{-\xi} = \begin{pmatrix} c_\xi\mathbb{1}_m & -s_\xi\mathbb{1}_m \\ -s_\xi\mathbb{1}_m & c_\xi\mathbb{1}_m \end{pmatrix}, \quad S_{U^\dagger} = \begin{pmatrix} U^T & \mathbb{0}_m \\ \mathbb{0}_m & U^\dagger \end{pmatrix}, \quad S_\zeta = \begin{pmatrix} c_\zeta\mathbb{1}_m & s_\zeta\mathbb{1}_m \\ s_\zeta\mathbb{1}_m & c_\zeta\mathbb{1}_m \end{pmatrix}. \tag{3.123}$$

With Eq. (3.122) we obtain

$$\begin{aligned}
\boldsymbol{V} &= \frac{1}{2}\begin{pmatrix} c_\xi\mathbb{1}_m & -s_\xi\mathbb{1}_m \\ -s_\xi\mathbb{1}_m & c_\xi\mathbb{1}_m \end{pmatrix}\begin{pmatrix} U^T & \mathbb{0}_m \\ \mathbb{0}_m & U^\dagger \end{pmatrix}\begin{pmatrix} c_\zeta\mathbb{1}_m & s_\zeta\mathbb{1}_m \\ s_\zeta\mathbb{1}_m & c_\zeta\mathbb{1}_m \end{pmatrix}\begin{pmatrix} c_\zeta\mathbb{1}_m & s_\zeta\mathbb{1}_m \\ s_\zeta\mathbb{1}_m & c_\zeta\mathbb{1}_m \end{pmatrix}\begin{pmatrix} U^* & \mathbb{0}_m \\ \mathbb{0}_m & U \end{pmatrix}\begin{pmatrix} c_\xi\mathbb{1}_m & -s_\xi\mathbb{1}_m \\ -s_\xi\mathbb{1}_m & c_\xi\mathbb{1}_m \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} c_\xi c_\zeta U^T - s_\xi s_\zeta U^\dagger & c_\xi s_\zeta U^T - s_\xi c_\zeta U^\dagger \\ -s_\xi c_\zeta U^T + c_\xi s_\zeta U^\dagger & -s_\xi s_\zeta U^T + c_\xi c_\zeta U^\dagger \end{pmatrix}\begin{pmatrix} c_\xi c_\zeta U^* - s_\xi s_\zeta U & -s_\xi c_\zeta U^* + c_\xi s_\zeta U \\ c_\xi s_\zeta U^* - s_\xi c_\zeta U & -s_\xi s_\zeta U^* + c_\xi c_\zeta U \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} [c_{2\xi}c_{2\zeta} - s_{2\xi}s_{2\zeta}\cos(2\phi)]\mathbb{1}_m & [-s_{2\xi}c_{2\zeta} + c_{2\xi}s_{2\zeta}\cos(2\phi)]\mathbb{1}_m + is_{2\zeta}\sin(2\phi)\Sigma \\ [-s_{2\xi}c_{2\zeta} + c_{2\xi}s_{2\zeta}\cos(2\phi)]\mathbb{1}_m - is_{2\zeta}\sin(2\phi)\Sigma & [c_{2\xi}c_{2\zeta} - s_{2\xi}s_{2\zeta}\cos(2\phi)]\mathbb{1}_m \end{pmatrix},
\end{aligned} \tag{3.124}$$

where in the third line we used $c_\chi^2+s_\chi^2=c_{2\chi}$, $2c_\chi s_\chi = s_{2\chi}$, and $c_\chi^2-s_\chi^2=1$, as well as $U=Oe^{i\phi\Sigma}$ with $O^TO=\mathbb{1}_m$ and $\Sigma^2=\mathbb{1}_m$, so that $U^\dagger U = U^T U^* = \mathbb{1}_m$, $U^TU=\cos(2\phi)\mathbb{1}_m + i\sin(2\phi)\Sigma$ and $U^\dagger U^* = \cos(2\phi)\mathbb{1}_m - i\sin(2\phi)\Sigma$. The matrix $\boldsymbol{V}+\frac{1}{2}\mathbb{1}_{2m}$ may thus be expressed as:

$$\frac{1}{2}\begin{pmatrix} [1+c_{2\xi}c_{2\zeta}-s_{2\xi}s_{2\zeta}\cos(2\phi)]\mathbb{1}_m & [-s_{2\xi}c_{2\zeta}+c_{2\xi}s_{2\zeta}\cos(2\phi)]\mathbb{1}_m + is_{2\zeta}\sin(2\phi)\Sigma \\ [-s_{2\xi}c_{2\zeta}+c_{2\xi}s_{2\zeta}\cos(2\phi)]\mathbb{1}_m - is_{2\zeta}\sin(2\phi)\Sigma & [1+c_{2\xi}c_{2\zeta}-s_{2\xi}s_{2\zeta}\cos(2\phi)]\mathbb{1}_m \end{pmatrix}. \tag{3.125}$$

With Eq. (3.120), we simply need to show that the inverse of the above matrix is

$$\mathbb{1}_{2m} - \begin{pmatrix}\mathbb{0}_m & \mathbb{1}_m \\ \mathbb{1}_m & \mathbb{0}_m\end{pmatrix}\begin{pmatrix} B^* & \mathbb{0}_m \\ \mathbb{0}_m & B\end{pmatrix} = \begin{pmatrix} \mathbb{1}_m & -B \\ -B^* & \mathbb{1}_m\end{pmatrix}. \tag{3.126}$$

A tedious but straightforward matrix multiplication with Eq. (3.125) concludes the proof, using $c_\chi^2-s_\chi^2=1$ and $\Sigma^2=\mathbb{1}_m$.

□

With Lemma 3.8 and Eq. (3.117) we have

$$\Pr_{\mathrm{CVS}}[\mathbf{0}] = \frac{1}{\pi^m\sqrt{\mathrm{Det}(\boldsymbol{V}+\mathbb{1}_{2m}/2)}}\,\mathrm{Haf}\begin{pmatrix} B_n^* & \mathbb{0}_n \\ \mathbb{0}_n & B_n\end{pmatrix}, \tag{3.127}$$

where

$$B_n = \frac{-s_{2\xi}c_{2\zeta}+c_{2\xi}s_{2\zeta}\cos(2\phi)}{1+c_{2\xi}c_{2\zeta}-s_{2\xi}s_{2\zeta}\cos(2\phi)}\mathbb{1}_n + i\frac{s_{2\zeta}\sin(2\phi)}{1+c_{2\xi}c_{2\zeta}-s_{2\xi}s_{2\zeta}\cos(2\phi)}\Sigma_n, \tag{3.128}$$

with Σ_n the $n\times n$ top left submatrix of Σ. Since the hafnian of a matrix does not depend on its diagonal entries, Eq. (3.127) can be rewritten as

$$\begin{aligned}\Pr_{\mathrm{CVS}}[\mathbf{0}] &= \frac{1}{\pi^m\sqrt{\mathrm{Det}(\boldsymbol{V}+\mathbb{1}_{2m}/2)}}\,\mathrm{Haf}\left[\frac{s_{2\zeta}\sin(2\phi)}{1+c_{2\xi}c_{2\zeta}-s_{2\xi}s_{2\zeta}\cos(2\phi)}\begin{pmatrix}-i\Sigma_n & \mathbb{0}_n \\ \mathbb{0}_n & i\Sigma_n\end{pmatrix}\right] \\ &= \frac{1}{\pi^m\sqrt{\mathrm{Det}(\boldsymbol{V}+\mathbb{1}_{2m}/2)}}\left[\frac{s_{2\zeta}\sin(2\phi)}{1+c_{2\xi}c_{2\zeta}-s_{2\xi}s_{2\zeta}\cos(2\phi)}\right]^n \mathrm{Haf}\begin{pmatrix}-i\Sigma_n & \mathbb{0}_n \\ \mathbb{0}_n & i\Sigma_n\end{pmatrix}.\end{aligned} \tag{3.129}$$

Now $\mathrm{Haf}(M\oplus N)=\mathrm{Haf}(M)\mathrm{Haf}(N)$, so the previous expression yields

$$\Pr_{\mathrm{CVS}}[\mathbf{0}] = \frac{1}{\pi^m\sqrt{\mathrm{Det}(\boldsymbol{V}+\mathbb{1}_{2m}/2)}}\left[\frac{s_{2\zeta}\sin(2\phi)}{1+c_{2\xi}c_{2\zeta}-s_{2\xi}s_{2\zeta}\cos(2\phi)}\right]^n \mathrm{Haf}(\Sigma_n)^2. \tag{3.130}$$

Finally, we compute $\mathrm{Det}(\boldsymbol{V}+\mathbb{1}_{2m}/2)$:

Lemma 3.9.

$$\mathrm{Det}(\boldsymbol{V}+\mathbb{1}_{2m}/2) = \frac{1}{2^m}\left[1+c_{2\xi}c_{2\zeta}-s_{2\xi}s_{2\zeta}\cos(2\phi)\right]^m. \tag{3.131}$$

Proof. From the proof of Lemma 3.8 we have

$$(\boldsymbol{V} + \mathbb{1}_{2m}/2)^{-1} = \begin{pmatrix} \mathbb{1}_m & -B \\ -B^* & \mathbb{1}_m \end{pmatrix}, \tag{3.132}$$

so that

$$\begin{aligned} \mathrm{Det}(\boldsymbol{V} + \mathbb{1}_{2m}/2) &= \frac{1}{\mathrm{Det}\begin{pmatrix} \mathbb{1}_m & -B \\ -B^* & \mathbb{1}_m \end{pmatrix}} \\ &= \frac{1}{\mathrm{Det}(\mathbb{1}_m - BB^*)}. \end{aligned} \tag{3.133}$$

Using the expression of the matrix B in Eq. (3.118) we obtain

$$\begin{aligned} BB^* &= \left(\frac{-s_{2\xi}c_{2\zeta} + c_{2\xi}s_{2\zeta}\cos(2\phi)}{1 + c_{2\xi}c_{2\zeta} - s_{2\xi}s_{2\zeta}\cos(2\phi)} \mathbb{1}_m + i \frac{s_{2\zeta}\sin(2\phi)}{1 + c_{2\xi}c_{2\zeta} - s_{2\xi}s_{2\zeta}\cos(2\phi)} \Sigma \right) \\ &\quad \times \left(\frac{-s_{2\xi}c_{2\zeta} + c_{2\xi}s_{2\zeta}\cos(2\phi)}{1 + c_{2\xi}c_{2\zeta} - s_{2\xi}s_{2\zeta}\cos(2\phi)} \mathbb{1}_m - i \frac{s_{2\zeta}\sin(2\phi)}{1 + c_{2\xi}c_{2\zeta} - s_{2\xi}s_{2\zeta}\cos(2\phi)} \Sigma \right) \\ &= \frac{\left[-s_{2\xi}c_{2\zeta} + c_{2\xi}s_{2\zeta}\cos(2\phi)\right]^2 + s_{2\zeta}^2 \sin^2(2\phi)}{\left[1 + c_{2\xi}c_{2\zeta} - s_{2\xi}s_{2\zeta}\cos(2\phi)\right]^2} \mathbb{1}_m, \end{aligned} \tag{3.134}$$

where we used $\Sigma^2 = \mathbb{1}_m$. Hence, with Eq. (3.133) we obtain

$$\begin{aligned} \mathrm{Det}(\boldsymbol{V} + \mathbb{1}_{2m}/2) &= \frac{1}{\mathrm{Det}(\mathbb{1}_m - BB^*)} \\ &= \frac{1}{\left[1 - \frac{\left[-s_{2\xi}c_{2\zeta} + c_{2\xi}s_{2\zeta}\cos(2\phi)\right]^2 + s_{2\zeta}^2\sin^2(2\phi)}{\left[1 + c_{2\xi}c_{2\zeta} - s_{2\xi}s_{2\zeta}\cos(2\phi)\right]^2}\right]^m} \\ &= \frac{\left[1 + c_{2\xi}c_{2\zeta} - s_{2\xi}s_{2\zeta}\cos(2\phi)\right]^{2m}}{\left[\left[1 + c_{2\xi}c_{2\zeta} - s_{2\xi}s_{2\zeta}\cos(2\phi)\right]^2 - \left[-s_{2\xi}c_{2\zeta} + c_{2\xi}s_{2\zeta}\cos(2\phi)\right]^2 - s_{2\zeta}^2\sin^2(2\phi)\right]^m} \\ &= \frac{1}{2^m}\left[1 + c_{2\xi}c_{2\zeta} - s_{2\xi}s_{2\zeta}\cos(2\phi)\right]^m. \end{aligned} \tag{3.135}$$

□

Combining Eq. (3.130) and Lemma 3.9, we finally obtain

$$\mathrm{Pr}_{\mathrm{CVS}}[\mathbf{0}] = \frac{2^{m/2} s_{2\zeta}^n \sin^n(2\phi)}{\pi^m \left[1 + c_{2\xi}c_{2\zeta} - s_{2\xi}s_{2\zeta}\cos(2\phi)\right]^{n+m/2}} \mathrm{Haf}(\Sigma_n)^2, \tag{3.136}$$

where $n = 2p$ is the number of single photons in the input.

■

Note that the matrix O appearing in the definition of the CVS circuit Eq. (3.108) does not

contribute to the output probability distribution. It provides additional degrees of freedom that may be useful for experimental considerations.

Note also that the expression of the prefactor $\kappa(\phi,\xi,\zeta)$ is left invariant when replacing ξ and ζ by $-\xi$ and $-\zeta$, which corresponds to changing which quadrature is squeezed both in input and output.

In the case of $\mathrm{CVS}_{\mathrm{SP}}$ circuits—with single photons as non-Gaussian inputs—the squeezing parameter ξ is equal to 0 and we have the following result, using $1+c_{2\zeta}=2c_\zeta^2$ and $s_{2\zeta}=2c_\zeta s_\zeta$:

Corollary 3.3. *We consider a CVS_{SP} circuit over m modes with $n=2p$ non-Gaussian input single photon states, evolution $U=Oe^{i\phi\Sigma}$ and unbalanced heterodyne detection $\zeta\in\mathbb{R}$. Then,*

$$Pr_{CVS_{SP}}[\mathbf{0}]=\kappa_{SP}(\phi,\zeta)\,\mathrm{Haf}\,(\Sigma_n)^2, \tag{3.137}$$

where Σ_n is the $n\times n$ top left submatrix of Σ and where

$$\kappa_{SP}(\phi,\zeta)=\frac{t_\zeta^n\sin^n(2\phi)}{\pi^m c_\zeta^m}, \tag{3.138}$$

with $t_\zeta=\tanh\zeta$ and $c_\zeta=\cosh\zeta$.

Next, we relate the output probability density evaluated at $(0,\ldots,0)$ of CVS circuits to the permanent of real matrices. Specifically, we provide an explicit construction holding for any real square matrix X.

Lemma 3.10. *Let $n=2p$ and let $X\in\mathbb{R}^{p\times p}$. For all $m\geq 2n$ and $\nu\leq 1/||X||$ there exists a matrix $\Sigma^X\in\mathcal{O}_S(M)$ such that its top left $n\times n$ submatrix is*

$$\Sigma_n^X=\nu\begin{pmatrix}0 & X\\ X^T & 0\end{pmatrix}. \tag{3.139}$$

Proof. Define $Y=\nu X$. The matrix $\mathbb{1}_p-Y^TY$ is symmetric positive semidefinite since $\|Y\|\leq 1$. It thus has a Cholesky decomposition $\mathbb{1}_p-Y^TY=Z^TZ$ for some square matrix Z. The columns of the $n\times p$ matrix

$$\begin{pmatrix}Y\\ Z\end{pmatrix} \tag{3.140}$$

form an orthonormal family that can be completed into an orthonormal basis of $\mathbb{R}^n$. The matrix obtained with these columns is orthogonal by construction and reads

$$\begin{pmatrix}Y & C\\ B^T & D\end{pmatrix}, \tag{3.141}$$

where B, C, D are $p \times p$ matrices. Finally, with the constraint $m \geq 2n$, setting

$$\Sigma^X = \begin{pmatrix} 0 & Y & 0 & C & 0 \\ Y^T & 0 & B & 0 & 0 \\ 0 & B^T & 0 & D & 0 \\ C^T & 0 & D^T & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbb{1}_{m-2n} \end{pmatrix} \tag{3.142}$$

yields an $m \times m$ symmetric orthogonal matrix—its columns are orthonormal by construction—which top left $n \times n$ submatrix is precisely given by Eq. (3.139).

■

Recall that a specific relation holds between the hafnian and the permanent. Namely, for any square matrix X, we have

$$\mathrm{Per}(X) = \mathrm{Haf}\begin{pmatrix} 0 & X \\ X^T & 0 \end{pmatrix}. \tag{3.143}$$

Using Lemma 3.7 with the matrix from Lemma 3.10, we get that for any square matrix X there exists a CVS circuit CVS_X which probability density at the origin reads:

$$\begin{aligned} \mathrm{Pr}_{\mathrm{CVS}_X}[\mathbf{0}] &= \kappa(\phi,\xi,\zeta)\,\mathrm{Haf}\left(\Sigma_n^X\right)^2 \\ &= \nu^n \kappa(\phi,\xi,\zeta)\left[\mathrm{Haf}\begin{pmatrix} 0 & X \\ X^T & 0 \end{pmatrix}\right]^2 \\ &= \nu^n \kappa(\phi,\xi,\zeta)\,\mathrm{Per}(X)^2, \end{aligned} \tag{3.144}$$

where $\nu \leq \frac{1}{\|X\|}$, and where

$$\kappa(\phi,\xi,\zeta) = \frac{2^{m/2} s_{2\zeta}^n \sin^n(2\phi)}{\pi^m \left[1 + c_{2\xi} c_{2\zeta} - s_{2\xi} s_{2\zeta} \cos(2\phi)\right]^{n+m/2}}. \tag{3.145}$$

By Theorem 28 of [AA13], multiplicative approximation of $\mathrm{Per}(X)^2$ is a #P-hard problem for real square matrices. Formally, for any $g \in [1, \mathrm{poly}\, n]$, the following problem is #P-hard: given a real matrix $X \in \mathbb{R}^{n\times n}$ such that $1/\|X\| \geq 2^{-\mathrm{poly(n)}}$, output a nonnegative real number P_X such that

$$\frac{\mathrm{Per}(X)^2}{g} \leq P_X \leq g\,\mathrm{Per}(X)^2. \tag{3.146}$$

The multiplying factor $\nu^n \kappa(\phi,\xi,\zeta)$ in Eq. (3.144) is finite and non-vanishing for some values of ξ, ζ and ϕ, so we obtain the following result:

Corollary 3.4. *For any $g \in [1, \mathrm{poly}\, n]$, the following problem is #P-hard: given a real matrix $X \in \mathbb{R}^{n\times n}$ such that $1/\|X\| \geq 2^{-poly(n)}$, output a nonnegative real number $\tilde{P}_X$ such that*

$$\frac{Pr_{CVS_X}[\mathbf{0}]}{g} \leq \tilde{P}_X \leq g Pr_{CVS_X}[\mathbf{0}]. \tag{3.147}$$

This is because by construction $\frac{\tilde{P}_X}{\nu^n \kappa(\phi,\xi,\zeta)}$ would then provide a multiplicative approximation of $\mathrm{Per}(X)^2$. As it turns out, this problem is easier to solve if one can perform weak simulation of CVS circuits classically:

Lemma 3.11. *Given access to a classical oracle which samples from the discretised output probability distribution of CVS circuits Pr^{η}_{CVS} of resolution η, for any $g \in [1, \mathrm{poly}\, n]$, the following problem can be solved in the third level of the polynomial hierarchy PH_3: given a real matrix $X \in \mathbb{R}^{n\times n}$, output a nonnegative real number $\tilde{P}_X$ such that*

$$\frac{Pr_{CVS_X}[\mathbf{0}]}{g} \leq \tilde{P}_X \leq g Pr_{CVS_X}[\mathbf{0}]. \tag{3.148}$$

By classical oracle, we mean here an oracle that takes a uniformly random input string as its only source of randomness (it has no built-in randomness as a quantum machine would). Note that we consider a classical oracle sampling from the discretised output probability distribution of CVS circuits $\Pr^{\eta}_{\mathrm{CVS}}$, rather than from the continuous probability density $\Pr_{\mathrm{CVS}}$. This is a strictly weaker oracle since one may obtain samples from $\Pr^{\eta}_{\mathrm{CVS}}$ using samples from $\Pr_{\mathrm{CVS}}$, with efficient classical post-processing.

Proof. With Eq. (3.111), the probability distribution for a CVS circuit with a finite resolution of the heterodyne detection $\eta \sim 2^{-\mathrm{poly}\, m}$, evaluated at $\mathbf{0}$ (in a slight abuse of notation we denote both the outcome and the corresponding discretised box by $\mathbf{0}$), reads:

$$\Pr^{\eta}_{\mathrm{CVS}}[\mathbf{0}] = \prod_{j=1}^{m} \int_{q_j=0}^{\eta} \int_{p_j=0}^{\eta} dq_j dp_j \Pr_{\mathrm{CVS}}[q_1, p_1, \ldots, q_m, p_m]. \tag{3.149}$$

Performing a Taylor expansion of the multivariate function $\boldsymbol{x} \mapsto \Pr_{\mathrm{CVS}}[\boldsymbol{x}]$ around the value $\mathbf{0} = (0, \ldots, 0)$, we obtain

$$\Pr_{\mathrm{CVS}}[\boldsymbol{x}] = \sum_{\boldsymbol{\gamma}\in\mathbb{N}^{2m}} \frac{\boldsymbol{x}^{\boldsymbol{\gamma}}}{\boldsymbol{\gamma}!} \partial^{\boldsymbol{\gamma}} \Pr_{\mathrm{CVS}}[\mathbf{0}]. \tag{3.150}$$

Plugging this expression in Eq. (3.149) and integrating we get

$$\begin{aligned} \Pr^{\eta}_{\mathrm{CVS}}[\mathbf{0}] &= \eta^{2m} \sum_{\boldsymbol{\gamma}\in\mathbb{N}^{2m}} \frac{\eta^{|\boldsymbol{\gamma}|}}{(\gamma_1+1)!\ldots(\gamma_{2m}+1)!} \partial^{\boldsymbol{\gamma}} \Pr_{\mathrm{CVS}}[\mathbf{0}] \\ &= \eta^{2m} \Pr_{\mathrm{CVS}}[\mathbf{0}] + \eta^{2m+1} \sum_{\substack{\boldsymbol{\gamma}\in\mathbb{N}^{2m} \\ |\boldsymbol{\gamma}|>0}} \frac{\eta^{|\boldsymbol{\gamma}|-1}}{(\gamma_1+1)!\ldots(\gamma_{2m}+1)!} \partial^{\boldsymbol{\gamma}} \Pr_{\mathrm{CVS}}[\mathbf{0}], \end{aligned} \tag{3.151}$$

so that

$$\frac{\Pr^{\eta}_{\mathrm{CVS}}[\mathbf{0}]}{\eta^{2m}} - \Pr_{\mathrm{CVS}}[\mathbf{0}] = \eta \sum_{\substack{\boldsymbol{\gamma}\in\mathbb{N}^{2m} \\ |\boldsymbol{\gamma}|>0}} \frac{\eta^{|\boldsymbol{\gamma}|-1}}{(\gamma_1+1)!\ldots(\gamma_{2m}+1)!} \partial^{\boldsymbol{\gamma}} \Pr_{\mathrm{CVS}}[\mathbf{0}]. \tag{3.152}$$

If η is small compared to $\Pr_{\mathrm{CVS}}[\mathbf{0}]$, a multiplicative approximation of $\Pr^{\eta}_{\mathrm{CVS}}[\mathbf{0}]/\eta^{2m}$ thus yields a multiplicative approximation of $\Pr_{\mathrm{CVS}}[\mathbf{0}]$.

We have $m = \operatorname{poly} n$ and $|\zeta| = \Omega(2^{-\operatorname{poly} m})$, by Eq. (3.109). When considering the circuit CVS_X associated to a real matrix X such that $1/\|X\| \geq 2^{-\operatorname{poly} m}$, we have $\Pr_{\mathrm{CVS}_X}[\mathbf{0}] = \Omega(2^{-\operatorname{poly} m})$ by Eq. (3.144). Hence, with $\eta \sim 2^{-\operatorname{poly} m}$, a multiplicative approximation of $\Pr^{\eta}_{\mathrm{CVS}_X}[\mathbf{0}]/\eta^{2m}$ is a multiplicative approximation of $\Pr_{\mathrm{CVS}_X}[\mathbf{0}]$.

We use Stockmeyer's approximate counting algorithm [Sto85] in order to conclude the proof: it is a classical algorithm which takes as input the classical description of a circuit sampling from a probability distribution and outputs a multiplicative approximation of the probability of a given outcome (see section 1.4.5). This algorithm sits in the third level of the polynomial hierarchy PH_3 and works as long as the probability to estimate is not superexponentially small, i.e., $o(2^{-\operatorname{poly} m})$ [LRKR17].

We have $\Pr^{\eta}_{\mathrm{CVS}}[\mathbf{0}]/\eta^{2m} = \Omega(2^{-\operatorname{poly} m})$, so with $\eta \sim 2^{-\operatorname{poly} m}$ the probability $\Pr^{\eta}_{\mathrm{CVS}}[\mathbf{0}]$ is not superexponentially small. Having at our disposal a classical oracle which samples from the probability distribution $\Pr^{\eta}_{\mathrm{CVS}}$ thus allows us to approximate multiplicatively the probability $\Pr^{\eta}_{\mathrm{CVS}}[\mathbf{0}]$ in the third level of the polynomial hierarchy, by making use of Stockmeyer's algorithm. Dividing the estimate obtained by η^{2m} finally yields a multiplicative approximation of $\Pr_{\mathrm{CVS}}[\mathbf{0}]$ in PH_3 (or rather in the class FPH_3 of search problems that may be solved by a PH_3 machine).

■

This result holds independently of the value of the squeezing parameter ξ, and when the detection parameter ζ satisfies $|\zeta| = \Omega(2^{-\operatorname{poly} m})$, i.e., even when the detection is very close to a balanced heterodyne detection. When $\zeta = 0$, however, the algorithm fails and the circuit is actually weakly simulable classically, because the output probability density factorises into products of single mode output probability densites, due to properties of balanced heterodyne detection. The same property will allow us to derive an efficient verification protocol for Boson Sampling and CVS circuits in the next chapter.

Combining Corollary 3.4 and Lemma 3.11 gives the main result of this section:

Theorem 3.3. *Sampling from the discretised output probability distribution of CVS circuits is classically hard, or the polynomial hierarchy collapses to its third level.*

Proof. Assuming that sampling from the discretised output probability distribution of CVS circuits can be done efficiently classically, Corollary 3.4 and Lemma 3.11 imply $\mathsf{P}^{\#\mathsf{P}} \subset \mathsf{PH}_3$ (where $\mathsf{P}^{\#\mathsf{P}}$ is the class of decision problems that can be solved efficiently using an oracle for the class of counting problems #P). On the other hand, by Toda's theorem [Tod91], $\mathsf{PH} \subset \mathsf{P}^{\#\mathsf{P}}$, so that $\mathsf{PH} \subset \mathsf{PH}_3$, i.e., the polynomial hierarchy collapses to its third level.

■

Theorem 3.3 implies that using enough non-Gaussian states as computational resources, weak simulation of Gaussian circuits is no longer classically efficient. This contrast with Theorem 3.2 from the previous section, i.e., the fact that strong simulation of Gaussian circuits with few non-Gaussian input states is classically efficient.

This statement is a *worst case* statement, i.e., there exists at least one CVS_X circuits which is hard to sample classically. In order to obtain an *average case* statement and identify a fraction of hard to sample CVS circuits, we define the *Real Gaussian Permanent Estimation* problem:

Problem 1 (Real Gaussian Permanent Estimation)**.** *Given as input a matrix $X \sim \mathcal{N}(0,1)_{\mathbb{R}}^{p\times p}$ of i.i.d. Gaussians together with error bounds $\epsilon, \delta > 0$, estimate $\mathrm{Per}(X)$ to within error $\pm\epsilon\cdot|\mathrm{Per}(X)|$, with probability at least $1-\delta$ over X, in $\mathrm{poly}(p, 1/\epsilon, 1/\delta)$ time.*

We can use the construction of Lemma 3.10 for the particular case of i.i.d. Gaussian matrices: for any $X \sim \mathcal{N}(0,1)_{\mathbb{R}}^{p\times p}$ of i.i.d. Gaussians, we obtain a circuit CVS_X such that Eq. ((3.144)) holds. Hence every instance of the RGPE is associated with a specific CVS circuit. In relation to the problem above, we introduce the *Permanent of Real Gaussians* Conjecture:

Conjecture 1 (Permanent of Real Gaussians)**.** *RGPE is #P-hard.*

We also introduce a second conjecture:

Conjecture 2 (Real Permanent Anti-Concentration)**.** *There exists a polynomial P such that for all p and $\delta > 0$,*

$$\Pr_{X\sim\mathcal{N}(0,1)_{\mathbb{R}}^{p\times p}}\left[|\mathrm{Per}(X)| < \frac{\sqrt{p!}}{P(p,1/\delta)}\right] < \delta. \tag{3.153}$$

This problem and these conjectures are precisely the real version of the *Gaussian Permanent Estimation* problem and the *Permanent-of-Gaussians* and *Permanent Anti-Concentration* conjectures introduced in [AA13]. This leads us to our average case hardness result.

Theorem 3.4. *Assuming Conjecture 2 is true, classical circuits sampling from the (discretised) probability distribution of CVS circuits can be used to solve Real Gaussian Permanent Estimation in the third level of the polynomial hierarchy. Assuming Conjecture 1 is also true, an efficient classical weak simulation of CVS_X circuits, where $X \sim \mathcal{N}(0,1)_{\mathbb{R}}^{p\times p}$, would imply a collapse of the polynomial hierarchy to its third level.*

Proof. With the same proof as Lemma 3.11, with $\eta = O(2^{-\mathrm{poly}\, m})$, classical circuits sampling from the (discretised) probability distribution of CVS circuits can be used to obtain a multiplicative approximation of $\mathrm{Pr}_{\mathrm{CVS}}[0,\ldots,0]$ in the third level of the polynomial hierarchy PH_3 by means of Stockmeyer algorithm. In particular, for $X \sim \mathcal{N}(0,1)_{\mathbb{R}}^{p\times p}$ a square matrix which entries are i.i.d. Gaussians and considering the circuit CVS_X, we obtain multiplicative approximation of $\mathrm{Per}(X)^2$.

RGPE however refers to estimating $\mathrm{Per}(X)$ rather than $\mathrm{Per}(X)^2$. It is easy to see that a multiplicative approximation of $\mathrm{Per}(X)^2$ can be turned into a multiplicative approximation $|\mathrm{Per}(X)|$ by taking the square root of the estimate. Then, in the case of real matrices, only the sign of the permanent remains to be determined.

A more general version of this question has been addressed in [AA13] where they showed that (the complex version of) Conjecture 2 allowed one to estimate the phase of $\mathrm{Per}(X)$ from multiplicative approximation of $|\mathrm{Per}(X)|^2$, for X i.i.d. *complex* Gaussian matrix. It implies in particular that Conjecture 2 allows one to determine the sign of $\mathrm{Per}(X)$ from $\mathrm{Per}(X)^2$ if X is i.i.d. *real* Gaussian matrix. Hence, assuming Conjecture 2, RGPE can be solved in the third level of the polynomial hierarchy using a classical circuit sampling from the output probability distribution of a CVS circuit as an oracle.

Assuming Conjecture 1 is true, RGPE is #P-hard. With the above, the existence of an efficient classical algorithm which approximates multiplicatively the output distribution of CVS_X circuits implies the existence of a classical algorithm sitting in the third level of the polynomial hierarchy able to solve a #P-hard problem. This in turn yields a collapse of the polynomial hierarchy to the third level, thanks to Toda's theorem [Tod91].

■

This result is an average case statement, i.e., it implies that a circuit CVS_X, where $X \sim \mathcal{N}(0,1)_{\mathbb{R}}^{p\times p}$, is hard to sample with high probability over X, assuming Conjectures 1 and 2 are true. Once again, we assumed the existence of a classical oracle sampling from the discretised output probability distribution of CVS circuits $\mathrm{Pr}_{\mathrm{CVS}}^{\eta}$, rather than the continuous probability density $\mathrm{Pr}_{\mathrm{CVS}}$. However, one may obtain samples from $\mathrm{Pr}_{\mathrm{CVS}}^{\eta}$ using samples from $\mathrm{Pr}_{\mathrm{CVS}}$, with efficient classical post-processing.

3.4 Discussion and open problems

We have considered various notions of classical simulation and have studied the transition from classically simulable models to models that are universal for quantum computing for continuous variables.

We have studied the case of adaptive linear optics, an intermediate model between Boson Sampling [AA13] and the Knill–Laflamme–Milburn scheme for universal quantum computing [KLM01], obtaining classical algorithms for both probability estimation and overlap estimation and analysing their running times. The conclusion to be drawn from our study is that achieving a quantum advantage for either probability estimation or overlap estimation using linear optics, input single photons and adaptive measurements, is challenging.

A quantum advantage is not ruled out for probability estimation only if the number of adaptive measurements scale at least logarithmically in the size of the interferometer. The challenge

posed by the implementation of a quantum algorithm with adaptive linear optics for probability estimation beyond classical capabilities thus comes from the number of adaptive measurements needed.

For overlap estimation, a quantum advantage is not ruled out for a constant number adaptive measurements, but many overlaps are easy to estimate classically in that case. It is only when a significant fraction of the input photons is detected at the stage of the adaptive measurements that a quantum advantage becomes possible. The challenge posed by the implementation of a quantum algorithm with adaptive linear optics for overlap estimation beyond classical capabilities thus comes from the need of photon number-resolving detection and the preparation of many photon number states.

For strong simulation, we have considered general Gaussian circuits with Gaussian measurements and non-Gaussian inputs and we have given sufficient conditions in terms of non-Gaussian resources for an efficient classical strong simulation. We have defined the G_{core} circuits, a broad family of Gaussian circuits supplemented with non-Gaussian input states, where the non-Gaussian states are multimode core states. We have identified various subclasses of these circuits:

- The Interleaved Photon-Added Gaussian circuits (IPAG), which are circuits that sample with Gaussian measurements from states which can be engineered from the vacuum using multimode Gaussian unitary operations and a finite number of photon additions.
- The G_{Fock} circuits, which are Gaussian circuits supplemented with Fock states in the input.
- The CVS_{PA}/CVS_{PS}/CVS_{SP} circuits, which are specific interferometers with unbalanced heterodyne detection, supplemented with photon-added squeezed states/photon-added squeezed states/single photons in the input.

The relation between these continuous variable quantum computational models is summarised as

$$\text{CVS}_{SP} \subset \text{CVS}_{PA} = \text{CVS}_{PS} \subset G_{\text{Fock}} \subset \text{IPAG} \subset G_{\text{core}}, \tag{3.154}$$

from the smallest class of circuits to the largest. The tools developped in this chapter also allows us to consider Gaussian circuits supplemented with non-Gaussian states and photon counters, by writing the photon counting POVM element as $|n\rangle\langle n| = \frac{1}{n!}(\hat{a}^\dagger)^n |0\rangle\langle 0| \hat{a}^n$, for $n \in \mathbb{N}$ and commuting the creation operators to the input through the Gaussian computation. Classical algorithms simulating this type of computational model have been derived recently [QA20].

For weak simulation, we have proven the computational hardness of a sampling problem that stems from the family of CVS circuits, relating their discretised output probability density to the permanent of real matrices. Introducing equivalent conjectures to those of [AA13] for real matrices, we have extended the hardness result to an average case hardness.

With this collection of results comes various related open problems:

One of the main outstanding problems is to prove the hardness of approximately sampling from CVS circuits. Following [AA13], this may involve making conjectures about anticoncentration and average case hardness of the loop hafnian rather than the permanent, as well as collecting evidence and ultimately proving these conjectures. These conjectures have already been extended from the permanent to the hafnian for the Gaussian Boson Sampling proposal [HKS$^+$16, KHS$^+$19].

A related problem is to prove the hardess of sampling from CVS circuits with a binning resolution which either scales as $\frac{1}{\text{poly}\, m}$ or is constant with respect to the number of modes, since an exponentially small resolution is not experimentally realistic.

Comparing more precisely IPAG and IPSG circuit families would give insight on the differences between photon addition and photon subtraction in the multimode case.

Whether the set of output states of IPAG circuits is dense in the set of all multimode states (the multimode equivalent of Lemma 2.6 from the previous chapter) is also an interesting question. In other words, is it possible to approximate with arbitrary precision (in trace distance) any multimode quantum state using only single photon additions and Gaussian unitary operations?

Another main open problem, which we solve in the next chapter, is the verification of the output of CVS circuits and Boson Sampling, necessary to a proper demonstration of quantum supremacy with these computational models.

inprepaVerifBS

CHAPTER 4

CERTIFICATION OF CONTINUOUS VARIABLE QUANTUM STATES

Out of the many properties featured by quantum physics, the impossibility to perfectly determine an unknown state [DY96] is specially interesting. This property is at the heart of quantum cryptography protocols such as quantum key distribution [BB84a]. On the other hand, it makes certification of the correct functioning of quantum devices a challenge, since the output of such devices can only be determined approximately, through repeated measurements over numerous copies of the output states. The involved configurations spaces have enormous dimensions, a serious burden for any characterization. What is more, certification comes along with an ironic twist: it is highly non-trivial in light of the fact that certain quantum computations are expected to exponentially outperform any attempt at classically solving the same problem. Determining an unknown state is difficult especially for continuous variable quantum states, which are described by possibly infinitely many complex parameters.

In this chapter, after introducing known methods for the characterisation of continuous variable quantum states, we develop new methods using heterodyne measurement in both the trusted and untrusted settings.

Firstly, based on quantum state tomography with heterodyne detection, we introduce a reliable method for continuous variable quantum state certification, which directly yields the elements of the density matrix of the state considered with analytical confidence intervals. This method requires neither mathematical reconstruction of the data nor discrete binning of the sample space, and uses a single Gaussian measurement setting, namely heterodyne detection.

Secondly, beyond quantum state tomography and without its identical copies assumption, we promote our reliable tomography method to an efficient protocol for verifying single-mode continuous variable pure quantum states with Gaussian measurements against fully malicious adversaries, i.e., making no assumptions whatsoever on the state generated by the adversary.

Thirdly, we generalise the previous protocols to the multimode case and obtain efficient protocols for verifying a large class of multimode continuous variable quantum states, with and without the identical copies assumption. In particular, we show how to efficiently verify the output state of a Boson Sampling experiment with a single-mode Gaussian measurement, thus enabling a proper demonstration of quantum supremacy with Boson Sampling.

This chapter is based on [EHW+20, CDG+19, CRW+20, CGKM20].

4.1 Building trust for a continuous variable quantum state

With rapidly developing quantum technologies for communication, simulation, computation and sensing, the ability to assess the correct functioning of quantum devices is of major importance, for near-term systems, the so-called noisy intermediate-scale quantum devices [Pre18], and for the more sophisticated devices. Depending on the desired level of trust and in particular the assumptions one is ready to make, several methods are available for certifying the output of quantum devices [EHW+20]. A common assumption is that the outcomes of the tested quantum device are *independent and identically distributed* (i.i.d.) over various uses of the device. This implies in particular that the conclusions drawn from test runs are also valid for future computational runs with the same device.

In the following, the task of checking the output state of a quantum device is denoted *tomography* for state independent methods, when i.i.d. behaviour is assumed, *certification* for a given a target state, when i.i.d. behaviour is assumed, and *verification* for a given target state, with no assumption whatsoever, and in particular without the i.i.d. assumption.

4.1.1 Tomography, certification and verification

Quantum state tomography [DPS03] is an important technique which aims at reconstructing a good approximation of the output state of a quantum device by performing multiple rounds of measurements on several copies of said output states. Given an ensemble of identically prepared systems, with measurement outcomes from the same observable, one can build up a histogram, from which a probability density can be estimated. According to Born's rule, this probability density is the square modulus of the state coefficients, taken in the basis corresponding to the measurement. However, a single measurement setting cannot yield the full state information since the phase of its coefficients are then lost. Many sets of measurements on many subensembles must be performed and combined to reconstruct the density matrix of the state. The data do not yield the state directly, but rather indirectly through data analysis. Quantum state tomography commonly assumes an i.i.d. behaviour for the device, i.e., that the density matrix of the output state considered is the same at each round of measurement. This assumption may be relaxed with a tradeoff in the efficiency of the protocol [CR12].

A certification task corresponds to a setting where one wants to benchmark an industrial quantum device, or check the output of a physical experiment. On the other hand, a verification task corresponds to a cryptographic scenario, where the device to be tested is untrusted, or the quantum data is given by a potentially malicious party, for example in the context of delegated quantum computing. In the latter case, the task of quantum verification is to ensure that either the device behaved properly, or the computation aborts with high probability. While delegated computing is a natural platform for the emerging quantum devices, one can provide a physical interpretation to this adversarial setting by emphasising that we aim for deriving verification schemes that make no assumptions whatsoever about the noise model of the underlying systems. Various methods for verification of quantum devices have been investigated, in particular for discrete variable quantum information [GKK19], and they provide different efficiencies and security parameters depending on the computational power of the verifier. The common feature for all these approaches is to utilise some basic obfuscation scheme that allows one to reduce the problem of dealing with a fully general noise model, or a fully general adversarial deviation of the device, to a simple error detection scheme [Vid18].

For continuous variable quantum devices, checking that the output state is close to a target state may be done with linear optics using optical homodyne tomography [LR09]. This method allows one to reconstruct the Wigner function of a generic state using only Gaussian measurements, namely homodyne detection. Because of the continuous character of its outcomes, one must proceed to a discrete binning of the sample space, in order to build probability histograms. Then, the state representation in phase space is determined by a mathematical reconstruction.

For cases where we have a specific target state, more efficient options are possible. For multimode Gaussian states, more efficient certification methods have been derived with Gaussian measurements [AGKE15]. These methods involve the computation of a fidelity witness, i.e., a lower bound on the fidelity, from the measured samples. The cubic phase state certification protocol of [LDT$^+$18] also introduces a fidelity witness, and is an example of certification of a specific non-Gaussian state with Gaussian measurements, which assumes an i.i.d. state preparation. The verification protocol for Gaussian continuous variable weighted hypergraph states of [TMM$^+$19] removes this assumption, again for this specific family of states.

4.1.2 General single-mode protocol

We address two main issues in what follows. First, existing continuous variable state tomography methods are not reliable in the sense of [CR12], because errors coming from the reconstruction procedure are indistinguishable from errors coming from the data. Second, there is no Gaussian verification protocol for non-Gaussian states without i.i.d. assumption.

We thus introduce a general *receive-and-measure* protocol for building trust for single-mode continuous variable quantum states, using solely Gaussian measurements, namely heterodyne

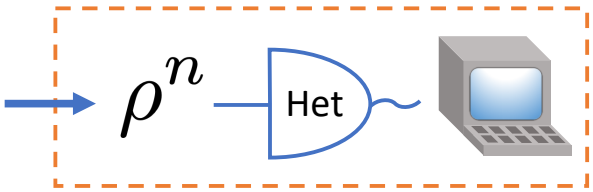

Figure 4.1: A schematic representation of the protocol. The tester (within the dashed rectangle) receives a continuous variable quantum state ρ^n over n subsystems. This state could be for example the outcome of n successive runs of a physical experiment, the output of a commercial quantum device, or directly sent by some untrusted quantum server. The tester measures with heterodyne detection some of the subsystems of ρ^n and uses the samples obtained and efficient classical post-processing to deduce information about the remaining subsystems.

detection (see section 1.4.2 and [FOP05, TMJ$^+$17]). This protocol allows us to perform reliable continuous variable quantum state tomography based on heterodyne detection, which we refer to as *heterodyne tomography* in what follows. This tomography technique only requires a single fixed measurement setting, compared to homodyne tomography. This protocol also provides a means for certifying single-mode continuous variable quantum states, under the i.i.d. assumption. Finally, the same protocol also allows us to verify single-mode continuous variable quantum states, without the i.i.d. assumption. For these three applications, the measurements performed are the same. It is only the selection of subsystems to be measured and the classical post-processing performed that differ from one application to another.

The structure of the protocol is depicted in Fig. 6.1: given a quantum state ρ^n over n subsystems, measure some of the subsystems with balanced heterodyne detection. Then, post-process the samples obtained to retrieve information about the remaining subsystems. We show in the following sections how this protocol may be used to perform reliable tomography, certification and verification of single-mode continuous variable quantum states, and we detail the corresponding choices of subsystems and the classical post-processing for each task.

4.2 Heterodyne estimator

In this section, we introduce a generalisation of the optical equivalence theorem for antinormal ordering [CG69a], which provides an estimator for the expected value of an operator acting on a state with bounded support over the Fock basis, from samples of heterodyne detection of the state. From this result, we derive various protocols in the following sections, ranging from state tomography to state verification.

We denote by $\mathop{\mathbb{E}}_{\alpha\leftarrow D}[f(\alpha)]$ the expected value of a function f for samples drawn from a distribution

D. Let us introduce for $k,l \geq 0$ the polynomials

$$\mathscr{L}_{k,l}(z) = e^{zz^*} \frac{(-1)^{k+l}}{\sqrt{k!}\sqrt{l!}} \frac{\partial^{k+l}}{\partial z^k \partial z^{*l}} e^{-zz^*}, \tag{4.1}$$

for $z \in \mathbb{C}$, which are, up to a normalisation, the Laguerre 2D polynomials, appearing in particular in the expressions of Wigner function of Fock states [Wün98]. For any operator $A = \sum_{k,l=0}^{+\infty} A_{kl} |k\rangle\langle l|$ and all $E \in \mathbb{N}$, we define with these polynomials the function

$$f_A(z,\eta) = \frac{1}{\eta} e^{\left(1-\frac{1}{\eta}\right)zz^*} \sum_{k,l=0}^{E} \frac{A_{kl}}{\sqrt{\eta^{k+l}}} \mathscr{L}_{k,l}\left(\frac{z}{\sqrt{\eta}}\right), \tag{4.2}$$

for all $z \in \mathbb{C}$, and all $0 < \eta < 1$. We omit the dependency in E for brevity. The function $z \mapsto f_A(z,\eta)$, being a polynomial multiplied by a converging Gaussian function, is bounded over $\mathbb{C}$. With the same notations, we also define the following constant:

$$K_A = \sum_{k,l=0}^{E} |A_{kl}| \sqrt{(k+1)(l+1)}. \tag{4.3}$$

Theorem 4.1. *Let $E \in \mathbb{N}$ and let $0 < \eta < \frac{2}{E}$. Let also $A = \sum_{k,l=0}^{+\infty} A_{kl} |k\rangle\langle l|$ be an operator and let $\rho = \sum_{k,l=0}^{E} \rho_{kl} |k\rangle\langle l|$ be a density operator with bounded support. Then,*

$$\left| \mathrm{Tr}\left(A\rho\right) - \mathop{\mathbb{E}}_{\alpha \leftarrow Q_\rho} [f_A(\alpha,\eta)] \right| \leq \eta K_A, \tag{4.4}$$

where the function f and the constant K are defined in Eqs. (4.2) and (4.3).

The function f_A defined in Eq. (4.2) is, up to a numerical factor of π, a bounded approximation of the Glauber–Sudarshan function P_A of the operator A. This approximation is parametrised by a precision η, and a cutoff value E. The optical equivalence theorem for antinormal ordering reads (see section 1.2 and [CG69a])

$$\mathrm{Tr}(A\rho) = \pi \int_{\alpha \in \mathbb{C}} Q_\rho(\alpha) P_A(\alpha) d^2\alpha. \tag{4.5}$$

Given that

$$\mathop{\mathbb{E}}_{\alpha \leftarrow Q_\rho} [f_A(\alpha,\eta)] = \int_{\alpha \in \mathbb{C}} Q_\rho(\alpha) f_A(\alpha,\eta) d^2\alpha, \tag{4.6}$$

we would expect that $\mathop{\mathbb{E}}_{\alpha \leftarrow Q_\rho} [f_A(\alpha,\eta)]$ is an approximation of $\mathrm{Tr}(A\rho)$ parametrised by η and E. Theorem 4.1 makes this statement more precise. We prove this theorem in what follows.

Proof. With Eq. (4.2) we obtain

$$\left| \mathrm{Tr}\left(A\rho\right) - \mathop{\mathbb{E}}_{\alpha \leftarrow Q_\rho} [f_A(\alpha,\eta)] \right| = \left| \sum_{k,l=0}^{+\infty} A_{lk} \mathrm{Tr}\left(|l\rangle\langle k| \rho\right) - \sum_{k,l=0}^{E} A_{lk} \mathop{\mathbb{E}}_{\alpha \leftarrow Q_\rho} [f_{|l\rangle\langle k|}(\alpha,\eta)] \right|$$

$$
\begin{aligned}
&= \left| \sum_{k,l=0}^{E} A_{lk} \left(\mathrm{Tr}\left(|l\rangle\langle k|\,\rho\right) - \mathop{\mathbb{E}}_{\alpha\leftarrow Q_\rho} [f_{|l\rangle\langle k|}(\alpha,\eta)] \right) \right| \\
&\leq \sum_{k,l=0}^{E} |A_{lk}| \left| \mathrm{Tr}\left(|l\rangle\langle k|\,\rho\right) - \mathop{\mathbb{E}}_{\alpha\leftarrow Q_\rho} [f_{|l\rangle\langle k|}(\alpha,\eta)] \right|,
\end{aligned} \tag{4.7}
$$

where we used in the second line the fact that ρ has a bounded support over the Fock basis. This shows that it is sufficient to prove the Theorem for $A = |l\rangle\langle k|$, for all k,l from 0 to E. We first introduce the following result:

Lemma 4.1. *For all* $0 \leq k,l \leq E$*,*

$$
\mathop{\mathbb{E}}_{\alpha\leftarrow Q_\rho} [f_{|l\rangle\langle k|}(\alpha,\eta)] = \rho_{kl} + \sum_{\substack{m>k,n>l \\ m-n=k-l}}^{E} \rho_{mn}\, \eta^{\frac{m+n-k-l}{2}} \sqrt{\binom{m}{k}\binom{n}{l}}. \tag{4.8}
$$

Proof. Let us fix k,l in $0,\ldots,E$. By Eqs. (4.1) and (4.2) we have, for all $z \in \mathbb{C}$,

$$
\begin{aligned}
f_{|l\rangle\langle k|}(z) &= \left(\frac{1}{\eta}\right)^{1+\frac{k+l}{2}} e^{\left(1-\frac{1}{\eta}\right)zz^*}\, \mathscr{L}_{l,k}\left(\frac{z}{\sqrt{\eta}}\right) \\
&= \left(\frac{1}{\eta}\right)^{1+\frac{k+l}{2}} e^{zz^*} \frac{(-1)^{k+l}}{\sqrt{k!}\sqrt{l!}} \frac{\partial^{k+l}}{\partial u^{*k} \partial u^{l}} e^{-uu^*}\Big|_{u=\frac{z}{\sqrt{\eta}}} \\
&= \frac{1}{\eta} e^{\left(1-\frac{1}{\eta}\right)zz^*} \sum_{p=0}^{\min(k,l)} \frac{(-1)^p \sqrt{k!}\sqrt{l!}}{p!(k-p)!(l-p)!} \left(\frac{1}{\eta}\right)^{k+l-p} z^{k-p} z^{*l-p}.
\end{aligned} \tag{4.9}
$$

Moreover, for all $\alpha \in \mathbb{C}$,

$$
\begin{aligned}
Q_\rho(\alpha) &= \frac{1}{\pi} \langle\alpha|\rho|\alpha\rangle \\
&= \frac{1}{\pi} \sum_{m,n=0}^{E} \rho_{mn} \langle\alpha|m\rangle \langle n|\alpha\rangle \\
&= \frac{1}{\pi} \sum_{m,n=0}^{E} \rho_{mn} \frac{\alpha^{*m}\alpha^{n}}{\sqrt{m!n!}} e^{-|\alpha|^2}.
\end{aligned} \tag{4.10}
$$

Combining these expressions we obtain

$$
\begin{aligned}
\mathop{\mathbb{E}}_{\alpha\leftarrow Q_\rho} [f_{|l\rangle\langle k|}(\alpha,\eta)] &= \int_{\alpha\in\mathbb{C}} Q_\rho(\alpha) f_{|l\rangle\langle k|}(\alpha,\eta)\, d^2\alpha \\
&= \frac{1}{\pi} \sum_{m,n=0}^{E} \frac{\rho_{mn}}{\sqrt{m!n!}} \int_{\alpha\in\mathbb{C}} \alpha^{*m}\alpha^{n} e^{-|\alpha|^2} f_{|l\rangle\langle k|}(\alpha,\eta)\, d^2\alpha \\
&= \frac{1}{\pi\eta} \sum_{m,n=0}^{E} \rho_{mn} \frac{\sqrt{k!}\sqrt{l!}}{\sqrt{m!}\sqrt{n!}} \sum_{p=0}^{\min(k,l)} \frac{(-1)^p}{p!(k-p)!(l-p)!} \left(\frac{1}{\eta}\right)^{k+l-p} \int_{\alpha\in\mathbb{C}} \alpha^{k+n-p} \alpha^{*(l+m-p)} e^{-\frac{1}{\eta}|\alpha|^2} d^2\alpha.
\end{aligned} \tag{4.11}
$$

Setting $\alpha = re^{i\theta}$, we have $d^2\alpha = rdrd\theta$ and the integral on the last line may be computed as

$$\begin{aligned}\int_{\alpha\in\mathbb{C}} \alpha^{k+n-p}\alpha^{*(l+m-p)}e^{-\frac{1}{\eta}|\alpha|^2}d^2\alpha &= \int_0^{+\infty} r^{k+l+m+n-2p+1}e^{-\frac{r^2}{\eta}}dr \int_0^{2\pi} e^{i(k+n-l-m)\theta}d\theta \\ &= \begin{cases} \pi\left(\frac{k+l+m+n}{2}-p\right)!\eta^{\frac{k+l+m+n}{2}-p+1} & \text{for } k-l=m-n, \\ 0 & \text{for } k-l\neq m-n, \end{cases}\end{aligned} \tag{4.12}$$

where we used $\int_0^{+\infty} r^{2t+1}e^{-\frac{r^2}{\eta}} = \frac{1}{2}t!\eta^{t+1}$ for $t = \frac{k+l+m+n}{2}-p$, which is obtained directly by induction and integration by parts (note that for $k-l=m-n$, and $p\leq\min(k,l)$, we have indeed $t\in\mathbb{N}$). Hence,

$$\begin{aligned}\mathop{\mathbb{E}}_{\alpha\leftarrow Q_\rho}[f_{|l\rangle\langle k|}(\alpha,\eta)] &= \sum_{\substack{m,n=0\\ m-n=k-l}}^{E}\rho_{mn}\frac{\sqrt{k!}\sqrt{l!}}{\sqrt{m!}\sqrt{n!}}\sum_{p=0}^{\min(k,l)}\frac{(-1)^p\left(\frac{k+l+m+n}{2}-p\right)!}{p!(k-p)!(l-p)!}\eta^{\frac{m+n-k-l}{2}} \\ &= \sum_{\substack{m,n=0\\ m-n=k-l}}^{E}\rho_{mn}\eta^{\frac{m+n-k-l}{2}}\frac{\left(\frac{k+l+m+n}{2}\right)!}{\sqrt{m!}\sqrt{n!}\sqrt{k!}\sqrt{l!}}\sum_{p=0}^{\min(k,l)}(-1)^p\frac{\binom{k}{p}\binom{l}{p}}{\binom{\frac{k+l+m+n}{2}}{p}}.\end{aligned} \tag{4.13}$$

Now for $k\leq l$ we have, for all $q\in\mathbb{N}$ (see, e.g., result 7.1 of [Gou72]),

$$\sum_{p=0}^{k}(-1)^p\frac{\binom{k}{p}\binom{l}{p}}{\binom{q}{p}} = \begin{cases}\frac{\binom{q-l}{k}}{\binom{q}{k}} & \text{for } q\geq k+l, \\ 0 & \text{for } q<k+l.\end{cases} \tag{4.14}$$

When $k\leq l$, Eq (4.13) thus yields

$$\begin{aligned}\mathop{\mathbb{E}}_{\alpha\leftarrow Q_\rho}[f_{|l\rangle\langle k|}(\alpha,\eta)] &= \sum_{\substack{m,n=0\\ m-n=k-l\\ m+n\geq k+l}}^{E}\rho_{mn}\eta^{\frac{m+n-k-l}{2}}\frac{\left(\frac{k+l+m+n}{2}\right)!}{\sqrt{m!}\sqrt{n!}\sqrt{k!}\sqrt{l!}}\frac{\binom{\frac{k+l+m+n}{2}-l}{k}}{\binom{\frac{k+l+m+n}{2}}{k}} \\ &= \sum_{\substack{m\geq k,n\geq l\\ m-n=k-l}}^{E}\rho_{mn}\eta^{\frac{m+n-k-l}{2}}\frac{1}{\sqrt{m!}\sqrt{n!}\sqrt{k!}\sqrt{l!}}\frac{\left(\frac{k-l+m+n}{2}\right)!\left(\frac{-k+l+m+n}{2}\right)!}{\left(\frac{-k-l+m+n}{2}\right)!} \\ &= \sum_{\substack{m\geq k,n\geq l\\ m-n=k-l}}^{E}\rho_{mn}\eta^{\frac{m+n-k-l}{2}}\frac{\sqrt{m!}\sqrt{n!}}{\sqrt{k!}\sqrt{l!}\sqrt{(m-k)!}\sqrt{(n-l)!}} \\ &= \sum_{\substack{m\geq k,n\geq l\\ m-n=k-l}}^{E}\rho_{mn}\eta^{\frac{m+n-k-l}{2}}\sqrt{\binom{m}{k}\binom{n}{l}},\end{aligned} \tag{4.15}$$

where we used that within the summation $m-n=k-l$. This formula is also valid for $l \leq k$, with the same reasoning. We finally obtain, for any k,l in $0,\ldots,E$

$$\begin{aligned} \mathop{\mathbb{E}}_{\alpha \leftarrow Q_\rho}[f_{|l\rangle\langle k|}(\alpha,\eta)] &= \sum_{\substack{m\geq k, n\geq l \\ m-n=k-l}}^{E} \rho_{mn} \eta^{\frac{m+n-k-l}{2}} \sqrt{\binom{m}{k}\binom{n}{l}} \\ &= \rho_{kl} + \sum_{\substack{m> k, n> l \\ m-n=k-l}}^{E} \rho_{mn} \eta^{\frac{m+n-k-l}{2}} \sqrt{\binom{m}{k}\binom{n}{l}}. \end{aligned} \tag{4.16}$$

□

Using Lemma 4.1, we obtain

$$\begin{aligned} \left| \mathrm{Tr}(|l\rangle\langle k|\,\rho) - \mathop{\mathbb{E}}_{\alpha \leftarrow Q_\rho}[f_{|l\rangle\langle k|}(\alpha,\eta)] \right| &= \left| \rho_{kl} - \mathop{\mathbb{E}}_{\alpha \leftarrow Q_\rho}[f_{|l\rangle\langle k|}(\alpha,\eta)] \right| \\ &= \left| \sum_{\substack{m> k, n> l \\ m-n=k-l}}^{E} \rho_{mn} \eta^{\frac{m+n-k-l}{2}} \sqrt{\binom{m}{k}\binom{n}{l}} \right| \\ &\leq \sum_{\substack{m> k, n> l \\ m-n=k-l}}^{E} |\rho_{mn}| \eta^{\frac{m+n-k-l}{2}} \sqrt{\binom{m}{k}\binom{n}{l}} \\ &= \sum_{s=1}^{E-\max(k,l)} |\rho_{s+k,s+l}| \eta^s \sqrt{\binom{s+k}{k}\binom{s+l}{l}} \\ &\leq \sum_{s=1}^{E-\max(k,l)} \eta^s \sqrt{\binom{s+k}{k}\binom{s+l}{l}} \sqrt{\rho_{s+k,s+k}} \sqrt{\rho_{s+l,s+l}}, \end{aligned} \tag{4.17}$$

where we set $s = m-k = n-l = \frac{m+n-k-l}{2}$ in the third line, and where we used $|\rho_{s+k,s+l}| \leq \sqrt{\rho_{s+k,s+k}}\sqrt{\rho_{s+l,s+l}}$ in the last line, since ρ is a positive semidefinite matrix. In order to obtain an upper bound independent of ρ, we now show for all s that $\eta^s \sqrt{\binom{s+k}{k}\binom{s+l}{l}} \leq \eta\sqrt{(k+1)(l+1)}$ for $\eta \leq \frac{2}{E}$. For all k,l in $0,\ldots,E$ and for all s in $2,\ldots,E-\max(k,l)$, we have

$$\frac{\sqrt{s+k}\sqrt{s+l}}{s} \leq \frac{E}{2}. \tag{4.18}$$

This in turn implies that for all s in $2,\ldots,E-\max(k,l)$

$$\begin{aligned} \eta^s \sqrt{\binom{s+k}{k}\binom{s+l}{l}} &= \eta \frac{\sqrt{(s+k)(s+l)}}{s} \eta^{s-1} \sqrt{\binom{s-1+k}{k}\binom{s-1+l}{l}} \\ &\leq \frac{\eta E}{2} \eta^{s-1} \sqrt{\binom{s-1+k}{k}\binom{s-1+l}{l}} \end{aligned} \tag{4.19}$$

$$\leq \eta^{s-1} \sqrt{\binom{s-1+k}{k}\binom{s-1+l}{l}},$$

since we assumed $\eta \leq \frac{2}{E}$. Hence by induction, for all s in $2,\dots,E-\max(k,l)$,

$$\eta^s \sqrt{\binom{s+k}{k}\binom{s+l}{l}} \leq \eta^1 \sqrt{\binom{1+k}{k}\binom{1+l}{l}} = \eta\sqrt{(k+1)(l+1)}. \tag{4.20}$$

Combining this with Eq. (4.17) yields

$$\begin{aligned} \left|\mathrm{Tr}(|l\rangle\langle k|\,\rho) - \mathop{\mathbb{E}}_{\alpha\leftarrow Q_\rho}[f_{|l\rangle\langle k|}(\alpha,\eta)]\right| &\leq \eta\sqrt{(k+1)(l+1)} \sum_{s=1}^{E-\max(k,l)} \sqrt{\rho_{s+k,s+k}}\sqrt{\rho_{s+l,s+l}} \\ &\leq \eta\sqrt{(k+1)(l+1)} \sqrt{\sum_{s=1}^{E-\max(k,l)} \rho_{s+k,s+k} \sum_{s=1}^{E-\max(k,l)} \rho_{s+l,s+l}} \\ &\leq \eta\sqrt{(k+1)(l+1)}, \end{aligned} \tag{4.21}$$

for all k,l in $0,\dots,E$, where we used Cauchy-Schwarz inequality and the fact that $\mathrm{Tr}(\rho)=1$. Note that the above bound still holds when $E \to +\infty$. Together with Eq. (4.7) we obtain

$$\begin{aligned} \left|\mathrm{Tr}\left(A\rho\right) - \mathop{\mathbb{E}}_{\alpha\leftarrow Q_\rho}[f_A(\alpha,\eta)]\right| &\leq \eta \sum_{k,l=0}^{E} |A_{kl}|\sqrt{(k+1)(l+1)} \\ &= \eta K_A, \end{aligned} \tag{4.22}$$

by Eq. (4.3).

■

This result provides an estimator for the expected value of any operator A acting on a continuous variable state ρ with bounded support over the Fock basis. This estimator is the expected value of a bounded function f_A over samples drawn from the Husimi Q function of ρ. This probability density corresponds to a Gaussian measurement of ρ, namely heterodyne detection (see section 1.4.2). The right hand side of Eq. (4.4) is an energy bound, which depends on the operator A and the value E.

When the operator A is the density matrix of a continuous variable pure state $|\psi\rangle$, the previous estimator approximates the fidelity $F(\psi,\rho) = \langle\psi|\rho|\psi\rangle$ between $|\psi\rangle\langle\psi|$ and ρ. With the same notations:

Corollary 4.1. *Let $E \in \mathbb{N}$ and let $0 < \eta < \frac{2}{E}$. Let also $|\psi\rangle\langle\psi| = \sum_{k,l=0}^{+\infty} \psi_k \psi_l^* |k\rangle\langle l|$ be a normalised pure state and let $\rho = \sum_{k,l=0}^{E} \rho_{kl} |k\rangle\langle l|$ be a density operator with bounded support. Then,*

$$\left|F\left(\psi,\rho\right) - \mathop{\mathbb{E}}_{\alpha\leftarrow Q_\rho}[f_\psi(\alpha,\eta)]\right| \leq \eta K_\psi \leq \frac{\eta}{2}(E+1)(E+2), \tag{4.23}$$

where the function f_A and the constant K_A are defined in Eqs. (4.2) and (4.3), for $A = |\psi\rangle\langle\psi|$.

Proof. We apply Theorem 4.1 for $A = |\psi\rangle\langle\psi|$ a pure state. We obtain

$$\begin{aligned}\left|\langle\psi|\rho|\psi\rangle - \mathop{\mathbb{E}}_{\alpha\leftarrow Q_\rho}[f_\psi(\alpha,\eta)]\right| &\leq \eta K_\psi \\ &= \eta\sum_{k,l=0}^{E}|\psi_k\psi_l|\sqrt{(k+1)(l+1)} \\ &= \eta\left(\sum_{n=0}^{E}|\psi_n|\sqrt{n+1}\right)^2 \\ &\leq \eta\sum_{n=0}^{E}|\psi_n|^2\sum_{n=0}^{E}(n+1) \\ &\leq \frac{\eta}{2}(E+1)(E+2),\end{aligned} \tag{4.24}$$

where we used Cauchy-Schwarz inequality, and $\sum_{n=0}^{E}|\psi_n|^2 \leq \mathrm{Tr}(|\psi\rangle\langle\psi|) = 1$. Since $|\psi\rangle$ is a pure state, we have $F(\psi,\rho) = \langle\psi|\rho|\psi\rangle$, which concludes the proof.

■

This result provides an estimator for the fidelity between any target pure state $|\psi\rangle$ and any continuous variable (mixed) state ρ with bounded support over the Fock basis. This estimator is the expected value of a bounded function f_ψ over samples drawn from the probability density corresponding to heterodyne detection of ρ. The right hand side of Eq. (4.23) is an energy bound, which may be refined depending on the expression of $|\psi\rangle$. In particular, the second bound is independent of the target state $|\psi\rangle$. The assumption of bounded support makes sense for tomography, where the energy range of the measured state is known, but not necessarily in a more adversarial setting.

Given these results, one may choose a target pure state $|\psi\rangle$ and measure with heterodyne detection various copies of the output (mixed) state ρ of a quantum device with bounded support over the Fock basis. Then, using the samples obtained, one may estimate the expected value of f_ψ, thus obtaining an estimate of the fidelity between the states $|\psi\rangle\langle\psi|$ and ρ. Using this result, we introduce a reliable method for performing continuous variable quantum state tomography using heterodyne detection.

4.3 Reliable heterodyne tomography

Continuous variable quantum state tomography methods usually make two assumptions: firstly that the measured states are independent identical copies (i.i.d. assumption, for *independently and identically distributed*), and secondly that the measured states have a bounded support over the Fock basis [LR09]. With the same assumptions, we present a reliable method for state tomography with heterodyne detection which has the advantage of providing analytical confidence intervals. Our method directly provides estimates of the elements of the state density

matrix, phase included. As such, neither mathematical reconstruction of the phase, nor binning of the sample space is needed, since the samples are used only to compute expected values of bounded functions. Moreover, only a single fixed Gaussian measurement setting is needed, namely heterodyne detection (Fig. 1.4).

The law of large numbers ensures that the sample average from independently and identically distributed (i.i.d.) random variables converges to the expected value of these random variables, when the number of samples goes to infinity. The following key lemma refines this statement and quantifies the speed of convergence:

Lemma 4.2. ***(Hoeffding inequality)*** *Let $\lambda > 0$, let $n \geq 1$, let $z_1, \ldots, z_n$ be i.i.d. complex random variables from a probability density D over $\mathbb{R}$, and let $f : \mathbb{C} \mapsto \mathbb{R}$ such that $|f(z)| \leq M$, for $M > 0$ and all $z \in \mathbb{C}$. Then*

$$\Pr\left[\left|\frac{1}{n}\sum_{i=1}^{n} f(z_i) - \mathop{\mathbb{E}}_{z\leftarrow D}[f(z)]\right| \geq \lambda\right] \leq 2\exp\left[-\frac{n\lambda^2}{2M^2}\right]. \tag{4.25}$$

This comes directly from Hoeffding inequality [Hoe63] applied to the real bounded i.i.d. random variables $f(z_1), \ldots, f(z_N)$. When dealing with complex random variables, we use the following result instead:

Lemma 4.3. ***(Hoeffding inequality for complex random variables)*** *Let $\lambda > 0$, let $n \geq 1$, let $z_1, \ldots, z_n$ be i.i.d. complex random variables from a probability density D over $\mathbb{C}$, and let $f : \mathbb{C} \mapsto \mathbb{C}$ such that $|f(z)| \leq M$, for $M > 0$ and all $z \in \mathbb{C}$. Then*

$$\Pr\left[\left|\frac{1}{n}\sum_{i=1}^{n} f(z_i) - \mathop{\mathbb{E}}_{z\leftarrow D}[f(z)]\right| \geq \lambda\right] \leq 4\exp\left[-\frac{n\lambda^2}{4M^2}\right]. \tag{4.26}$$

Proof. For all $a > 0$ and all $z \in \mathbb{C}$, $|z| = \sqrt{\Re(z)^2 + \Im(z)^2} \geq a$ implies $|\Re(z)| \geq a/\sqrt{2}$ or $|\Im(z)| \geq a/\sqrt{2}$. Hence,

$$\Pr[|z| \geq a] \leq \Pr\left[|\Re(z)| \geq \frac{a}{\sqrt{2}}\right] + \Pr\left[|\Im(z)| \geq \frac{a}{\sqrt{2}}\right], \tag{4.27}$$

so applying twice Lemma 4.2 for the real random variables $\Re(f(z))$ and $\Im(f(z))$, respectively, yields Lemma 4.3.

■

For tomographic application, all copies of the state are measured. For $n \geq 1$, let $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ be samples from heterodyne detection of n copies of a quantum state ρ. For $\epsilon > 0$ and $k, l \in \mathbb{N}$, we define

$$\rho^{\epsilon}_{kl} = \frac{1}{n}\sum_{i=1}^{n} f_{|l\rangle\langle k|}\left(\alpha_i, \frac{\epsilon}{K_{|l\rangle\langle k|}}\right), \tag{4.28}$$

where the function f_A and the constant K_A are defined in Eqs. (4.2) and (4.3), for $A = |l\rangle\langle k|$, and where $\epsilon > 0$ is a free parameter. The quantity ρ^{ϵ}_{kl} is the average of the function $f_{|l\rangle\langle k|}$ over the samples $\alpha_1, \ldots, \alpha_n$. The next result shows that this estimator approximates the matrix element k,l of this state with high probability. We use the notations of Theorem 4.1.

Theorem 4.2 (Reliable heterodyne tomography)**.** *Let $\epsilon, \epsilon' > 0$, let $n \geq 1$, and let $\alpha_1, \ldots, \alpha_n$ be samples obtained by measuring with heterodyne detection n copies of a state $\rho = \sum_{k,l=0}^{E} \rho_{kl} |k\rangle\langle l|$ with bounded support, for $E \in \mathbb{N}$. Then*

$$\left|\rho_{kl} - \rho^{\epsilon}_{kl}\right| \leq \epsilon + \epsilon', \tag{4.29}$$

for all $0 \leq k,l \leq E$, with probability greater than

$$1 - 4 \sum_{0 \leq k \leq l \leq E} \exp\left[-\frac{n\epsilon^{2+k+l}\epsilon'^2}{4C_{kl}}\right], \tag{4.30}$$

where the estimate ρ^{ϵ}_{kl} is defined in Eq. (4.28), and where

$$C_{kl} := [(k+1)(l+1)]^{1+\frac{k+l}{2}}\, 2^{|l-k|} \binom{\max(k,l)}{\min(k,l)} \tag{4.31}$$

is a constant independent of ρ.

Proof. In order to prove Theorem 4.2, we apply Lemma 4.3 to the functions $z \mapsto f_{|l\rangle\langle k|}(z,\eta)$ defined in Eq. (4.2). We first bound these functions:

Lemma 4.4. *For all $k,l \geq 0$, define*

$$M_{kl} := \sqrt{2^{|l-k|} \binom{\max(k,l)}{\min(k,l)}}. \tag{4.32}$$

Then for all k,l and all $z \in \mathbb{C}$,

$$\left|f_{|k\rangle\langle l|}(z,\eta)\right| \leq \frac{M_{kl}}{\eta^{1+\frac{k+l}{2}}}. \tag{4.33}$$

Proof. For k or $l > E$ the inequality is trivial. For all $k,l \leq E$ and all $z \in \mathbb{C}$,

$$\begin{aligned} \left|f_{|k\rangle\langle l|}(z,\eta)\right| &= \left(\frac{1}{\eta}\right)^{1+\frac{k+l}{2}} e^{\left(1-\frac{1}{\eta}\right)|z|^2} \left|\mathscr{L}_{k,l}\left(\frac{z}{\sqrt{\eta}}\right)\right| \\ &= \frac{1}{\eta} e^{\left(1-\frac{1}{\eta}\right)|z|^2} \frac{1}{\sqrt{k!}\sqrt{l!}} \left|\sum_{p=0}^{\min(k,l)} \frac{(-1)^p k! l!}{p!(k-p)!(l-p)!} \frac{1}{\eta^{k+l-p}} z^{l-p} z^{*(k-p)}\right|, \end{aligned} \tag{4.34}$$

where we used Eq. (4.1). Now for all $z \in \mathbb{C}^*$ and all $a > 0$ we have [Wün98]

$$\begin{aligned}\left|\sum_{p=0}^{\min(k,l)} \frac{(-1)^p k!l!}{p!(k-p)!(l-p)!} a^{k+l-p} z^{l-p} z^{*(k-p)}\right| &= a^k l! |z|^{k-l} \left|L_l^{(k-l)}\left(a|z|^2\right)\right| \\ &= a^l k! |z|^{l-k} \left|L_k^{(l-k)}\left(a|z|^2\right)\right|,\end{aligned} \tag{4.35}$$

where

$$L_n^{(\alpha)}(x) = \sum_{q=0}^{n} \frac{(-1)^q}{q!} \binom{n+\alpha}{n-q} x^q \tag{4.36}$$

are the generalised Laguerre polynomials [AS65], defined for $\alpha \in \mathbb{R}$ and $n \in \mathbb{N}$. Plugging this relation into Eq. (4.34) we obtain

$$\begin{aligned}\left|f_{|k\rangle\langle l|}(z,\eta)\right| &= e^{\left(1-\frac{1}{\eta}\right)|z|^2} \frac{|z|^{l-k}}{\eta^{1+l}} \frac{\sqrt{k!}}{\sqrt{l!}} \left|L_k^{(l-k)}\left(\frac{|z|^2}{\eta}\right)\right| \\ &= e^{\left(1-\frac{1}{\eta}\right)|z|^2} \frac{|z|^{k-l}}{\eta^{1+k}} \frac{\sqrt{l!}}{\sqrt{k!}} \left|L_l^{(k-l)}\left(\frac{|z|^2}{\eta}\right)\right|,\end{aligned} \tag{4.37}$$

for all $z \in \mathbb{C}$. The generalised Laguerre polynomials are bounded as [Roo85]

$$\left|L_n^{(\alpha)}(x)\right| \leq \frac{\Gamma(n+\alpha+1)}{n!\Gamma(\alpha+1)} e^{\frac{x}{2}}, \tag{4.38}$$

for all $x \geq 0$, all $\alpha \geq 0$ and all $n \in \mathbb{N}$, and as

$$\left|L_n^{(\alpha)}(x)\right| \leq 2^{-\alpha} e^{\frac{x}{2}}, \tag{4.39}$$

for all $x \geq 0$, all $\alpha \leq -\frac{1}{2}$ and all $n \in \mathbb{N}$.

Let $a > 0$. Assuming $k < l$, we have $|z|^{l-k} \leq a^{l-k}$ for $|z| \leq a$, and $|z|^{k-l} \leq a^{k-l}$ for $|z| \geq a$. Thus, the first line of Eq. (4.37), together with Eq. (4.38), give

$$\begin{aligned}\left|f_{|k\rangle\langle l|}(z,\eta)\right| &\leq e^{\left(1-\frac{1}{\eta}\right)|z|^2} \frac{a^{l-k}}{\eta^{1+l}} \frac{\sqrt{k!}}{\sqrt{l!}} \frac{l!}{k!(l-k)!} e^{\frac{|z|^2}{2\eta}} \\ &\leq \frac{a^{l-k}}{\eta^{1+l}} \frac{\sqrt{l!}}{(l-k)!\sqrt{k!}},\end{aligned} \tag{4.40}$$

for $|z| \leq a$ and $k < l$. Similarly, the second line of Eq. (4.37), together with Eq. (4.39), give

$$\begin{aligned}\left|f_{|k\rangle\langle l|}(z,\eta)\right| &\leq e^{\left(1-\frac{1}{\eta}\right)|z|^2} \frac{a^{k-l}}{\eta^{1+k}} \frac{\sqrt{l!}}{\sqrt{k!}} 2^{l-k} e^{\frac{|z|^2}{2\eta}} \\ &\leq \frac{a^{k-l}}{\eta^{1+k}} \frac{\sqrt{l!}}{\sqrt{k!}} 2^{l-k},\end{aligned} \tag{4.41}$$

for $|z| \geq a$ and $k < l$.

These two last bounds in Eqs. (4.40) and (4.41) are equal for $a^{l-k} = (2\eta)^{\frac{l-k}{2}}\sqrt{(l-k)!}$, yielding the bound

$$\left|f_{|k\rangle\langle l|}(z,\eta)\right| \leq \sqrt{\frac{2^{l-k}}{\eta^{2+k+l}}\binom{l}{k}}, \tag{4.42}$$

for all $z \in \mathbb{C}$ and $k < l$. For $l < k$ the same reasoning gives

$$\left|f_{|k\rangle\langle l|}(z,\eta)\right| \leq \sqrt{\frac{2^{k-l}}{\eta^{2+k+l}}\binom{k}{l}}. \tag{4.43}$$

Finally, for $k = l$ the previous bounds also hold, by combining Eqs. (4.37) and (4.38), and this proves the lemma.

□

Let $k, l \geq 0$, $n \in \mathbb{N}$ and $\epsilon' > 0$. Applying Lemma 4.3 to the function $f_{|l\rangle\langle k|}$, with the bound from Lemma 4.4 yields

$$\Pr\left[\left|\frac{1}{n}\sum_{i=1}^{n} f_{|l\rangle\langle k|}(\alpha_i,\eta) - \mathop{\mathbb{E}}_{\alpha\leftarrow Q_\rho}[f_{|l\rangle\langle k|}(\alpha,\eta)]\right| \geq \epsilon'\right] \leq 4\exp\left[-\frac{n\eta^{2+k+l}\epsilon'^2}{4M_{kl}^2}\right]. \tag{4.44}$$

Applying Theorem 4.1 for $A = |l\rangle\langle k|$ we also obtain

$$\left|\rho_{kl} - \mathop{\mathbb{E}}_{\alpha\leftarrow Q_\rho}\left[f_{|l\rangle\langle k|}(\alpha,\eta)\right]\right| \leq \eta\sqrt{k+1}\sqrt{l+1}. \tag{4.45}$$

Let $\alpha_1, \ldots, \alpha_n$ be samples from the Q function of ρ. Combining Eqs. (4.44) and (4.45), we obtain with the triangular inequality

$$\left|\rho_{kl} - \frac{1}{n}\sum_{i=1}^{n} f_{|l\rangle\langle k|}(\alpha_i,\eta)\right| \leq \eta\sqrt{k+1}\sqrt{l+1} + \epsilon', \tag{4.46}$$

with probability greater than

$$1 - 4\exp\left[-\frac{n\eta^{2+k+l}\epsilon'^2}{4M_{kl}^2}\right]. \tag{4.47}$$

We have $K_{|l\rangle\langle k|} = \sqrt{(k+1)(l+1)}$ by Eq. (4.3). Taking $\eta = \frac{\epsilon}{K_{|l\rangle\langle k|}}$ yields

$$\left|\rho_{kl} - \frac{1}{n}\sum_{i=1}^{n} f_{|l\rangle\langle k|}\left(\alpha_i, \frac{\epsilon}{K_{|l\rangle\langle k|}}\right)\right| \leq \epsilon + \epsilon', \tag{4.48}$$

with probability greater than

$$1 - 4\exp\left[-\frac{n\epsilon^{2+k+l}\epsilon'^2}{4C_{kl}}\right], \tag{4.49}$$

where we defined

$$\begin{aligned} C_{kl} &:= [(k+1)(l+1)]^{1+\frac{k+l}{2}} M_{kl}^2 \\ &= [(k+1)(l+1)]^{1+\frac{k+l}{2}} 2^{|l-k|} \binom{\max(k,l)}{\min(k,l)}. \end{aligned} \tag{4.50}$$

Now this holds for $0 \leq k,l \leq E$. Together with the union bound, this proves the theorem.

■

In light of this result, the principle for performing reliable heterodyne tomography is straightforward and as follows: n identical copies $\rho^{\otimes n}$ of the output quantum state of a physical experiment or quantum device are measured with heterodyne detection, yielding the values $\alpha_1, \ldots, \alpha_n$. These values are used to compute the estimates ρ_{kl}^{ϵ}, defined in Eq. (4.28), for all k,l in the range of energy of the experiment. Then, Theorem 4.2 directly provides confidence intervals for all these estimates of ρ_{kl}, the matrix elements of the density operator ρ, without the need for a binning of the sample space or any additional data reconstruction, using a single measurement setting. For a desired precision ϵ and a failure probability δ, the number of samples needed scales as $n = \text{poly}(1/\epsilon, \log(1/\delta))$.

Both homodyne and heterodyne quantum state tomography assume a bounded support over the Fock basis for the output state considered, i.e., that all matrix elements are equal to zero beyond a certain value, and that the output quantum states are i.i.d., i.e., that all measured output states are independent and identical. While these assumptions are natural when looking at the output of a physical experiment, corresponding to a noisy partially trusted quantum device with bounded energy, they may be questionable in the context of untrusted devices. We remove these assumptions in what follows: we first drop the bounded support assumption, deriving a certification protocol for continuous variable quantum states of an i.i.d. device with heterodyne detection ; then, we drop both assumptions, deriving a general verification protocol for continuous variable quantum states against an adversary who can potentially be fully malicious.

4.4 Continuous variable quantum state certification protocol

Given an untrusted source of quantum states, the purpose of state certification and state verification protocols is to check whether if its output state is close to a given target state, or far from it. To achieve this, a verifier tests the output state of the source. Ideally, one would like to obtain an upper bound on the probability that the state is not close from the target state, given that it passed a test. However, this is known to be impossible without prior knowledge of the tested state distribution [GKK19]. Indeed, writing this conditional probability

$$\Pr[\text{incorrect}|\text{accept}] = \frac{\Pr[\text{incorrect} \cap \text{accept}]}{\Pr[\text{accept}]}, \tag{4.51}$$

in a situation where the device always produces a bad output state, it is rejected by the verifier's test most of the time, so the acceptance probability is very small while the conditional probability

is equal to 1. Therefore, the quantity that will always be bounded in certification and verification protocols in which one does not have prior knowledge of the device is the joint probability that the tested state is not close to the target state *and* that it passes the test. Equivalently, we obtain lower bounds on the probability that the tested state is close to the target state or that it fails the test.

We first consider the certification of the output of an i.i.d. quantum device, i.e., which output state is the same at each round. However, we do not assume that the output states of the device have bounded support over the Fock basis anymore. This is instead ensured probabilistically using the samples from heterodyne detection.

Let us define the following operators for $E \geq 0$:

$$U = \sum_{n=E+1}^{+\infty} |n\rangle\langle n| = 1 - \Pi_E, \tag{4.52}$$

where $\Pi_E = \sum_{n=0}^{E} |n\rangle\langle n|$ is the projector onto the Hilbert space $\bar{\mathscr{H}}$ of states with less than E photons, and

$$T = \frac{1}{\pi} \int_{|\alpha|^2 \geq E} |\alpha\rangle\langle\alpha| \, d^2\alpha, \tag{4.53}$$

where $|\alpha\rangle$ is a coherent state. We have the following result, proven in [LGPRC13] by expanding T in the Fock basis:

$$U \leq 2T. \tag{4.54}$$

In particular,

$$\mathrm{Tr}(U\rho) \leq 2\,\mathrm{Tr}(T\rho). \tag{4.55}$$

The probability P_r that exactly r among n values of $|\alpha_i|^2$ are bigger than E and $n-r$ values are lower, and that the projection Π_E of the state ρ onto the Hilbert space $\bar{\mathscr{H}}$ of states with less than E photons fails is bounded as

$$\begin{aligned}
P_r &= \binom{n}{r} \mathrm{Tr}\left[(1-\Pi_E) T^r (1-T)^{n-r} \rho^{\otimes n+1}\right] \\
&= \binom{n}{r} \mathrm{Tr}\left[U T^r (1-T)^{n-r} \rho^{\otimes n+1}\right] \\
&\leq 2\binom{n}{r} \mathrm{Tr}\left(T\rho\right)^{r+1} \mathrm{Tr}\left[(1-T)\rho\right]^{n-r} \\
&\leq 2\binom{n}{r} \max_p \left| p^{r+1}(1-p)^{n-r} \right| \\
&= 2\binom{n}{r} \left(\frac{r+1}{n+1}\right)^{r+1} \left(1 - \frac{r+1}{n+1}\right)^{n-r}
\end{aligned} \tag{4.56}$$

$$
\begin{aligned}
&\leq \frac{2n^r}{r!}\left(\frac{r+1}{n+1}\right)^{r+1}\left(1-\frac{r+1}{n+1}\right)^{n-r} \\
&\leq \frac{2n^r}{r!}\frac{(r+1)^{r+1}}{n^{r+1}}\exp\left[-\frac{(n-r)(r+1)}{n+1}\right] \\
&\leq \frac{2}{n}\frac{r+1}{\sqrt{2\pi(r+1)}}\exp\left[\frac{(r+1)^2}{n+1}\right] \\
&\leq \frac{\sqrt{r+1}}{n}\exp\left[\frac{(r+1)^2}{n+1}\right],
\end{aligned}
$$

where we used Eq. (4.55), $1-x\leq e^{-x}$ and $(r+1)!\geq\sqrt{2\pi(r+1)}(r+1)^{r+1}e^{-(r+1)}$. For $s\in\mathbb{N}$, and for all $r\leq s$,

$$
\frac{\sqrt{r+1}}{n}\exp\left[\frac{(r+1)^2}{n+1}\right]\leq\frac{\sqrt{s+1}}{n}\exp\left[\frac{(s+1)^2}{n+1}\right], \tag{4.57}
$$

hence the probability that at most s among n values of $|\alpha_i|^2$ are bigger than E, and that the projection Π_E of the state ρ onto the Hilbert space $\bar{\mathscr{H}}$ of states with less than E photons fails is bounded by

$$
P^{iid}_{\text{support}} := \frac{(s+1)^{3/2}}{n}\exp\left[\frac{(s+1)^2}{n+1}\right]. \tag{4.58}
$$

For $1\ll s\ll n$, this implies that either ρ is contained in a lower dimensional subspace, or the score at the support estimation step is higher than s, with high probability.

Our continuous variable quantum state certification protocol is then as follows: let $|\psi\rangle$ be a target pure state, of which one wants to certify m copies. The values s and E are free parameters of the protocol. One instructs the i.i.d. device to prepare $n+m$ copies of $|\psi\rangle$, and the device outputs an i.i.d. (mixed) state $\rho^{\otimes n+m}$. One keeps m copies $\rho^{\otimes m}$, and measures the n others with heterodyne detection, obtaining the samples $\alpha_1,\ldots,\alpha_n$. One records the number r of samples such that $|\alpha_i|^2>E$. We refer to this step as *support estimation*. For a given $\epsilon>0$, one also computes with the same samples the estimate

$$
F_\psi(\rho)=\left[\frac{1}{n}\sum_{i=1}^{n}f_\psi\left(\alpha_i,\frac{\epsilon}{mK_\psi}\right)\right]^m, \tag{4.59}
$$

where the function f_A and the constant K_A are defined in Eqs. (4.2) and (4.3), for $A=|\psi\rangle\langle\psi|$, and where $\epsilon>0$ is a free parameter. Note that the support estimation step is no longer necessary if the target state has a bounded support over the Fock basis.

The next result quantifies how close this estimate is from the fidelity between the remaining m copies of the output state $\rho^{\otimes m}$ of the tested device and m copies of the target state $|\psi\rangle\langle\psi|^{\otimes m}$.

Theorem 4.3 (Gaussian certification of continuous variable quantum states)**.** *Let $\epsilon,\epsilon'>0$, let $s\leq n$, and let $\alpha_1,\ldots,\alpha_n$ be samples obtained by measuring with heterodyne detection n copies of a state ρ. Let E in $\mathbb{N}$, and let r be the number of samples such that $|\alpha_i|^2>E$. Let also $|\psi\rangle$ be a pure state. Then for all $m\in\mathbb{N}^*$,*

$$
\left|F(\psi^{\otimes m},\rho^{\otimes m})-F_\psi(\rho)\right|\leq\epsilon+\epsilon', \tag{4.60}
$$

or $r > s$, with probability greater than

$$1 - \left(P^{iid}_{Support} + P^{iid}_{Hoeffding} \right), \tag{4.61}$$

where

$$P^{iid}_{Support} = \frac{(s+1)^{3/2}}{n} \exp\left[\frac{(s+1)^2}{n+1} \right], \tag{4.62}$$

$$P^{iid}_{Hoeffding} = 2\exp\left[-\frac{n\epsilon^{2+2E}\epsilon'^2}{2m^{4+2E} C_\psi^2} \right], \tag{4.63}$$

where the estimate $F_\psi(\rho)$ is defined in Eq. (4.59), and where

$$C_\psi = \sum_{k,l=0}^{E} |\psi_k \psi_l| \left(\frac{\epsilon}{m} \right)^{E - \frac{k+l}{2}} K_\psi^{1+\frac{k+l}{2}} \sqrt{2^{|l-k|} \binom{\max(k,l)}{\min(k,l)}} \tag{4.64}$$

is a constant independent of ρ, with the constant K defined in Eq. (4.3).

In order to prove this theorem we make use of the following simple result:

Lemma 4.5. *Let $\eta > 0$ and $a, b \in [0,1]$ such that $|a-b| \leq \eta$. Then for all $m \geq 1$,*

$$\left| a^m - b^m \right| \leq m|a-b| \leq m\eta. \tag{4.65}$$

Proof. With the notations of the lemma,

$$\begin{aligned} \left| a^m - b^m \right| &= |a-b| \left| \sum_{j=0}^{m-1} a^j b^{m-j-1} \right| \\ &\leq m|a-b| \\ &\leq m\eta. \end{aligned} \tag{4.66}$$

■

We first consider the case of $m = 1$ from which we deduce the general case with the lemma.

Proof. Let us write $|\psi\rangle = \sum_{n\geq 0} \psi_n |n\rangle$. For $\eta > 0$, the function $z \mapsto f_\psi(z,\eta)$ is real-valued, since $|\psi\rangle\langle\psi|$ is hermitian. It is bounded as

$$\begin{aligned} \left| f_\psi(\alpha,\eta) \right| &= \left| \sum_{k,l=0}^{E} \psi_k \psi_l^* f_{|k\rangle\langle l|}(\alpha,\eta) \right| \\ &\leq \sum_{k,l=0}^{E} \left| \psi_k \psi_l^* f_{|k\rangle\langle l|}(\alpha,\eta) \right| \end{aligned}$$

$$\begin{aligned} &\leq \sum_{k,l=0}^{E} \left|\psi_k \psi_l\right| \frac{M_{kl}}{\eta^{1+\frac{k+l}{2}}} \\ &= \frac{1}{\eta^{1+E}} \sum_{k,l=0}^{E} \left|\psi_k \psi_l\right| \eta^{E-(k+l)/2} M_{kl} \\ &= \frac{M_\psi(\eta)}{\eta^{1+E}}, \end{aligned} \tag{4.67}$$

where we used Lemma 4.4, and where we defined

$$M_\psi(\eta) := \sum_{k,l=0}^{E} \left|\psi_k \psi_l\right| \eta^{E-(k+l)/2} M_{kl}. \tag{4.68}$$

Applying Lemma 4.2 to the real-valued function $z \mapsto f_\psi(z,\eta)$ thus yields

$$\Pr\left[\left|\frac{1}{n}\sum_{i=1}^{n} f_\psi(\alpha_i,\eta) - \mathop{\mathbb{E}}_{\alpha \leftarrow Q_\rho}[f_\psi(\alpha,\eta)]\right| \geq \epsilon'\right] \leq 2\exp\left[-\frac{n\eta^{2+2E}\epsilon'^2}{2M_\psi^2(\eta)}\right], \tag{4.69}$$

for $\epsilon', \eta > 0$, where the probability is over i.i.d. samples from heterodyne detection of ρ.

In what follows, we first assume that $\rho \in \bar{\mathscr{H}}$. By Corollary 4.1 we have

$$\left|F\left(\psi,\rho\right) - \mathop{\mathbb{E}}_{\alpha \leftarrow Q_\rho}[f_\psi(\alpha,\eta)]\right| \leq \eta K_\psi. \tag{4.70}$$

Combining Eqs. (4.69) and (4.70) yields

$$\left|F\left(\psi,\rho\right) - \frac{1}{n}\sum_{i=1}^{n} f_\psi(\alpha_i,\eta)\right| \leq \eta K_\psi + \epsilon', \tag{4.71}$$

with probability greater than $1 - 2\exp\left[-\frac{n\eta^{2+2E}\epsilon'^2}{2M_\psi^2(\eta)}\right]$. Setting $\eta = \frac{\epsilon}{K_\psi}$ yields

$$\left|F\left(\psi,\rho\right) - \frac{1}{n}\sum_{i=1}^{n} f_\psi\left(\alpha_i, \frac{\epsilon}{K_\psi}\right)\right| \leq \epsilon + \epsilon', \tag{4.72}$$

with probability greater than $1 - 2\exp\left[-\frac{n\epsilon^{2+2E}\epsilon'^2}{2C_{\psi,1}^2(\epsilon)}\right]$, where we defined

$$\begin{aligned} C_{\psi,1}(\epsilon) &:= K_\psi^{1+E} M_\psi\left(\frac{\epsilon}{K_\psi}\right) \\ &= \sum_{k,l=0}^{E} |\psi_k \psi_l| \epsilon^{E-\frac{k+l}{2}} K_\psi^{1+\frac{k+l}{2}} \sqrt{2^{|l-k|}\binom{\max(k,l)}{\min(k,l)}}. \end{aligned} \tag{4.73}$$

Combining Lemma 4.5 and Eq. (4.72) we obtain

$$\left| F\left(\psi,\rho\right)^m - \left[\frac{1}{n}\sum_{i=1}^{n} f_\psi\left(\alpha_i, \frac{\epsilon}{K_\psi}\right)\right]^m \right| \leq m(\epsilon+\epsilon'), \tag{4.74}$$

with probability greater than $1-2\exp\left[-\frac{n\epsilon^{2+2E}\epsilon'^2}{2C_{\psi,1}^2(\epsilon)}\right]$. Note that we excluded the pathological case $\frac{1}{n}\sum_{i=1}^{n} f_\psi(\alpha_i,\epsilon/K_\psi) > 1$: when that is the case we instead set $\frac{1}{n}\sum_{i=1}^{n} f_\psi(\alpha_i,\epsilon/K_\psi) = 1$. The target state ψ is pure so $F(\psi^{\otimes m},\rho^{\otimes m}) = F\left(\psi,\rho\right)^m$. Hence, replacing ϵ and ϵ' by ϵ/m and ϵ'/m, respectively, gives

$$\left|F\left(\psi,\rho\right)^m - F_\psi(\rho)\right| \leq \epsilon+\epsilon', \tag{4.75}$$

with probability greater than

$$P^{iid}_{\text{Hoeffding}} := 1-2\exp\left[-\frac{n\epsilon^{2+2E}\epsilon'^2}{2m^{4+2E}C_\psi^2}\right], \tag{4.76}$$

where

$$F_\psi(\rho) = \left[\frac{1}{n}\sum_{i=1}^{n} f_\psi\left(\alpha_i, \frac{\epsilon}{mK_\psi}\right)\right]^m, \tag{4.77}$$

and where

$$\begin{aligned} C_\psi &:= C_{\psi,1}(\epsilon/m) \\ &= \sum_{k,l=0}^{E} |\psi_k\psi_l| \left(\frac{\epsilon}{m}\right)^{E-\frac{k+l}{2}} K_\psi^{1+\frac{k+l}{2}} \sqrt{2^{|l-k|}\binom{\max(k,l)}{\min(k,l)}}. \end{aligned} \tag{4.78}$$

Until now we have assumed $\rho \in \bar{\mathscr{H}}$. By Eq. (4.58), the probability that at most s among n values of $|\alpha_i|^2$ are bigger than E, and that the projection Π_E of the state ρ onto the Hilbert space $\bar{\mathscr{H}}$ of states with less than E photons fails is bounded by

$$P^{iid}_{\text{support}} = \frac{(s+1)^{3/2}}{n}\exp\left[\frac{(s+1)^2}{n+1}\right]. \tag{4.79}$$

With the union bound we thus obtain

$$\left|F\left(\psi,\rho\right)^m - F_\psi(\rho)\right| \leq \epsilon+\epsilon', \tag{4.80}$$

or $r > s$, with probability greater than $1-\left(P^{iid}_{\text{support}} + P^{iid}_{\text{Hoeffding}}\right)$.

■

This result implies that the quantity $F_\psi(\rho)$ is a good estimate of the fidelity $F(\psi^{\otimes m},\rho^{\otimes m})$, or the score at the support estimation step is higher than s, with high probability. The values of the energy parameters E and s should be chosen to guarantee completeness, i.e., that if the correct state $|\psi\rangle$ is sent, then $r \leq s$ with high probability. This theorem is valid for all continuous variable target pure states $|\psi\rangle$, and the failure probability may be greatly reduced depending on

the expression of $|\psi\rangle$. The number of samples needed for certifying a given number of copies m with a precision ϵ and a failure probability δ scales as $n = \text{poly}(m, 1/\epsilon, 1/\delta)$.

This certification protocol is promoted to a verification protocol for single-mode states in the following section, by removing the i.i.d. assumption.

4.5 Continuous variable quantum state verification protocol

We now consider an adversarial setting, where a verifier delegates the preparation of a continuous variable quantum state to a potentially malicious party, called the *prover*. One could see the verifier as the experimentalist in the laboratory and the prover as the noisy device, where we aim not to make any assumptions about its correct functionality or noise model. Given the absence of any direct error correction mechanism that permits a fault tolerant run of the device, the aim of verification is to ensure that a wrong outcome is not being accepted. In the context of state verification, this amounts to making sure that the output state of the tested device is close to an ideal target state.

The prover is not supposed to have i.i.d. behaviour. In particular, when asked for various copies of the same state, the prover may actually send a large state entangled over all subsystems, possibly also entangled with a quantum system on his side. In that case, the certification protocol derived in the previous section is not reliable. With usual tomography measurements, the number of samples needed for a given precision of the fidelity estimate scales exponentially in the number of copies to verify. This is an essential limitation of quantum tomography techniques, because they check all possible correlations between the different subsystems.

However we prove that, because of the symmetry of the protocol, the verifier can assume that the prover is sending permutation-invariant states, i.e., states that are invariant under any permutation of their subsystems. After a specific support estimation step, reduced states of permutation-invariant states are close to mixture almost-i.i.d. states, i.e., states that are i.i.d. on almost all subsystems. At the heart of this reduction is a de Finetti theorem for infinite-dimensional systems [RC09], which allows us to restrict to an almost-i.i.d. prover.

4.5.1 Description of the protocol

The verification protocol is as follows: the verifier wants to verify m copies of a target pure state $|\psi\rangle$. The numbers n, k, q, s and E are free parameters of the protocol.

- The prover is instructed to prepare $n+k$ copies of $|\psi\rangle$ and send them to the verifier. We denote by ρ^{n+k} the state received by the verifier.
- The verifier picks k subsystems of the state ρ^{n+k} at random and measures them with heterodyne detection, obtaining the remaining state ρ^n and the samples $\beta_1, \ldots, \beta_k$. The verifier records the number r of values $|\beta_i|^2 > E$ (support estimation step).

- The verifier discards $4q$ subsystems at random, obtaining the remaining state state ρ^{n-4q}, and measures all the others subsystems but m chosen at random with heterodyne detection, obtaining the remaining state ρ^m and the samples $\alpha_1,\ldots,\alpha_{n-4q-m}$.

- The verifier computes with these samples the estimate

$$F_\psi(\rho)=\left[\frac{1}{n-4q-m}\sum_{i=1}^{n-4q-m} f_\psi\left(\alpha_i,\frac{\epsilon}{mK_\psi}\right)\right]^m, \tag{4.81}$$

  where the function f_A and the constant K_A are defined in Eqs. (4.2) and (4.3), for $A=|\psi\rangle\langle\psi|$ and where $\epsilon>0$ is a free parameter.

Note that this estimate is identical to the one defined in Eq. (4.59) for the certification protocol, replacing n by $n-4q-m$. In order to show that this is a good estimate of the fidelity between the remaining state ρ^m and m copies of the target state $|\psi\rangle$, we generalise results from [Ren08, RC08, RC09]. More precisely, we obtain the following results:

- *Support estimation for permutation-invariant states*: with high probability, most of the subsystems of the permutation-invariant state ρ^{n-4q} lie in a lower dimensional subspace, or the score of the state ρ^{n+k} at the support estimation step is high (section 4.5.2).

- *De Finetti reduction*: any permutation-invariant state with most of its subsystems in a lower dimensional subspace has a purification in the symmetric subspace that still has most of its subsystems in a lower dimensional subspace. This purification is well approximated by a mixture of almost-i.i.d. states (section 4.5.3).

- *Hoeffding inequality for almost-i.i.d. states*: mixtures of almost-i.i.d. states can be certified in a similar fashion as i.i.d. states (section 4.5.4).

Using these intermediate results, we obtain the following theorem:

Theorem 4.4 (Gaussian verification of continuous variable quantum states)**.** *Let $n\geq 1$, let $s\leq k$, and let ρ^{n+k} be a state over $n+k$ subsystems. Let $\beta_1,\ldots,\beta_k$ be samples obtained by measuring k subsystems at random with heterodyne detection and let ρ^n be the remaining state after the measurement. Let E in $\mathbb{N}$, and let r be the number of samples such that $|\beta_i|^2>E$. Let also $q\geq m$, and let ρ^m be the state remaining after discarding $4q$ subsystems of ρ^n at random, and measuring $n-4q-m$ other subsystems at random with heterodyne detection, yielding the samples $\alpha_1,\ldots,\alpha_{n-4q-m}$. Let $\epsilon,\epsilon'>0$ and let $|\psi\rangle$ be a target pure state. Then,*

$$\left|F\left(\psi^{\otimes m},\rho^m\right)-F_\psi(\rho)\right|\leq\epsilon+\epsilon'+P_{deFinetti}, \tag{4.82}$$

or $r>s$, with probability greater than

$$1-\left(P_{support}+P_{deFinetti}+P_{choice}+P_{Hoeffding}\right), \tag{4.83}$$

where

$$P_{support} = 8k^{3/2} \exp\left[-\frac{k}{9}\left(\frac{q}{n} - \frac{2s}{k}\right)^2\right], \tag{4.84}$$

$$P_{deFinetti} = q^{(E+1)^2/2} \exp\left[-\frac{2q(q+1)}{n}\right], \tag{4.85}$$

$$P_{choice} = \frac{m(4q+m-1)}{n-4q}, \tag{4.86}$$

$$P_{Hoeffding} = 2\binom{n-4q}{4q} \exp\left[-\frac{n-8q}{2m^{4+2E}}\left(\frac{\epsilon^{1+E}\epsilon'}{C_\psi} - \frac{8qm^{2+E}}{n-4q-m}\right)^2\right], \tag{4.87}$$

where the estimate $F_\psi(\rho)$ is defined in Eq. (4.81), and where

$$C_\psi = \sum_{k,l=0}^{E} |\psi_k \psi_l| \left(\frac{\epsilon}{m}\right)^{E-\frac{k+l}{2}} K_\psi^{1+\frac{k+l}{2}} \sqrt{2^{|l-k|}\binom{\max(k,l)}{\min(k,l)}} \tag{4.88}$$

is a constant independent of ρ, with the constant K defined in Eq. (4.3).

We defer the proof of this result to section 4.5.5. It implies that either the quantity $F_\psi(\rho)$ is a good estimate of the fidelity $F(\psi^{\otimes m}, \rho^m)$, or the score at the support estimation step is higher than s, with high probability. Like for the certification protocol, the values of the energy parameters E and s should be chosen by the verifier to guarantee completeness, i.e., that if the prover sends the correct state $|\psi\rangle$, then $r \leq s$ with high probability.

For specific choices of the free parameters of the protocol, detailed in the proof of the theorem, either the estimate $F_\psi(\rho)$ is polynomially precise in m, or $r > s$, with polynomial probability in m, with $n, k, q = \text{poly}\, m$. In particular, the efficiency of the protocol may be greatly refined by taking into account the expression of $|\psi\rangle$ in the Fock basis, and optimizing over the free parameters.

This verification protocol let the verifier gain confidence about the precision of the estimate of the fidelity in Eq. (4.81). If the value of the estimate is close enough to 1, the verifier may decide to use the state to run a computation. Indeed, statements on the fidelity of a state allow one to infer the correctness of any trusted computation done afterwards using this state. Let $\beta > 0$, and let $\mathcal{O}$ be the observable corresponding to the result of the trusted computation performed on ρ^m, the reduced state over m subsystems instead of $|\psi\rangle^{\otimes m}$, m copies of the target state $|\psi\rangle$. In other words, $\mathcal{O}$ encodes the resources which the verifier can perform perfectly (ancillary states, evolution and measurements), the imperfections being encoded in ρ. Then, $F\left(\psi^{\otimes m}, \rho^m\right) \geq 1 - \beta$ implies the following bound on the total variation distance between the probability densities of the computation output of the actual and the target computations:

$$\|P^{\mathcal{O}}_{\psi^{\otimes m}} - P^{\mathcal{O}}_{\rho^m}\|_{tvd} \leq D(\psi^{\otimes m}, \rho^m) \leq \sqrt{\beta}, \tag{4.89}$$

by standard properties of the trace distance D (see section 1.1.2 and [FVDG99]). What this means is that the distribution of outcomes for the state ρ^m sent by the prover is almost indistinguishable from the distribution of outcomes for m copies of the ideal target state $|\psi\rangle$, when the fidelity is close enough to one.

In what follows, we detail the intermediate steps described above and prove Theorem 4.4.

4.5.2 Support estimation for permutation-invariant states

We first derive a support estimation step for permutation-invariant states. We will use in this section the following operators, already introduced in section 4.4: for $E \geq 0$:

$$U = \sum_{n=E+1}^{+\infty} |n\rangle\langle n| = 1 - \Pi_E, \tag{4.90}$$

where $\Pi_E = \sum_{n=0}^{E} |n\rangle\langle n|$ is the projector onto the Hilbert space $\bar{\mathscr{H}}$ of states with at most E photons, and

$$T = \frac{1}{\pi} \int_{|\alpha|^2 \geq E} |\alpha\rangle\langle\alpha| \, d^2\alpha, \tag{4.91}$$

where $|\alpha\rangle$ is a coherent state. We also recall the following result, from Eq. (4.54), proven in [LGPRC13]:

$$U \leq 2T. \tag{4.92}$$

We recall a few notations and results from [RC08]: let $\mathscr{A} = \{A_0, A_1\}, \mathscr{B} = \{B_0, B_1\}$ be two binary POVMs over $\mathscr{H}$. Define for $\delta > 0$,

$$\gamma_{A\to B}(\delta) = \sup_{\psi} \left\{ \mathrm{Tr}(B\psi), \text{s.t.} \, \mathrm{Tr}(A\psi) \leq \delta \right\}. \tag{4.93}$$

In particular,

$$\gamma_{T\to U}(\delta) \leq 2\delta, \tag{4.94}$$

by Eq. (4.92). We recall the following result (Lemma III.1. of [RC08]):

Lemma 4.6. *Let $n \geq 2k$, let $\delta > 0$, let $\mathscr{A} = \{A_0, A_1\}$ and $\mathscr{B} = \{B_0, B_1\}$ be two binary POVMs over $\mathscr{H}$, and let $x_1, \ldots, x_{n+k}$ the $(n+k)$-partite classical outcome of the measurement $\mathscr{A}^{\otimes n} \otimes \mathscr{B}^{\otimes k}$ applied to any permutation-invariant state ρ^{n+k}. Then*

$$\Pr\left[\frac{x_1 + \cdots + x_n}{n} > \gamma_{B_1 \to A_1}\left(\frac{x_{n+1} + \cdots + x_{n+k}}{k} + \delta \right) + \delta \right] \leq 8k^{3/2} e^{-k\delta^2}. \tag{4.95}$$

This result is a refined version of Serfling's bound [Ser74]. It relates the outcomes of a measurement on some subsystems of a symmetric state with the outcomes of a related measurement on the rest of the subsystems. With this technical Lemma, we derive in what follows a support estimation step for permutation-invariant states using samples from heterodyne detection.

Let ρ^{n+k} be a state over $n+k$ subsystems. Applying a random permutation to this state and measuring its last k subsystems with heterodyne detection is equivalent to measuring k subsystems at random. We thus assume in the following that the state ρ^{n+k} is a permutation-invariant state, without loss of generality, and that the verifier measures its last k subsystems with heterodyne detection.

Let $\mathscr{T} = \{1-T, T\}$ and $\mathscr{U} = \{1-U, U\}$. Let $x_1, \ldots, x_{n+k}$ the $(n+k)$-partite classical outcome of the measurement $\mathscr{U}^{\otimes n} \otimes \mathscr{T}^{\otimes k}$ applied to the permutation-invariant state ρ^{n+k} sent by the prover. A value $x_i = 1$ for $i \in 1, \ldots, n$ means that the projection of the i^{th} subsystem onto $\bar{\mathscr{H}}$ failed, while a value $x_j = 1$ for $j \in n+1, \ldots, n+k$ means that the value $|\beta|^2$ obtained when measuring the j^{th} subsystem with heterodyne detection was bigger than E. In particular, the number of values β_i satisfying $|\beta_i|^2 > E$, is expressed as $x_{n+1} + \cdots + x_{n+k}$. Let $\mathscr{T}_{\leq s}^k$ be the event that at most s of the k values β_i satisfy $|\beta_i|^2 > E$, and let $\mathscr{F}_q^n$ be the event that the projection onto $\bar{\mathscr{H}}$ fails for more than q subsystems of the remaining state ρ^n. Then:

Lemma 4.7 (Support estimation for permutation-invariant states)**.**

$$\Pr\left[\mathscr{F}_q^n \cap \mathscr{T}_{\leq s}^k\right] \leq P_{support}. \tag{4.96}$$

where $P_{support} = 8k^{3/2} \exp\left[-\frac{k}{9}\left(\frac{q}{n} - \frac{2s}{k}\right)^2\right]$.

Proof. With Eq. (4.94), we have for all $\delta > 0$

$$\gamma_{T\to U}\left(\frac{x_{n+1} + \cdots + x_{n+k}}{k} + \delta\right) + \delta \leq 2\frac{x_{n+1} + \cdots + x_{n+k}}{k} + 3\delta. \tag{4.97}$$

Taking $\delta_0 = \frac{1}{3}\left(\frac{q}{n} - \frac{2s}{k}\right)$ we obtain

$$\gamma_{T\to U}\left(\frac{x_{n+1} + \cdots + x_{n+k}}{k} + \delta_0\right) + \delta_0 \leq \frac{q}{n} + 2\left(\frac{x_{n+1} + \cdots + x_{n+k}}{k} - \frac{s}{k}\right), \tag{4.98}$$

so if $x_1 + \cdots + x_n > q$ and $x_{n+1} + \cdots + x_{n+k} \leq s$, then

$$\gamma_{T\to U}\left(\frac{x_{n+1} + \cdots + x_{n+k}}{k} + \delta_0\right) + \delta_0 < \frac{x_1 + \cdots + x_n}{n}. \tag{4.99}$$

Hence,

$$\begin{aligned}
\Pr\left[\mathscr{F}_q^n \cap \mathscr{T}_{\leq s}^k\right] &= \Pr[(x_1 + \cdots + x_n > q) \cap (x_{n+1} + \cdots + x_{n+k} \leq s)] \\
&\leq \Pr\left[\left(\frac{x_1 + \cdots + x_n}{n} > \gamma_{T\to U}\left(\frac{x_{n+1} + \cdots + x_{n+k}}{k} + \delta_0\right) + \delta_0\right)\right] \\
&\leq 8k^{3/2} e^{-k\delta_0^2} \\
&= 8k^{3/2} \exp\left[-\frac{k}{9}\left(\frac{q}{n} - \frac{2s}{k}\right)^2\right],
\end{aligned} \tag{4.100}$$

where we used Lemma 4.6 for $\mathscr{A} = \mathscr{U}$ and $\mathscr{B} = \mathscr{T}$.

■

Recall that $\bar{\mathscr{H}}$ is the Hilbert space of states with at most E photons, of dimension $E+1$. For $q \leq n$, let us define the set of permutation-invariant states over n subsystems, with at most q subsystems out of this lower dimensional subspace (introduced in [RC09]):

$$\mathscr{S}_{\bar{\mathscr{H}}^{\otimes n-q}}^n := \text{span} \bigcup_\pi \pi\left(\bar{\mathscr{H}}^{\otimes n-q} \otimes \mathscr{H}^{\otimes q}\right)\pi^{-1}, \tag{4.101}$$

where the union is taken over all permutations. Lemma 4.7 then gives

$$\Pr\left[\mathscr{F}_q^n \cap \mathscr{T}_{\leq s}^k\right] \leq P_{\text{support}}, \tag{4.102}$$

where $\mathscr{F}_q^n$ is the event that the projection of ρ^n (the remaining state after the support estimation step) onto $\mathscr{S}^n_{\bar{\mathscr{H}}^{\otimes n-q}}$ fails, and where $P_{\text{support}} = 8k^{3/2}\exp\left[-\frac{k}{9}\left(\frac{q}{n}-\frac{2s}{k}\right)^2\right]$. For $1 \ll q \ll n$ and $q/n \ll s/k$, this implies that either ρ^n has most of its subsystems in a lower dimensional subspace, or the score at the support estimation step is higher than s, with high probability.

4.5.3 De Finetti reduction

We recall in this section two results from [RC09].

- The first result says that any permutation-invariant state with most of its subsystems in a lower dimensional subspace has a purification in the symmetric subspace that still has most of its subsystems in a lower dimensional subspace. Formally, for $n \in \mathbb{N}$, and given a Hilbert space $\mathscr{K}$, let us write $\text{Sym}^n(\mathscr{K}) = \{\phi \in \mathscr{K}^{\otimes n},\ \pi\phi = \phi\ (\forall \pi)\}$ the symmetric subspace of a Hilbert space $\mathscr{K}^{\otimes n}$, then (Lemma 3 of [RC09]):

  Lemma 4.8. *For all $q \leq n$, any permutation-invariant state $\rho^n \in \mathscr{S}^n_{\bar{\mathscr{H}}^{\otimes n-q}}$ has a purification $\tilde{\rho}^n$ in $\text{Sym}^n(\mathscr{H}\otimes\mathscr{H}) \bigcap \mathscr{S}^n_{(\bar{\mathscr{H}}\otimes\bar{\mathscr{H}})^{\otimes n-2q}}$.*

The states of the form $|v\rangle^{\otimes n}$ are the so-called *i.i.d. states*. For all $n, r \geq 0$ and all $|v\rangle \in \bar{\mathscr{H}}\otimes\bar{\mathscr{H}}$, the set of *almost-i.i.d. states along* $|v\rangle$, $\mathscr{S}^n_{v^{\otimes n-r}}$, is defined as the span of all vectors that are, up to reorderings, of the form $|v\rangle^{\otimes n-r}\otimes|\phi\rangle$, for an arbitrary $\phi \in (\mathscr{H}\otimes\mathscr{H})^{\otimes r}$. In the following, we simply refer to these states as *almost-i.i.d. states* (which becomes relevant when $r \ll n$).

- The second result is a de Finetti theorem for states in $\text{Sym}^n(\mathscr{H}\otimes\mathscr{H}) \bigcap \mathscr{S}^n_{(\bar{\mathscr{H}}\otimes\bar{\mathscr{H}})^{\otimes n-2q}}$, which says that reduced states from them are well approximated by mixtures of almost-i.i.d. states. Formally (Theorem 4 of [RC09], applied to $\mathscr{K} = \mathscr{H}\otimes\mathscr{H}$ and $\bar{\mathscr{K}} = \bar{\mathscr{H}}\otimes\bar{\mathscr{H}}$, with $\dim(\bar{\mathscr{K}}) = (E+1)^2$):

  Theorem 4.5. *Let $\tilde{\rho}^n \in \text{Sym}^n(\mathscr{H}\otimes\mathscr{H}) \bigcap \mathscr{S}^n_{(\bar{\mathscr{H}}\otimes\bar{\mathscr{H}})^{\otimes n-2q}}$ and let $\tilde{\rho}^{n-4q} = \text{Tr}_{4q}(\tilde{\rho}^n)$. Then, there exist a finite set $\mathscr{V}$ of unit vectors $|v\rangle \in \bar{\mathscr{H}}\otimes\bar{\mathscr{H}}$, a probability distribution $\{p_v\}_{v\in\mathscr{V}}$ over $\mathscr{V}$, and almost-i.i.d. states $\tilde{\rho}_v^{n-4q} \in \mathscr{S}^{n-4q}_{v^{\otimes n-8q}}$ such that*

$$F\left(\tilde{\rho}^{n-4q}, \sum_{v\in\mathscr{V}} p_v \tilde{\rho}_v^{n-4q}\right) > 1 - q^{(E+1)^2}\exp\left[-\frac{4q(q+1)}{n}\right]. \tag{4.103}$$

Given a state $\rho^n \in \mathscr{S}^n_{\bar{\mathscr{H}}^{\otimes n-q}}$, applying Theorem 4.5 to the purification $\tilde{\rho}^n$ given by Lemma 4.8 shows that the reduced state $\tilde{\rho}^{n-4q}$ is close in fidelity to a mixture of states that are i.i.d. on $n-8q$ subsystems.

4.5.4 Hoeffding inequality for almost-i.i.d. states

We recall here Lemma 4.2, in the context of a product measurement applied to an i.i.d. state $|v\rangle\langle v|^{\otimes n}$:

Lemma 4.9. ***(Hoeffding inequality for i.i.d. states)*** *Let $M > 0 \in \mathbb{R}$ and let $f : \mathbb{C} \mapsto \mathbb{R}$ be a function bounded as $|f(\alpha)| < M$ for all $\alpha \in \mathbb{C}$. Let $\lambda > 0$, let $p \in \mathbb{N}^*$, and let $|v\rangle \in \mathscr{H}$. Let $\mathscr{M} = \{\mathscr{M}_\alpha\}_{\alpha\in\mathbb{C}}$ be a POVM on $\mathscr{H}$ and let $D_{|v\rangle}$ be the probability density function of the outcomes of the measurement $\mathscr{M}$ applied to $|v\rangle\langle v|$. Then*

$$\Pr_{\boldsymbol{\alpha}}\left[\left|\frac{1}{p}\sum_{i=1}^{p} f(\alpha_i) - \mathop{\mathbb{E}}_{\beta\leftarrow D_{|v\rangle}}[f(\beta)]\right| \geq \lambda\right] \leq 2\exp\left[-\frac{p\lambda^2}{2M^2}\right], \tag{4.104}$$

where the probability is taken over the outcomes $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)$ of the product measurement $\mathscr{M}^{\otimes p}$ applied to $|v\rangle\langle v|^{\otimes p}$.

The next result gives an equivalent statement for almost-i.i.d. states along a state $|v\rangle$, measured with a product measurement. It generalises Theorem 4.5.2 of [Ren08], where the probability distributions over finite sets, corresponding to product measurements with finite number of outcomes, are replaced by continuous variable probability densities, corresponding to product measurements with continuous variable outcomes. Frequencies estimators are also replaced with estimators of expected values of bounded functions. We will use this result for the POVM corresponding to a product heterodyne detection.

Lemma 4.10 (Hoeffding inequality for almost-i.i.d. states)**.** *Let $M > 0 \in \mathbb{R}$ and let $f : \mathbb{C} \mapsto \mathbb{R}$ be a function bounded as $|f(\alpha)| \leq M$ for all $\alpha \in \mathbb{C}$. Let $\mu > 0$ and $1 \leq m \leq r < t$ such that*

$$(t-m)\mu > 2Mr. \tag{4.105}$$

Let also $|v\rangle \in \bar{\mathscr{H}}$ and $|\Phi\rangle \in \mathscr{S}^t_{v^{\otimes t-r}}$. Let $\mathscr{M} = \{\mathscr{M}_\alpha\}_{\alpha\in\mathbb{C}}$ be a POVM on $\mathscr{H}$ and let $D_{|v\rangle}$ be the probability density function of the outcomes of the measurement $\mathscr{M}$ applied to $|v\rangle\langle v|$. Then

$$\Pr_{\boldsymbol{\alpha}}\left[\left|\frac{1}{t-m}\sum_{i=1}^{t-m} f(\alpha_i) - \mathop{\mathbb{E}}_{\beta\leftarrow D_{|v\rangle}}[f(\beta)]\right| \geq \mu\right] \leq 2\binom{t}{r}\exp\left[-\frac{t-r}{2}\left(\frac{\mu}{M} - \frac{2r}{t-m}\right)^2\right], \tag{4.106}$$

where the probability is taken over the outcomes $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{t-m})$ of the product measurement $\mathscr{M}^{\otimes t-m}$ applied to $|\Phi\rangle\langle\Phi|$.

In essence, this lemma says that a product measurement on all but m subsystems of an almost-i.i.d. state along a state $|v\rangle$ will yield statistics that are similar to the ones that would be obtained by measuring the i.i.d. state $|v\rangle^{\otimes t-m}$.

Proof. $|\Phi\rangle \in \mathscr{S}^t_{v^{\otimes t-r}}$, so by Lemma 4.1.6 of [Ren08], there exist a finite set $\mathscr{S}$ of size at most $\binom{t}{r}$, a family of states $|\tilde{\Phi}^s\rangle \in \mathscr{H}^{\otimes r}$ for $s \in \mathscr{S}$, complex amplitudes $\{\gamma_s\}_{s\in\mathscr{S}}$ and permutations $\{\pi_s\}_{s\in\mathscr{S}}$ over $[1,\dots,t]$ such that

$$\begin{aligned} |\Phi\rangle &:= \sum_{s\in\mathscr{S}} \gamma_s |\Phi^s\rangle \\ &= \sum_{s\in\mathscr{S}} \gamma_s \pi_s \left(|v\rangle^{\otimes t-r} \otimes |\tilde{\Phi}^s\rangle\right). \end{aligned} \tag{4.107}$$

With the notations of the Lemma, let us define for $\mu > 0$:

$$\Omega_\mu = \left\{ \boldsymbol{\alpha} \in \mathbb{C}^{t-m}, \left| \frac{1}{t-m} \sum_{i=1}^{t-m} f(\alpha_i) - \mathop{\mathbb{E}}_{\beta\leftarrow D_{|v\rangle}} [f(\beta)] \right| > \mu \right\}. \tag{4.108}$$

We recall here Lemma of 4.5.1 of [Ren08]:

Lemma 4.11. *Let $|\mathscr{X}|$ be a finite set and $|\psi\rangle = \sum_{x\in\mathscr{X}} |\psi^x\rangle$, and let A be a non-negative operator. Then*

$$\langle\psi|A|\psi\rangle \leq |\mathscr{X}| \sum_{x\in\mathscr{X}} \langle\psi^x|A|\psi^x\rangle. \tag{4.109}$$

In particular, using Eq. (4.107) and this lemma when A is a POVM element of the product measurement $\mathscr{M}_{\boldsymbol{\alpha}} \equiv \mathscr{M}_{\alpha_1} \otimes \cdots \otimes \mathscr{M}_{\alpha_{t-m}}$, we obtain:

$$\begin{aligned} \Pr_{\boldsymbol{\alpha}\leftarrow|\Phi\rangle}[\boldsymbol{\alpha} \in \Omega_\mu] &= \int_{\Omega_\mu} \langle\Phi|\mathscr{M}_{\boldsymbol{\alpha}}|\Phi\rangle \, d^{2(t-m)}\boldsymbol{\alpha} \\ &\leq \int_{\Omega_\mu} |\mathscr{S}| \sum_{s\in\mathscr{S}} |\gamma_s|^2 \langle\Phi^s|\mathscr{M}_{\boldsymbol{\alpha}}|\Phi^s\rangle d^{2(t-m)}\boldsymbol{\alpha} \\ &\leq |\mathscr{S}| \sum_{s\in\mathscr{S}} |\gamma_s|^2 \int_{\Omega_\mu} \langle\Phi^s|\mathscr{M}_{\boldsymbol{\alpha}}|\Phi^s\rangle d^{2(t-m)}\boldsymbol{\alpha} \\ &= |\mathscr{S}| \sum_{s\in\mathscr{S}} |\gamma_s|^2 \Pr_{\boldsymbol{\alpha}\leftarrow|\Phi^s\rangle}[\boldsymbol{\alpha} \in \Omega_\mu], \end{aligned} \tag{4.110}$$

where we write $\boldsymbol{\alpha} \leftarrow |\chi\rangle$ to indicate that $\boldsymbol{\alpha} = (\alpha_1,\dots,\alpha_{t-m})$ is distributed according to the outcomes of the product measurement $\mathscr{M}^{\otimes t-m}$ applied to $|\chi\rangle$.

Let $\boldsymbol{\alpha} \leftarrow |\Phi^s\rangle$. We have $|\Phi^s\rangle = \pi_s(|v\rangle^{\otimes t-r} \otimes |\tilde{\Phi}^s\rangle)$, and in particular $(\alpha_{\pi_s(1)},\dots,\alpha_{\pi_s(t-r)})$ is distributed according to the outcomes of the product measurement $\mathscr{M}^{\otimes t-r}$ applied to $|v\rangle^{\otimes t-r}$. We also have, for $|f| \leq M$,

$$\begin{aligned} \left| \frac{1}{t-r}\sum_{i=1}^{t-r} f(\alpha_{\pi_s(i)}) - \frac{1}{t-m}\sum_{i=1}^{t-m} f(\alpha_i) \right| &= \left| \frac{1}{t-r}\sum_{i=1}^{t-r} f(\alpha_{\pi_s(i)}) - \frac{1}{t-m}\left(\sum_{i=1}^{t} f(\alpha_i) - \sum_{i=t-m+1}^{t} f(\alpha_i)\right)\right| \\ &= \left| \frac{1}{t-r}\sum_{i=1}^{t-r} f(\alpha_{\pi_s(i)}) - \frac{1}{t-m}\left(\sum_{i=1}^{t} f(\alpha_{\pi_s(i)}) - \sum_{i=t-m+1}^{t} f(\alpha_i)\right)\right| \\ &= \left| \left(\frac{1}{t-r} - \frac{1}{t-m}\right)\sum_{i=1}^{t-r} f(\alpha_{\pi_s(i)}) + \frac{1}{t-m}\left(\sum_{i=t-m+1}^{t} f(\alpha_i) - \sum_{i=t-r+1}^{t} f(\alpha_{\pi_s(i)})\right)\right| \end{aligned}$$

$$\begin{aligned}
&\leq \left|\frac{1}{t-r}-\frac{1}{t-m}\right|\sum_{i=1}^{t-r}|f(\alpha_{\pi_s(i)})|+\frac{1}{t-m}\left(\sum_{i=t-m+1}^{t}|f(\alpha_i)|+\sum_{i=t-r+1}^{t}|f(\alpha_{\pi_s(i)})|\right)\\
&\leq \frac{|r-m|}{t-m}M+\frac{(m+r)}{t-m}M\\
&= \frac{2rM}{t-m},
\end{aligned} \tag{4.111}$$

where we used $r \geq m$. Now for all $s \in \mathscr{S}$,

$$\begin{aligned}
\Pr_{\boldsymbol{\alpha}\leftarrow|\Phi^s\rangle}[\boldsymbol{\alpha}\in\Omega_\mu] &= \Pr_{\boldsymbol{\alpha}\leftarrow|\Phi^s\rangle}\left[\left|\frac{1}{t-m}\sum_{i=1}^{t-m}f(\alpha_i)-\mathop{\mathbb{E}}_{\beta\leftarrow D_{|v\rangle}}[f(\beta)]\right|>\mu\right]\\
&\leq \Pr_{\boldsymbol{\alpha}\leftarrow|\Phi^s\rangle}\left[\left|\frac{1}{t-r}\sum_{i=1}^{t-r}f(\alpha_{\pi_s(i)})-\mathop{\mathbb{E}}_{\beta\leftarrow D_{|v\rangle}}[f(\beta)]\right|+\left|\frac{1}{t-m}\sum_{i=1}^{t-m}f(\alpha_i)-\frac{1}{t-r}\sum_{i=1}^{t-r}f(\alpha_{\pi_s(i)})\right|>\mu\right]\\
&\leq \Pr_{\boldsymbol{\alpha}\leftarrow|\Phi^s\rangle}\left[\left|\frac{1}{t-r}\sum_{i=1}^{t-r}f(\alpha_{\pi_s(i)})-\mathop{\mathbb{E}}_{\beta\leftarrow D_{|v\rangle}}[f(\beta)]\right|>\mu-\frac{2rM}{t-m}\right]\\
&\leq 2\exp\left[-\frac{t-r}{2}\left(\frac{\mu}{M}-\frac{2r}{t-m}\right)^2\right],
\end{aligned} \tag{4.112}$$

where we used triangular inequality in the second line, Eq. (4.111) in the third line and Lemma 4.9 in the fourth line with $p = t-r$ and $\lambda = \mu - \frac{2rM}{t-m} > 0$. Combining this last equation with Eq. (4.110), and using $|\mathscr{S}| \leq \binom{t}{r}$ we finally obtain,

$$\Pr_{\boldsymbol{\alpha}\leftarrow|\psi\rangle}\left[\left|\frac{1}{t-m}\sum_{i=1}^{t-m}f(\alpha_i)-\mathop{\mathbb{E}}_{\beta\leftarrow D_{|v\rangle}}[f(\beta)]\right|\geq\mu\right]\leq 2\binom{t}{r}\exp\left[-\frac{t-r}{2}\left(\frac{\mu}{M}-\frac{2r}{t-m}\right)^2\right]. \tag{4.113}$$

■

We recall the bound on $z \mapsto f_\psi(z,\eta)$ obtained in Eq. (4.68): for all $\alpha \in \mathbb{C}$,

$$|f_\psi(\alpha,\eta)| \leq \frac{M_\psi(\eta)}{\eta^{1+E}}, \tag{4.114}$$

where

$$M_\psi(\eta) = \sum_{k,l=0}^{E}\left|\psi_k\psi_l\right|\eta^{E-(k+l)/2}\sqrt{2^{|l-k|}\binom{\max(k,l)}{\min(k,l)}}. \tag{4.115}$$

Let $\mu,\eta>0$, $E\in\mathbb{N}$, let $|v\rangle \in \bar{\mathscr{H}}\otimes\bar{\mathscr{H}}$, and let $|\Phi_v\rangle^{n-4q} \in \mathscr{S}_{v^{\otimes n-8q}}^{n-4q}$. Applying Lemma 4.10 for the real-valued function f_ψ, for $t = n-4q$, for $r = 4q$, for $D_{|v\rangle} = Q_{|v\rangle\langle v|}$, and with the bound from Eq. (4.114), we obtain

$$\begin{aligned}
&\Pr_{\boldsymbol{\alpha}}\left[\left|\frac{1}{n-4q-m}\sum_{i=1}^{n-4q-m}f_\psi(\alpha_i,\eta)-\mathop{\mathbb{E}}_{\beta\leftarrow Q_{|v\rangle\langle v|}}[f_\psi(\beta,\eta)]\right|\geq\mu\right]\\
&\qquad\leq 2\binom{n-4q}{4q}\exp\left[-\frac{n-8q}{2}\left(\frac{\eta^{1+E}\mu}{M_\psi(\eta)}-\frac{8q}{n-4q-m}\right)^2\right],
\end{aligned} \tag{4.116}$$

where the probability is over the outcomes $\boldsymbol{\alpha}$ of a product heterodyne measurement of the first $n-4q-m$ subsystems of $|\Phi_v\rangle^{n-4q} \in \mathscr{S}_{v^{\otimes n-8q}}^{n-4q}$.

4.5.5 Proof of Theorem 4.4

We introduce the following simple result:

Lemma 4.12. *Let $0 < \beta < 1$. Let ρ_1, ρ_2 be two states such that $F(\rho_1, \rho_2) > 1 - \beta$. Let $|\Phi\rangle$ be a pure state, then*

$$\left|F(\Phi, \rho_1) - F(\Phi, \rho_2)\right| \leq D(\rho_1, \rho_2) \leq \sqrt{\beta}. \tag{4.117}$$

Proof. Let us write $P^{\Phi}_{\rho_1}$ and $P^{\Phi}_{\rho_2}$ the probability distributions associated to the binary measurement $\{|\Phi\rangle\langle\Phi|, I - |\Phi\rangle\langle\Phi|\}$ of the states ρ_1 and ρ_2, respectively. Then, $P^{\Phi}_{\rho_1}(0) + P^{\Phi}_{\rho_1}(1) = P^{\Phi}_{\rho_2}(0) + P^{\Phi}_{\rho_2}(1) = 1$, and

$$\begin{aligned} \|P^{\Phi}_{\rho_1} - P^{\Phi}_{\rho_2}\|_{tvd} &= \frac{1}{2}\left(|P^{\Phi}_{\rho_1}(0) - P^{\Phi}_{\rho_2}(0)| + |P^{\Phi}_{\rho_1}(1) - P^{\Phi}_{\rho_2}(1)|\right) \\ &= |P^{\Phi}_{\rho_1}(0) - P^{\Phi}_{\rho_2}(0)|. \end{aligned} \tag{4.118}$$

Hence,

$$\begin{aligned} \left|F(\Phi, \rho_1) - F(\Phi, \rho_2)\right| &= \left|\langle\Phi|\rho_1|\Phi\rangle - \langle\Phi|\rho_2|\Phi\rangle\right| \\ &= |P^{\Phi}_{\rho_1}(0) - P^{\Phi}_{\rho_2}(0)| \\ &= \|P^{\Phi}_{\rho_1} - P^{\Phi}_{\rho_2}\|_{tvd} \\ &\leq D(\rho_1, \rho_2) \\ &\leq \sqrt{1 - F(\rho_1, \rho_2)} \\ &\leq \sqrt{\beta}, \end{aligned} \tag{4.119}$$

where we used Eqs. (1.11, 1.15).

■

With these intermediate results, we are now in position to prove Theorem 4.4.

Proof. Let $|\psi\rangle\langle\psi|$ be the target pure state, and let ρ^{n+k} be a state sent over $n+k$ subsystems. Let $\beta_1, \ldots, \beta_k$ be samples obtained by measuring k subsystems at random of ρ^{n+k} with heterodyne detection. Let ρ^n be the remaining state after the support estimation step. In what follows, we first assume that $\rho^n \in \mathscr{S}^n_{\bar{\mathscr{H}}^{\otimes n-q}}$.

Let ρ^{n-4q} be the state obtained from ρ^n by tracing over the first $4q$ subsystems. In that case, by section 4.5.3, there exist a finite set $\mathscr{V}$ of unit vectors $|v\rangle \in \bar{\mathscr{H}} \otimes \bar{\mathscr{H}}$, a probability distribution $\{p_v\}_{v\in\mathscr{V}}$ over $\mathscr{V}$, and almost-i.i.d. states $\tilde{\rho}^{n-4q}_v \in \mathscr{S}^{n-4q}_{v^{\otimes n-8q}}$ such that

$$F\left(\rho^{n-4q}, \sum_{v\in\mathscr{V}} p_v \rho^{n-4q}_v\right) > 1 - q^{(E+1)^2} \exp\left[-\frac{4q(q+1)}{n}\right], \tag{4.120}$$

where ρ_v^{n-4q} is the remaining state after tracing over the purifying subsystems, since the fidelity is non-decreasing under quantum operations [BCF+96]. We also obtain

$$F\left(\rho^m, \sum_{v\in\mathcal{V}} p_v \rho_v^m\right) > 1 - q^{(E+1)^2} \exp\left[-\frac{4q(q+1)}{n}\right], \tag{4.121}$$

where ρ^m (resp. ρ_v^m) is the remaining state after measuring the first $n-4q-m$ subsystems of ρ^{n-4q} (resp. ρ_v^{n-4q}) with heterodyne detection.

Let $\alpha_1, \ldots, \alpha_{n-4q-m}$ be the samples obtained by measuring the first $n-4q-m$ subsystems of ρ^{n-4q} with heterodyne detection. The verifier computes the estimate (4.81)

$$F_\psi(\rho) = \left[\frac{1}{n-4q-m} \sum_{i=1}^{n-4q-m} f_\psi\left(\alpha_i, \frac{\epsilon}{mK_\psi}\right)\right]^m, \tag{4.122}$$

and whenever $F_\psi \geq 1$ we instead set $F_\psi = 1$. Let us define the completely positive map $\mathscr{E}$ on $\mathscr{H}^{n-4q}$ associated to the classical post-processing of the protocol as:

$$\sigma \mapsto \mathscr{E}(\sigma) = \sum_e \Pr\left[F_\psi(\sigma) = e\right] |e\rangle\langle e|. \tag{4.123}$$

The sum ranges over the values that the estimate may take. With Eq. (4.120) and Lemma 4.12 we obtain

$$D\left(\rho^{n-4q}, \sum_{v\in\mathcal{V}} p_v \rho_v^{n-4q}\right) \leq q^{\frac{(E+1)^2}{2}} \exp\left[-\frac{2q(q+1)}{n}\right], \tag{4.124}$$

The trace distance is non-increasing under quantum operations, so Eq. (4.124) implies

$$D\left(\mathscr{E}\left(\rho^{n-4q}\right), \mathscr{E}\left(\sum_{v\in\mathcal{V}} p_v \rho_v^{n-4q}\right)\right) \leq q^{\frac{(E+1)^2}{2}} \exp\left[-\frac{2q(q+1)}{n}\right]. \tag{4.125}$$

Using the definition of the map $\mathscr{E}$, we obtain a bound in total variation distance:

$$\left\| P\left[F_\psi(\rho)\right] - P\left[F_\psi\left(\sum_{v\in\mathcal{V}} p_v \rho_v^{n-4q}\right)\right] \right\|_{tvd} \leq q^{\frac{(E+1)^2}{2}} \exp\left[-\frac{2q(q+1)}{n}\right], \tag{4.126}$$

where P denotes the probability distributions for the values of the estimates $F_\psi(\rho)$ and $F_\psi\left(\sum_{v\in\mathcal{V}} p_v \rho_v^{n-4q}\right)$.

In particular, this bound implies that for all $\lambda > 0$,

$$\begin{aligned}\left|\Pr\left[\left|F(\psi^{\otimes m}, \rho^m) - F_\psi(\rho)\right| > \lambda\right] - \Pr\left[\left|F(\psi^{\otimes m}, \rho^m) - F_\psi\left(\sum_{v\in\mathcal{V}} p_v \rho_v^{n-4q}\right)\right| > \lambda\right]\right| \\ \leq q^{\frac{(E+1)^2}{2}} \exp\left[-\frac{2q(q+1)}{n}\right],\end{aligned} \tag{4.127}$$

and thus

$$\begin{aligned}\Pr\left[\left|F(\psi^{\otimes m},\rho^m)-F_\psi(\rho)\right|>\lambda\right] \leq q^{\frac{(E+1)^2}{2}}\exp\left[-\frac{2q(q+1)}{n}\right] \\ +\Pr\left[\left|F(\psi^{\otimes m},\rho^m)-F_\psi\left(\sum_{v\in\mathcal{V}}p_v\rho_v^{n-4q}\right)\right|>\lambda\right].\end{aligned} \tag{4.128}$$

With Eq. (4.121) and Lemma 4.12 we obtain

$$\left|F\left(\psi^{\otimes m},\rho^m\right)-F\left(\psi^{\otimes m},\sum_{v\in\mathcal{V}}p_v\rho_v^m\right)\right| \leq q^{\frac{(E+1)^2}{2}}\exp\left[-\frac{2q(q+1)}{n}\right], \tag{4.129}$$

where $\psi^{\otimes m}$ is m copies of the target pure state $|\psi\rangle$. With the triangular inequality,

$$\begin{aligned}&\left|F(\psi^{\otimes m},\rho^m)-F_\psi\left(\sum_{v\in\mathcal{V}}p_v\rho_v^{n-4q}\right)\right| \\ &\leq \left|F(\psi^{\otimes m},\rho^m)-F\left(\psi^{\otimes m},\sum_{v\in\mathcal{V}}p_v\rho_v^m\right)\right|+\left|F\left(\psi^{\otimes m},\sum_{v\in\mathcal{V}}p_v\rho_v^m\right)-F_\psi\left(\sum_{v\in\mathcal{V}}p_v\rho_v^{n-4q}\right)\right| \\ &\leq q^{\frac{(E+1)^2}{2}}\exp\left[-\frac{2q(q+1)}{n}\right]+\left|F\left(\psi^{\otimes m},\sum_{v\in\mathcal{V}}p_v\rho_v^m\right)-F_\psi\left(\sum_{v\in\mathcal{V}}p_v\rho_v^{n-4q}\right)\right|,\end{aligned} \tag{4.130}$$

where we used Eq. (4.129) in the last line. With Eq. (4.128) we obtain, for all $\lambda>0$

$$\begin{aligned}\Pr\left[\left|F(\psi^{\otimes m},\rho^m)-F_\psi(\rho)\right|>\lambda\right] \leq q^{\frac{(E+1)^2}{2}}\exp\left[-\frac{2q(q+1)}{n}\right] \\ +\Pr\left[\left|F\left(\psi^{\otimes m},\sum_{v\in\mathcal{V}}p_v\rho_v^m\right)-F_\psi\left(\sum_{v\in\mathcal{V}}p_v\rho_v^{n-4q}\right)\right|>\lambda-q^{\frac{(E+1)^2}{2}}e^{-\frac{2q(q+1)}{n}}\right].\end{aligned} \tag{4.131}$$

By linearity of the probabilities, it suffices to bound $\Pr\left[\left|F(\psi^{\otimes m},\Phi^m)-F_\psi(\Phi)\right|>\mu\right]$, for $\mu=\lambda-q^{\frac{(E+1)^2}{2}}\exp\left[-\frac{2q(q+1)}{n}\right]$, where $|\Phi\rangle\in\mathscr{S}^{n-4q}_{v^{\otimes n-8q}}$, for $|v\rangle\in\bar{\mathscr{H}}\otimes\bar{\mathscr{H}}$, and where Φ^m is the state obtained from $|\Phi\rangle\langle\Phi|$ by measuring the first $n-4q-m$ subsystems with heterodyne detection and tracing over the purifying subsystems.

Lemma 4.13. *Let $|\Phi\rangle\in\mathscr{S}^{n-4q}_{v^{\otimes n-8q}}$. For all $\epsilon'>0$,*

$$\begin{aligned}\Pr\left[\left|F(\psi^{\otimes m},\Phi^m)-F_\psi(\Phi)\right|>\epsilon+\epsilon'\right] \leq 2\binom{n-4q}{4q}\exp\left[-\frac{n-8q}{2m^{4+2E}}\left(\frac{\epsilon^{1+E}\epsilon'}{C_\psi}-\frac{8qm^{2+E}}{n-4q-m}\right)^2\right] \\ +\frac{m(4q+m-1)}{n-4q},\end{aligned} \tag{4.132}$$

where

$$C_\psi = \sum_{k,l=0}^{E} |\psi_k \psi_l| \left(\frac{\epsilon}{m}\right)^{E-\frac{k+l}{2}} K_\psi^{1+\frac{k+l}{2}} \sqrt{2^{|l-k|}\binom{\max(k,l)}{\min(k,l)}} \underset{\epsilon\to 0}{\longrightarrow} |\psi_E|^2 K_\psi^{1+E}. \tag{4.133}$$

Proof. Let $\alpha_1,\ldots,\alpha_{n-4q-m}$ be samples obtained by measuring the first $n-4q-m$ subsystems of $|\Phi\rangle\langle\Phi|$ with heterodyne detection. We have (4.81)

$$F_\psi(\Phi) = \left[\frac{1}{n-4q-m}\sum_{i=1}^{n-4q-m} f_\psi\left(\alpha_i, \frac{\epsilon}{mK_\psi}\right)\right]^m, \tag{4.134}$$

and

$$\begin{aligned}
\left|F(\psi^{\otimes m},\Phi^m) - F_\psi(\Phi)\right| &\leq \left|F(\psi^{\otimes m},\Phi^m) - F(\psi^{\otimes m},|v\rangle\langle v|^{\otimes m})\right| \\
&\quad + \left|F(\psi^{\otimes m},|v\rangle\langle v|^{\otimes m}) - \left(\underset{\beta\leftarrow Q_{|v\rangle\langle v|}}{\mathbb{E}}\left[f_\psi\left(\beta,\frac{\epsilon}{mK_\psi}\right)\right]\right)^m\right| \\
&\quad + \left|\left(\underset{\beta\leftarrow Q_{|v\rangle\langle v|}}{\mathbb{E}}\left[f_\psi\left(\beta,\frac{\epsilon}{mK_\psi}\right)\right]\right)^m - F_\psi(\Phi)\right| \\
&= \left|F(\psi^{\otimes m},\Phi^m) - F(\psi^{\otimes m},|v\rangle\langle v|^{\otimes m})\right| \\
&\quad + \left|F(\psi,|v\rangle\langle v|)^m - \left(\underset{\beta\leftarrow Q_{|v\rangle\langle v|}}{\mathbb{E}}\left[f_\psi\left(\beta,\frac{\epsilon}{mK_\psi}\right)\right]\right)^m\right| \\
&\quad + \left|\left(\underset{\beta\leftarrow Q_{|v\rangle\langle v|}}{\mathbb{E}}\left[f_\psi\left(\beta,\frac{\epsilon}{mK_\psi}\right)\right]\right)^m - \left(\frac{1}{n-4q-m}\sum_{i=1}^{n-4q-m} f_\psi\left(\alpha_i,\frac{\epsilon}{mK_\psi}\right)\right)^m\right| \\
&\leq \left|F(\psi^{\otimes m},\Phi^m) - F(\psi^{\otimes m},|v\rangle\langle v|^{\otimes m})\right| \\
&\quad + m\left|F(\psi,|v\rangle\langle v|) - \underset{\beta\leftarrow Q_{|v\rangle\langle v|}}{\mathbb{E}}\left[f_\psi\left(\beta,\frac{\epsilon}{mK_\psi}\right)\right]\right| \\
&\quad + m\left|\underset{\beta\leftarrow Q_{|v\rangle\langle v|}}{\mathbb{E}}\left[f_\psi\left(\beta,\frac{\epsilon}{mK_\psi}\right)\right] - \frac{1}{n-4q-m}\sum_{i=1}^{n-4q-m} f_\psi\left(\alpha_i,\frac{\epsilon}{mK_\psi}\right)\right|,
\end{aligned} \tag{4.135}$$

where we used Lemma 4.5. We bound these three terms in the following.

When selecting at random m subsystems from an almost-i.i.d. state over $n-4q$ subsystems which is i.i.d. on $n-8q$ subsystems, the probability that all of the selected states are from the $n-8q$ i.i.d. subsystems is

$$\frac{\binom{n-8q}{m}}{\binom{n-4q}{m}} = \frac{(n-8q)(n-8q-1)\ldots(n-8q-m+1)}{(n-4q)(n-4q-1)\ldots(n-4q-m+1)}, \tag{4.136}$$

and we have

$$\begin{aligned} 1-\frac{(n-8q)(n-8q-1)\ldots(n-8q-m+1)}{(n-4q)(n-4q-1)\ldots(n-4q-m+1)} &\leq 1-\frac{(n-8q-m+1)^m}{(n-4q)^m} \\ &= 1-\left(1-\frac{4q+m-1}{n-4q}\right)^m \\ &\leq \min\left(1,\frac{m(4q+m-1)}{n-4q}\right) \\ &\leq \frac{m(4q+m-1)}{n-4q}, \end{aligned} \tag{4.137}$$

where we used $1-(1-x)^a \leq ax$ for all $a \geq 1$ and $x \in [0,1]$. In particular, for $|\Phi\rangle \in \mathscr{S}_{v^{\otimes n-8q}}^{n-4q}$, and Φ^m its reduced state over m modes chosen at random, we have

$$\Phi^m = |v\rangle\langle v|^{\otimes m}, \tag{4.138}$$

with probability greater than $1-\frac{m(4q+m-1)}{n-4q}$, where we used the definition of $\mathscr{S}_{v^{\otimes n-8q}}^{n-4q}$, and Eq. (4.137). Using Lemma 4.12, the *first term* in Eq. (4.135) vanishes with probability greater than:

$$1-\frac{m(4q+m-1)}{n-4q}. \tag{4.139}$$

The bound for the *second term* is given by Corollary 4.1 applied to the state $|v\rangle$, for $\eta = \frac{\epsilon}{mK_\psi}$:

$$m\left|F(\psi,|v\rangle\langle v|)-\mathop{\mathbb{E}}_{\beta\leftarrow Q_{|v\rangle\langle v|}}\left[f_\psi\left(\beta,\frac{\epsilon}{mK_\psi}\right)\right]\right| \leq \epsilon. \tag{4.140}$$

The bound for the *third term* is probabilistic, given by Eq. (4.116), for $\eta = \frac{\epsilon}{mK_\psi}$ and $\mu = \frac{\epsilon'}{m}$. For all $\epsilon' > 0$,

$$\begin{aligned} \Pr_{\boldsymbol{\alpha}}\left[\left|\frac{1}{n-4q-m}\sum_{i=1}^{n-4q-m} f_\psi\left(\alpha_i,\frac{\epsilon}{mK_\psi}\right)-\mathop{\mathbb{E}}_{\beta\leftarrow Q_{|v\rangle\langle v|}}\left[f_\psi\left(\beta,\frac{\epsilon}{mK_\psi}\right)\right]\right| \geq \frac{\epsilon'}{m}\right] \\ \leq 2\binom{n-4q}{4q}\exp\left[-\frac{n-8q}{2}\left(\frac{\epsilon^{1+E}\epsilon'}{m^{2+E}K_\psi^{1+E}M_\psi(\frac{\epsilon}{mK_\psi})}-\frac{8q}{n-4q-m}\right)^2\right]. \end{aligned} \tag{4.141}$$

We now bring together the previous bounds in order to prove Lemma 4.13. Combining

Eqs. (4.135), (4.139), (4.140) and (4.141) yields

$$\begin{aligned}
&\Pr\left[\left|F(\psi^{\otimes m},\Phi^m)-F_\psi(\Phi)\right|>\epsilon+\epsilon'\right]\\
&\quad\le \Pr_{\boldsymbol{\alpha}}\left[\left|\frac{1}{n-4q-m}\sum_{i=1}^{n-4q-m} f_\psi\left(\alpha_i,\frac{\epsilon}{mK_\psi}\right)-\mathop{\mathbb{E}}_{\beta\leftarrow Q_{|v\rangle\langle v|}}\left[f_\psi\left(\beta,\frac{\epsilon}{mK_\psi}\right)\right]\right|\ge\frac{\epsilon'}{m}\right]\\
&\quad\le 2\binom{n-4q}{4q}\exp\left[-\frac{n-8q}{2}\left(\frac{\epsilon^{1+E}\epsilon'}{m^{2+E}K_\psi^{1+E}M_\psi(\frac{\epsilon}{mK_\psi})}-\frac{8q}{n-4q-m}\right)^2\right]\\
&\qquad+\frac{m(4q+m-1)}{n-4q}\\
&\quad= 2\binom{n-4q}{4q}\exp\left[-\frac{n-8q}{2m^{4+2E}}\left(\frac{\epsilon^{1+E}\epsilon'}{C_\psi}-\frac{8qm^{2+E}}{n-4q-m}\right)^2\right]+\frac{m(4q+m-1)}{n-4q},
\end{aligned}\tag{4.142}$$

where

$$\begin{aligned}
C_\psi&=K_\psi^{1+E}M_\psi\left(\frac{\epsilon}{mK_\psi}\right)\\
&=\sum_{k,l=0}^{E}|\psi_k\psi_l|\left(\frac{\epsilon}{m}\right)^{E-\frac{k+l}{2}}K_\psi^{1+\frac{k+l}{2}}\sqrt{2^{|l-k|}\binom{\max(k,l)}{\min(k,l)}}\underset{\epsilon\to0}{\longrightarrow}|\psi_E|^2K_\psi^{1+E}.
\end{aligned}\tag{4.143}$$

□

Combining Eq. (4.131) and Lemma 4.13, we finally obtain

$$\begin{aligned}
&\Pr\left[\left|F(\psi^{\otimes m},\rho^m)-F_\psi(\rho)\right|>\epsilon+\epsilon'+q^{\frac{(E+1)^2}{2}}e^{-\frac{2q(q+1)}{n}}\right]\\
&\quad\le q^{\frac{(E+1)^2}{2}}\exp\left[-\frac{2q(q+1)}{n}\right]+2\binom{n-4q}{4q}\exp\left[-\frac{n-8q}{2m^{4+2E}}\left(\frac{\epsilon^{1+E}\epsilon'}{C_\psi}-\frac{8qm^{2+E}}{n-4q-m}\right)^2\right]\\
&\qquad+\frac{m(4q+m-1)}{n-4q}.
\end{aligned}\tag{4.144}$$

Setting

$$P_{\text{Hoeffding}}=2\binom{n-4q}{4q}\exp\left[-\frac{n-8q}{2m^{4+2E}}\left(\frac{\epsilon^{1+E}\epsilon'}{C_\psi}-\frac{8qm^{2+E}}{n-4q-m}\right)^2\right],\tag{4.145}$$

$$P_{\text{choice}}=\frac{m(4q+m-1)}{n-4q},\tag{4.146}$$

and

$$P_{\text{deFinetti}}=q^{\frac{(E+1)^2}{2}}\exp\left[-\frac{2q(q+1)}{n}\right],\tag{4.147}$$

we obtain

$$\Pr\left[\left|F(\psi^{\otimes m},\rho^m)-F_\psi(\rho)\right|>\epsilon+\epsilon'+P_{\text{deFinetti}}\right]\le P_{\text{deFinetti}}+P_{\text{choice}}+P_{\text{Hoeffding}}.\tag{4.148}$$

Until now we have assumed $\rho^n \in \mathscr{S}^n_{\bar{\mathcal{H}}^{\otimes n-q}}$. By section 4.5.2,

$$\Pr\left[\mathscr{F}_q^n \cap \mathscr{T}_{\leq s}^k\right] \leq P_{\text{support}}. \tag{4.149}$$

where $\mathscr{F}_q^n$ is the event that the projection of ρ^n (the remaining state after the support estimation step) onto $\mathscr{S}^n_{\bar{\mathcal{H}}^{\otimes n-q}}$ fails, where $\mathscr{T}_{\leq s}^k$ is the event that at most s of the k values β_i from the support estimation step satisfy $|\beta_i|^2 > E$, and where $P_{\text{support}} = 8k^{3/2}\exp\left[-\frac{k}{9}\left(\frac{q}{n}-\frac{2s}{k}\right)^2\right]$. With the union bound we thus obtain

$$\Pr\left[\left(\left|F(\psi^{\otimes m},\rho^m) - F_\psi(\rho)\right| > \epsilon + \epsilon' + P_{\text{deFinetti}}\right) \cap \mathscr{T}_{\leq s}^k\right] \leq P_{\text{support}} + P_{\text{deFinetti}} + P_{\text{choice}} + P_{\text{Hoeffding}}, \tag{4.150}$$

where

$$P_{\text{support}} = 8k^{3/2}\exp\left[-\frac{k}{9}\left(\frac{q}{n}-\frac{2s}{k}\right)^2\right],$$

$$P_{\text{deFinetti}} = q^{\frac{(E+1)^2}{2}}\exp\left[-\frac{2q(q+1)}{n}\right], \tag{4.151}$$

$$P_{\text{choice}} = \frac{m(4q+m-1)}{n-4q}, \tag{4.152}$$

$$P_{\text{Hoeffding}} = 2\binom{n-4q}{4q}\exp\left[-\frac{n-8q}{2m^{4+2E}}\left(\frac{\epsilon^{1+E}\epsilon'}{C_\psi} - \frac{8qm^{2+E}}{n-4q-m}\right)^2\right].$$

The variables $\epsilon, \epsilon', n, m, q, k, s, E$ are free parameters of the protocol. Let us fix, e.g., $E = O(1)$, $s = O(1)$, $n = O\left(m^{19+8E}\right)$, $k = O(m^{19+8E})$, $q = O\left(m^{10+4E}\right)$, and $\epsilon = \epsilon' = O(\frac{1}{m})$. Then, either the estimate $F_\psi(\rho)$ of the fidelity $F(\psi^{\otimes m},\rho^m)$ is polynomially precise (in m), or the score at the support estimation step is higher than s, with polynomial probability (in m), by plugging the different scalings in the above expressions.

■

These general single-mode state certification and verification protocols may be used for various usecases. We present selected applications in the following section, relating to the certification of non-Gaussian properties of quantum states.

4.6 Certification of non-Gaussian properties

4.6.1 Certifying the stellar rank

The stellar hierarchy can be certified with the previous protocol using the estimate of the fidelity obtained as a witness for the stellar rank. We recall a few definitions and results from chapter 2. The stellar rank of a single-mode normalised pure quantum state corresponds to the minimal number of photon additions necessary to engineer the state from the vacuum, together with Gaussian unitary operations. Moreover, a mixed state which has a stellar rank equal to n cannot

be expressed as a mixture of pure states of ranks strictly lower than n. Given $k \in \mathbb{N}^*$ and a target pure state $|\psi\rangle$, if a mixed state ρ satisfies

$$F(\psi,\rho) > 1 - [R_k^\star(\psi)]^2, \tag{4.153}$$

where $R_k^\star(\psi)$ is the k-robustness of the state $|\psi\rangle$, then it has a stellar rank greater or equal to k. This in turn can be checked by computing the robustness profile of the state $|\psi\rangle$.

With Theorem 4.3 for $m = 1$, we obtain the following protocol for certifying the stellar rank under the i.i.d. assumption, where E, s, ϵ and ϵ' are free parameters:

Let $|\psi\rangle$ be a target pure state. First, measure with heterodyne detection n copies of the (mixed) state ρ, obtaining the samples $\alpha_1,\dots,\alpha_n$. Then, record the number r of samples such that $|\alpha_i|^2 > E$. Compute with the same samples the estimate

$$F_\psi(\rho) = \frac{1}{n}\sum_{i=1}^{n} f_\psi\left(\alpha_i, \frac{\epsilon}{K_\psi}\right), \tag{4.154}$$

where the function f_A and the constant K_A are defined in Eqs. (4.2) and (4.3), for $A = |\psi\rangle\langle\psi|$. Then,

$$\left|F(\psi,\rho) - F_\psi(\rho)\right| \leq \epsilon + \epsilon', \tag{4.155}$$

or $r > s$, with probability greater than

$$1 - \left(\frac{(s+1)^{3/2}}{n}\exp\left[\frac{(s+1)^2}{n+1}\right] + 2\exp\left[-\frac{n\epsilon^{2+2E}\epsilon'^2}{2C_\psi^2}\right]\right), \tag{4.156}$$

where

$$C_\psi = \sum_{k,l=0}^{E} |\psi_k\psi_l|\epsilon^{E-\frac{k+l}{2}} K_\psi^{1+\frac{k+l}{2}} \sqrt{2^{|l-k|}\binom{\max(k,l)}{\min(k,l)}} \tag{4.157}$$

is a constant independent of ρ, with the constant K defined in Eq. (4.3). In particular, if the estimate obtained satisfies

$$F_\psi(\rho) > 1 - [R_k^\star(\psi)]^2 + \epsilon + \epsilon', \tag{4.158}$$

which can be readily checked from the robustness profile of the target state $|\psi\rangle$, then either the score at the support estimation step is high or the state ρ has stellar rank greater or equal to k, with high probability for a large number of samples. An analogous statement holds for the case of verification, without the i.i.d. assumption, with Theorem 4.4.

4.6.2 Certifying Wigner negativity

In the previous section, we detail how to certify a nonzero stellar rank of any experimental (mixed) state, which implies that this state is non-Gaussian. However, such a mixed state may still have positive Wigner function. Since processes with positive Wigner functions are classically simulable [ME12], negativity of the Wigner function is also a crucial property to look for. In

this section, we show how our certification protocol with heterodyne detection allows for the certification of Wigner negativity without the need for a full tomography.

The Wigner function of a state ρ evaluated at $\alpha \in \mathbb{C}$ is related to the expected value of the parity operator displaced by α [Roy77]:

$$W_\rho(\alpha) = \frac{2}{\pi} \mathrm{Tr}\left[\hat{D}(\alpha)\hat{\Pi}\hat{D}^\dagger(\alpha)\rho\right], \tag{4.159}$$

where

$$\hat{\Pi} = \sum_{n\geq 0} (-1)^n |n\rangle\langle n| \tag{4.160}$$

is the parity operator. Hence, we can use the certification protocol to obtain mean value estimations of the operator $\frac{2}{\pi}\hat{D}(\alpha)\hat{\Pi}\hat{D}^\dagger(\alpha)$ and retrieve the value of the Wigner function at α. Moreover, since the displacement can be reverted in post-processing by translating the samples by α, we can alternatively obtain mean value estimations of the operator $\frac{2}{\pi}\hat{\Pi}$ (which has a simpler expression in Fock basis) using translated samples.

Using either the certification protocol from Theorem 4.3 or the verification protocol from Theorem 4.4 allows us to witness Wigner negativity under or wihtout the i.i.d. assumption, respectively.

4.7 Certifying multimode continuous variable quantum states

The certification and verification protocols described in the previous sections allow us to obtain efficiently estimates of fidelities of any single-mode continuous variable quantum state with any target single-mode pure state, with analytical confidence intervals, either with i.i.d. assumption or with no assumption whatsoever. These protocols also allow us to estimate efficiently fidelities with multimode i.i.d. pure states. However, translating them directly to efficient protocols for general multimode states seems hopeless, since verifying a multimode state implies accounting for all possible correlations between its subsystems, of which there is an exponential number in the size of the state.

On the other hand, we show in what follows that being able to estimate single-mode fidelities with heterodyne detection is enough to provide fidelity witnesses for a large class of multimode states. This result combines the following two observations:

- If all the single-mode subsystems ρ_i of a multimode quantum state $\boldsymbol{\rho}$ are close enough to single-mode pure states $|\psi_i\rangle\langle\psi_i|$, then $\boldsymbol{\rho}$ is close to the tensor product of these pure states (Lemma 4.14). In particular, being able to estimate single-mode fidelities is enough to provide fidelity witnesses for product of pure states.

- Passive linear transformations followed by single-mode Gaussian unitary operations and product of single-mode balanced heterodyne detections can be simulated by performing unbalanced heterodyne detections first, then post-processing efficiently the samples

(Lemma 4.15). In particular, for such an operation $\hat{V}$, if the multimode state $\boldsymbol{\rho}$ can be efficiently certified using heterodyne detection, then it is also the case for the state $\hat{V}\boldsymbol{\rho}\hat{V}^\dagger$.

This allows us to verify efficiently a large class of multimode continuous variable quantum states, with and without the i.i.d. assumption, including the m-mode states of the form

$$\left(\bigotimes_{i=1}^{m}\hat{G}_i\right)\hat{U}\left(\bigotimes_{i=1}^{m}|\psi_i\rangle\right), \tag{4.161}$$

where $\hat{U}$ is a passive linear transformation (a unitary transformation of the creation and annihilation operators of the modes) and where, for all $i \in \{1,\dots,m\}$, the state $|\psi_i\rangle$ is a single-mode pure state with constant energy (which does not scale with the number of modes m) and the operation $\hat{G}_i$ is a single-mode Gaussian unitary which may be written as a combination of a single-mode displacement and a single-mode squeezing (see section 1.3). In particular, these states includes multimode Gaussian states and the output states of Boson Sampling interferometers and of CVS circuits (see sections 1.4.5 and 3.3.3).

The fidelity witnesses presented here extend the work of [AGKE15] in various respects. Their work provides fidelity witnesses for multimode photonic state preparations with Gaussian measurements, under the i.i.d. assumption. However, the witnesses are for a more restricted class of target states and are efficient for Gaussian pure states only. In particular, the number of copies needed to certify with constant precision the output of a Boson Sampling interferometer with n input photons over m modes with their protocol scales as $\Omega(m^{n+4})$, which is worse than exponential in the antibunching regime $n = O(\sqrt{m})$, while we show that our protocol provides tight fidelity witnesses with constant precision with $O(m^4 \log m)$ copies. Moreover, we are able to remove the i.i.d. state preparation assumption, at the cost of an increased—though still polynomial—number of measurements needed for the same estimate precision and confidence interval.

In the following sections, we present the general protocol and detail its application in the case of Boson Sampling.

4.7.1 General multimode protocol

We present the two versions of the multimode verification protocol, with or without i.i.d. assumption. Under the i.i.d. assumption:

1. The verifier chooses an m-mode target pure state $|\boldsymbol{\tau}_{U,\boldsymbol{\xi},\boldsymbol{\beta}}\rangle := \hat{S}(\boldsymbol{\xi})\hat{D}(\boldsymbol{\beta})\hat{U}(\bigotimes_{i=1}^{m}|\psi_i\rangle)$, as in Eq. (4.161), where for all $i \in \{1,\dots,m\}$ the state $|\psi_i\rangle$ has constant energy, where $\hat{U}$ is an m-mode passive linear transformation with $m \times m$ unitary matrix U and where $\boldsymbol{\xi}, \boldsymbol{\beta} \in \mathbb{C}^m$. The verifier also chooses a precision parameter $0 < \eta < 1$ and energy cutoff values $E_1,\dots,E_m$.

2. The verifier asks the prover for $N = O(\text{poly } m)$ copies of the target state $|\boldsymbol{\tau}_{U,\boldsymbol{\xi},\boldsymbol{\beta}}\rangle$. Let $\boldsymbol{\rho}^{\otimes N}$ be the $(N \times m)$-mode (mixed) state sent by the prover, where $\boldsymbol{\rho}$ is an m-mode (mixed) state.

3. The verifier measures with unbalanced heterodyne detection with unbalancing parameters $\boldsymbol{\xi}$ all the m subsystems of all the N copies of $\boldsymbol{\rho}$, obtaining the N vectors of samples $\boldsymbol{\gamma}^{(1)},\dots,\boldsymbol{\gamma}^{(N)} \in \mathbb{C}^m$.

4. For all $k \in \{1,\dots,N\}$, the verifier computes $\boldsymbol{\alpha}^{(k)} = U^\dagger(\boldsymbol{\gamma}^{(k)} - \boldsymbol{\beta})$. We write $\boldsymbol{\alpha}^{(k)} = (\alpha_1^{(k)},\dots,\alpha_m^{(k)})$.

5. For all $i \in \{1,\dots,m\}$, the verifier records the number r_i of values among $\alpha_i^{(1)},\dots,\alpha_i^{(N)}$ such that $|\alpha_i^{(k)}|^2 > E_i$ (*support estimation*).

6. For all $i \in \{1,\dots,m\}$, the verifier computes the mean $\tilde{F}_i$ of the function $z \mapsto f_{\psi_i}(z,\epsilon,E_i)$ over the same values $\alpha_i^{(1)},\dots,\alpha_i^{(N)}$, where the function f is defined in Eq.(4.2).

7. The verifier computes $\tilde{W} = 1 - \sum_{i=1}^m (1-\tilde{F}_i)$.

The cutoff values $E_1,\dots,E_m$ should be chosen by the verifier to guarantee completeness for the estimation of the single-mode fidelities, i.e., that if the prover is sending a near-ideal state it is accepted with high probability. For a sufficiently large number of copies $N = O(\text{poly}\, m)$, we show in what follows that $\tilde{W}$ is a tight lower bound on the fidelity with inverse polynomial precision, or one of the scores $r_1,\dots,r_m$ is high, with high probability.

Without i.i.d. assumption, an equivalent protocol is obtained by using the version of the protocol which does not assume i.i.d. state preparation for estimating the single-mode fidelities in Theorem 4.4. In that case, the final protocol is nearly identical, up to slight differences for the classical post-processing: a small fraction of the measured subsystems have to be discarded at random and the samples used for the support estimation step must be randomly chosen and cannot be used to compute the fidelity estimates. This comes at the cost of an increased number of measurements necessary for the same witness precision and confidence interval, which corresponds however to a polynomial overhead in m.

For both protocols, note that the efficiency may be greatly refined by taking into account the expression of the single-mode target pure states $|\psi_i\rangle$ in Fock basis. We give an example of such optimisation in the next section, in the case of Boson Sampling output states, when the single-mode target pure states are either single-photon Fock states or vacuum states. In particular, for the protocol under i.i.d. assumption, if the single-mode target states have a finite support over the Fock basis then the support estimation step is no longer needed.

We now show that the estimate $\tilde{W}$ is a tight fidelity witness for a number of samples $O(\text{poly}\, m)$. We first prove the following result:

Lemma 4.14. *Let $\boldsymbol{\rho}$ be an m-mode state. For all $i \in \{1,\dots,m\}$, we denote by ρ_i the single-mode reduced state of $\boldsymbol{\rho}$ over the i^{th} mode. Let $|\psi_1\rangle,\dots,|\psi_m\rangle$ be single-mode pure states. For all $i \in \{1,\dots,m\}$, we write $F(\rho_i,\psi_i) = 1-\epsilon_i$, where F is the fidelity. Then,*

$$1 - \sum_{i=1}^m \epsilon_i \leq F(\boldsymbol{\rho}, \psi_1 \otimes \cdots \otimes \psi_m) \leq \prod_{i=1}^m (1-\epsilon_i). \tag{4.162}$$

In particular, when $\epsilon_1 = \cdots = \epsilon_m = \epsilon$,

$$1 - m\epsilon \leq F(\rho, \psi_1 \otimes \cdots \otimes \psi_m) \leq (1-\epsilon)^m. \tag{4.163}$$

Proof. Since $|\psi_1\rangle, \ldots, |\psi_m\rangle$ are pure states,

$$F(\boldsymbol{\rho}, \psi_1 \otimes \cdots \otimes \psi_m) = \mathrm{Tr}[\boldsymbol{\rho}\, |\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_m\rangle\langle\psi_m|], \tag{4.164}$$

and

$$\begin{aligned} F(\boldsymbol{\rho}_i, \psi_i) &= \mathrm{Tr}[\boldsymbol{\rho}_i\, |\psi_i\rangle\langle\psi_i|] \\ &= \mathrm{Tr}[\boldsymbol{\rho}\, \mathbb{1}_{i-1} \otimes |\psi_i\rangle\langle\psi_i| \otimes \mathbb{1}_{m-i}] \end{aligned} \tag{4.165}$$

for all $i \in \{1, \ldots, m\}$. The left hand side of Eq. (4.162) is obtained by writing $F(\boldsymbol{\rho}, \psi_1 \otimes \cdots \otimes \psi_m)$ as a telescopic sum:

$$\begin{aligned} \mathrm{Tr}[\boldsymbol{\rho}\, |\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_m\rangle\langle\psi_m|] &= \mathrm{Tr}[\boldsymbol{\rho}\, \mathbb{1}_m] \\ &\quad - \mathrm{Tr}[\boldsymbol{\rho}(\mathbb{1} - |\psi_1\rangle\langle\psi_1|) \otimes \mathbb{1}_{m-1}] \\ &\quad - \mathrm{Tr}[\boldsymbol{\rho}\, |\psi_1\rangle\langle\psi_1| \otimes (\mathbb{1} - |\psi_2\rangle\langle\psi_2|) \otimes \mathbb{1}_{m-2}] \\ &\quad - \mathrm{Tr}[\boldsymbol{\rho}\, |\psi_1\rangle\langle\psi_1| \otimes |\psi_2\rangle\langle\psi_2| \otimes (\mathbb{1} - |\psi_3\rangle\langle\psi_3|) \otimes \mathbb{1}_{m-3}] \\ &\quad - \ldots \\ &\quad - \mathrm{Tr}[\boldsymbol{\rho}\, |\psi_1\rangle\langle\psi_1| \otimes |\psi_2\rangle\langle\psi_2| \otimes \cdots \otimes (\mathbb{1} - |\psi_m\rangle\langle\psi_m|)] \\ &\geq 1 - \sum_{i=1}^{m} \big(1 - \mathrm{Tr}[\boldsymbol{\rho}\, \mathbb{1}_{i-1} \otimes |\psi_i\rangle\langle\psi_i| \otimes \mathbb{1}_{m-i}]\big), \end{aligned} \tag{4.166}$$

by linearity of the trace, where we used $\mathrm{Tr}(\boldsymbol{\rho}) = 1$. This gives

$$F(\boldsymbol{\rho}, \psi_1 \otimes \cdots \otimes \psi_m) \geq 1 - \sum_{i=1}^{m} \big(1 - F(\rho_i, \psi_i)\big), \tag{4.167}$$

with Eqs. (4.164) and (4.165).

The right hand side of Eq. (4.162) is obtained by Cauchy-Schwarz inequality and a simple induction:

$$\begin{aligned} \mathrm{Tr}[\boldsymbol{\rho}\, |\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_m\rangle\langle\psi_m|] &= \mathrm{Tr}\left[\left(\sqrt{\boldsymbol{\rho}}\, |\psi_1\rangle\langle\psi_1| \otimes \mathbb{1}_{m-1}\right)\left(\mathbb{1} \otimes |\psi_2\rangle\langle\psi_2| \otimes \cdots \otimes |\psi_m\rangle\langle\psi_m|\, \sqrt{\boldsymbol{\rho}}\right)\right] \\ &\leq \mathrm{Tr}\left[\boldsymbol{\rho}\, |\psi_1\rangle\langle\psi_1| \otimes \mathbb{1}_{m-1}\right] \mathrm{Tr}\left[\boldsymbol{\rho}\, \mathbb{1} \otimes |\psi_2\rangle\langle\psi_2| \otimes \cdots \otimes |\psi_m\rangle\langle\psi_m|\right] \\ &\leq \ldots \\ &\leq \prod_{i=1}^{m} \mathrm{Tr}[\boldsymbol{\rho}\, \mathbb{1}_{i-1} \otimes |\psi_i\rangle\langle\psi_i| \otimes \mathbb{1}_{m-i}], \end{aligned} \tag{4.168}$$

where we used the cyclicity of the trace and the fact that $|\psi_1\rangle, \ldots, |\psi_m\rangle$ are pure states. This gives

$$F(\boldsymbol{\rho}, \psi_1 \otimes \cdots \otimes \psi_m) \leq \prod_{i=1}^{m} F(\rho_i, \psi_i), \tag{4.169}$$

with Eqs. (4.164) and (4.165).

Writing $F(\rho_i, \psi_i) = 1 - \epsilon_i$, we obtain, with Eqs. (4.167) and (4.169),

$$1 - \sum_{i=1}^{m} \epsilon_i \leq F(\boldsymbol{\rho}, \psi_1 \otimes \cdots \otimes \psi_m) \leq \prod_{i=1}^{m} (1 - \epsilon_i), \tag{4.170}$$

which concludes the proof. Additionnally, by the inequality of arithmetic and geometric means,

$$\begin{aligned} \prod_{i=1}^{m} (1 - \epsilon_i) &\leq \left(1 - \frac{1}{m} \sum_{i=1}^{m} \epsilon_i\right)^m \\ &\leq \exp\left(-\sum_{i=1}^{m} \epsilon_i\right), \end{aligned} \tag{4.171}$$

which gives a looser bound in terms of the total single-mode deviation $\sum_{i=1}^{m} \epsilon_i$:

$$1 - \sum_{i=1}^{m} \epsilon_i \leq F(\boldsymbol{\rho}, \psi_1 \otimes \cdots \otimes \psi_m) \leq \exp\left(-\sum_{i=1}^{m} \epsilon_i\right). \tag{4.172}$$

■

Note that Eq. (4.163) is tight for small ϵ, since its right hand side is then equivalent to $1 - m\epsilon$. Lemma 4.14 implies that if the fidelities of single-mode subsystems of an m-mode quantum state with some target pure states are higher than $1 - \frac{\lambda}{m}$, for some $\lambda > 0$, then the m-mode state has fidelity at least $1 - \lambda$ with the target m-mode product state.

Together with the union bound and the single-mode certification and verification protocols from Theorems 4.3 and 4.4, this provides a means for obtaining efficiently tight fidelity witnesses with any target tensor product of single-mode pure states with analytical confidence intervals, with and without i.i.d. assumption.

At this point, we can obtain fidelity witnesses only for pure product states, with no entanglement, using a fidelity estimation protocol for each of the single-mode subsystems in parallel. We make use of the properties of heterodyne detection in order to extend the class of target states for which fidelity witnesses can be efficiently obtained, from pure product states to the multimode states that are obtained from a pure product state with a passive linear transformation followed by single-mode Gaussian unitary operations, as in Eq. (4.161).

The POVM elements of product single-mode unbalanced heterodyne detection over m modes

with unbalancing parameters $\boldsymbol{\xi} \in \mathbb{C}^m$ are given by (see section 1.4.2)

$$\Pi_{\boldsymbol{\alpha}}^{\boldsymbol{\xi}} = \frac{1}{\pi^m} |\boldsymbol{\alpha}, \boldsymbol{\xi}\rangle\langle\boldsymbol{\alpha}, \boldsymbol{\xi}|, \tag{4.173}$$

for all $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m) \in \mathbb{C}^m$, where $|\boldsymbol{\alpha}, \boldsymbol{\xi}\rangle = \bigotimes_{i=1}^m |\alpha_i, \xi_i\rangle$ is a product of squeezed coherent states $\hat{S}(\xi_i)\hat{D}(\alpha_i)|0\rangle$.

The POVM elements of product single-mode balanced heterodyne detection are given by $\Pi_{\boldsymbol{\alpha}}^{\mathbf{0}}$, for all $\boldsymbol{\alpha} \in \mathbb{C}^m$, and we have $\Pi_{\boldsymbol{\alpha}}^{\boldsymbol{\xi}} = \hat{S}(\boldsymbol{\xi})\Pi_{\boldsymbol{\alpha}}^{\mathbf{0}}\hat{S}^\dagger(\boldsymbol{\xi})$. In particular, a single-mode squeezing followed by a single-mode balanced heterodyne detection can be simulated by performing directly an unbalanced heterodyne detection according to the squeezing parameter. One retrieves balanced heterodyne detection by setting the unbalancing parameter to 0 and homodyne detection by letting the modulus of the unbalancing parameter go to infinity.

Passive linear transformations correspond to unitary transformations of the creation and annihilation operators of the modes. These transformations, which may be implemented by unitary optical interferometers, map coherent states to coherent states: if $\hat{U}$ is a passive linear transformation and U is the unitary matrix describing its action on the creation and annihilation operators of the modes, an input coherent state $|\boldsymbol{\alpha}\rangle$ is mapped to an output coherent state $\hat{U}|\boldsymbol{\alpha}\rangle = |U\boldsymbol{\alpha}\rangle$, where $U\boldsymbol{\alpha}$ is obtained by multiplying the vector $\boldsymbol{\alpha}$ by the unitary matrix U. Hence, the POVM elements corresponding to a passive linear transformation $\hat{U}$ followed by a product of single-mode balanced heterodyne detection are given by $\hat{U}\Pi_{\boldsymbol{\alpha}}^{\mathbf{0}}\hat{U}^\dagger = \Pi_{U\boldsymbol{\alpha}}^{\mathbf{0}}$, for all $\boldsymbol{\alpha} \in \mathbb{C}^m$. This implies that the passive linear transformation $\hat{U}^\dagger$ followed by a product of single-mode heterodyne detections can be simulated by performing the heterodyne detections first, then multiplying the vector of samples obtained by U.

A similar property holds with single-mode displacements: since displacements map coherent states to coherent states, up to a global phase, by displacing their amplitude, a single-mode displacement followed by a single-mode heterodyne detection can be simulated by performing the heterodyne detection first, then translating the sample obtained according to the displacement amplitude. In particular we have $\hat{D}(\boldsymbol{\beta})\Pi_{\boldsymbol{\alpha}}^{\mathbf{0}}\hat{D}^\dagger(\boldsymbol{\beta}) = \Pi_{\boldsymbol{\alpha}+\boldsymbol{\beta}}^{\mathbf{0}}$ for all $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{C}^m$, where $\boldsymbol{\alpha} + \boldsymbol{\beta} = (\alpha_1 + \beta_1, \ldots, \alpha_m + \beta_m)$.

Combining the properties of heterodyne detection we obtain the following result:

Lemma 4.15. *Let $\boldsymbol{\beta}, \boldsymbol{\xi} \in \mathbb{C}^m$ and let $\hat{V} = \hat{S}(\boldsymbol{\xi})\hat{D}(\boldsymbol{\beta})\hat{U}$, where $\hat{U}$ is an m-mode passive linear transformation with $m \times m$ unitary matrix U. For all $\boldsymbol{\alpha} \in \mathbb{C}^m$, let $\boldsymbol{\gamma} = U\boldsymbol{\alpha} + \boldsymbol{\beta}$. Then,*

$$\Pi_{\boldsymbol{\gamma}}^{\boldsymbol{\xi}} = \hat{V}\Pi_{\boldsymbol{\alpha}}^{\mathbf{0}}\hat{V}^\dagger. \tag{4.174}$$

Proof. We have $\Pi_{\boldsymbol{\alpha}}^{\boldsymbol{\xi}} = \frac{1}{\pi^m} |\boldsymbol{\alpha}, \boldsymbol{\xi}\rangle\langle\boldsymbol{\alpha}, \boldsymbol{\xi}|$, for all $\boldsymbol{\alpha}, \boldsymbol{\xi} \in \mathbb{C}^m$, where $|\boldsymbol{\alpha}, \boldsymbol{\xi}\rangle = \hat{S}(\boldsymbol{\xi})\hat{D}(\boldsymbol{\alpha})|\mathbf{0}\rangle$ is a tensor

product of squeezed coherent states. We also have

$$\begin{cases} \hat{U}\Pi^{\mathbf{0}}_{\boldsymbol{\alpha}}\hat{U}^\dagger = \Pi^{\mathbf{0}}_{U\boldsymbol{\alpha}}, \\ \hat{D}(\boldsymbol{\beta})\Pi^{\mathbf{0}}_{\boldsymbol{\alpha}}\hat{D}^\dagger(\boldsymbol{\beta}) = \Pi^{\mathbf{0}}_{\boldsymbol{\alpha}+\boldsymbol{\beta}}, \\ \hat{S}(\boldsymbol{\xi})\Pi^{\mathbf{0}}_{\boldsymbol{\alpha}}\hat{S}^\dagger(\boldsymbol{\xi}) = \Pi^{\boldsymbol{\xi}}_{\boldsymbol{\alpha}}, \end{cases} \tag{4.175}$$

for all $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\xi} \in \mathbb{C}^m$ and all m-mode passive linear transformations $\hat{U}$ with $m \times m$ unitary matrix U. Writing $\hat{V} = \hat{S}(\boldsymbol{\xi})\hat{D}(\boldsymbol{\beta})\hat{U}$, we obtain

$$\begin{aligned} \hat{V}\Pi^{\mathbf{0}}_{\boldsymbol{\alpha}}\hat{V}^\dagger &= \hat{S}(\boldsymbol{\xi})\hat{D}(\boldsymbol{\beta})\hat{U}\Pi^{\mathbf{0}}_{\boldsymbol{\alpha}}\hat{U}^\dagger\hat{D}^\dagger(\boldsymbol{\beta})\hat{S}^\dagger(\boldsymbol{\xi}) \\ &= \hat{S}(\boldsymbol{\xi})\hat{D}(\boldsymbol{\beta})\Pi^{\mathbf{0}}_{U\boldsymbol{\alpha}}\hat{D}^\dagger(\boldsymbol{\beta})\hat{S}^\dagger(\boldsymbol{\xi}) \\ &= \hat{S}(\boldsymbol{\xi})\Pi^{\mathbf{0}}_{U\boldsymbol{\alpha}+\boldsymbol{\beta}}\hat{S}^\dagger(\boldsymbol{\xi}) \\ &= \Pi^{\boldsymbol{\xi}}_{U\boldsymbol{\alpha}+\boldsymbol{\beta}}. \end{aligned} \tag{4.176}$$

■

Lemma 4.15 implies that the POVM $\{\hat{V}\Pi^{\mathbf{0}}_{\boldsymbol{\alpha}}\hat{V}^\dagger\}_{\boldsymbol{\alpha}\in\mathbb{C}^m}$ can be simulated with the POVM $\{\Pi^{\boldsymbol{\xi}}_{\boldsymbol{\gamma}}\}_{\boldsymbol{\gamma}\in\mathbb{C}^m}$ by computing $\boldsymbol{\alpha} = U^\dagger(\boldsymbol{\gamma}-\boldsymbol{\beta})$, i.e., translating the vector of samples $\boldsymbol{\gamma}$ by the vector of complex amplitudes $-\boldsymbol{\beta}$ and multiplying the vector obtained by the $m \times m$ unitary matrix $U^\dagger$. This means that a passive linear transformation followed by single-mode Gaussian unitary operations before balanced heterodyne detection can be simulated by performing unbalanced heterodyne detection directly, then post-processing efficiently the classical outcomes. In particular, for such a transformation $\hat{V}$, if a multimode pure product state $\bigotimes_{i=1}^m |\psi_i\rangle$ can be efficiently verified using balanced heterodyne detection, then the state $\hat{V}(\bigotimes_{i=1}^m |\psi_i\rangle)$ can be efficiently verified using unbalanced heterodyne detection.

Formally, let ρ be an m-mode (mixed) state. Let $|\psi_1\rangle,\ldots,|\psi_m\rangle$ be single-mode pure states and let $\hat{V} = \hat{S}(\boldsymbol{\xi})\hat{D}(\boldsymbol{\beta})\hat{U}$, with $\boldsymbol{\beta},\boldsymbol{\xi} \in \mathbb{C}^m$, where $\hat{U}$ is a passive linear transformation over m modes with an associated $m \times m$ unitary matrix U. Then,

$$\begin{aligned} F(\boldsymbol{\rho}, \hat{V}|\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_m\rangle\langle\psi_m|\hat{V}^\dagger) &= F(\hat{V}^\dagger\boldsymbol{\rho}\hat{V}, |\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_m\rangle\langle\psi_m|) \\ &\geq 1 - \sum_{i=1}^m \Big(1 - F(|\psi_i\rangle\langle\psi_i|, (\hat{V}^\dagger\boldsymbol{\rho}\hat{V})_i)\Big), \end{aligned} \tag{4.177}$$

where we have used Lemma 4.14 and where $(\hat{V}^\dagger\boldsymbol{\rho}\hat{V})_i$ is the i^{th} single-mode reduced density matrix of the state $\hat{V}^\dagger\boldsymbol{\rho}\hat{V}$.

The single-mode fidelities $F(|\psi_i\rangle\langle\psi_i|, (\hat{V}^\dagger\boldsymbol{\rho}\hat{V})_i)$ can be estimated with analytical confidence intervals by measuring multiple copies of the m-mode state $\hat{V}^\dagger\boldsymbol{\rho}\hat{V}$ with product balanced heterodyne detection and post-processing the samples for individual subsystems according to the protocols from Theorems 4.3 and 4.4. By Lemma 4.15, this is equivalent to measuring the state $\boldsymbol{\rho}$ directly with a product of single-mode unbalanced heterodyne detections with unbalancing

parameters $\boldsymbol{\xi}$, translating the vector of samples $\boldsymbol{\gamma}$ obtained by the vector of complex amplitudes $-\boldsymbol{\beta}$ and multiplying the vector obtained by the unitary matrix $U^\dagger$. Then, the obtained samples may be post-processed according to the heterodyne certification or verification protocols.

If all the single-mode fidelity estimates obtained are precise to $\frac{1}{\text{poly}\, m}$ and greater than $1-\frac{1}{\text{poly}\, m}$ with high probability, which can be checked in time $O(\text{poly}\, m)$, then with the union bound for the failure probabilities, the fidelity between the m-mode state $\boldsymbol{\rho}$ and the target state $\hat{V}|\psi_1\rangle\langle\psi_1|\otimes\cdots\otimes|\psi_m\rangle\langle\psi_m|\hat{V}^\dagger$ is greater than $1-\frac{1}{\text{poly}\, m}$, with high probability. Hence the single-mode fidelity estimation protocols give a verification protocol for obtaining tight multimode fidelity witnesses, under or without the i.i.d. assumption.

The single-mode protocols from Theorems 4.3 and 4.4 are efficient as long as the energy of the single-mode target pure state is constant, i.e., it does not scale with the number of modes. Note that additional displacements may be introduced to reduce the energy of the single-mode target pure states, since by modifying their amplitudes these displacements can be braided through the transformation $\hat{V}$ and accounted for by translating the heterodyne detection samples. The efficiently verifiable states thus are the pure states of the form $\hat{S}(\boldsymbol{\xi})\hat{D}(\boldsymbol{\beta})\hat{U}(\bigotimes_{i=1}^m|\psi_i\rangle)$, such that for all $i\in\{1,\ldots,m\}$, the state $|\psi_i\rangle$ can be displaced onto a state of constant energy.

In particular, multimode Gaussian pure states with constant squeezing parameter can be efficiently verified, since these can be written as a product of pure single-mode squeezed coherent states followed by a passive linear transformation (see section 1.3). Note however that under the i.i.d. assumption the witnesses from [AGKE15] may provide a more efficient certification method for Gaussian states.

Remarkably, the class of efficiently verifiable states also includes the output states of CVS circuits and Boson Sampling interferometers. Our verification protocol may thus be used to verify quantum supremacy, as we detail in the following section.

4.7.2 Quantum supremacy with Boson Sampling: from validation to verification

The experimental demonstration of quantum computational supremacy is regarded as an important milestone in the field of quantum information. It involves a quantum device solving efficiently a computational task which is provably hard for classical computers, together with a verification of its correct functionning [HM17]. While the former has been recently accomplished with superconducting circuits [AAB$^+$19], the latter is still partial or relying on various computational assumptions.

Demonstrating quantum supremacy is inherently difficult because the computational task at hand is a sampling task from an anti-concentrating probability distribution over an exponential sample space. For that reason, direct non-interactive verification of quantum computational supremacy with a verifier restricted to classical computations is impossible [HKEG19]. Possible verification with a classical verifier includes interactive protocols with additional computational

assumptions [Reg09, AC16, Mah18], or partial verification, which ultimately relies on making assumptions about the inner functionning of the quantum device.

If one is reluctant to rely on additional assumptions, another way for performing verification is to allow the verifier to have quantum capabilities. However, the computational power of the verifier needs to be as small as possible, as it would not make sense if the verifier had enough computational power to perform the sampling task directly. In the context of discrete variable quantum computing, a minimal quantum capability would correspond to being able to prepare only single-qubit states or to perform only simple local measurements. For example, protocols for verification of IQP circuits [SB09] with these minimal requirements have been derived under the i.i.d. assumption with single-qubit states [MPKK17] or with local measurements [HKSE16] and more recently without the i.i.d. assumption with single-qubit states [KD19] or with local measurements [TM18].

In the context of continuous variable quantum computing, this minimal quantum capability would corresponds to being able to prepare only single-mode Gaussian states, or to perform only single-mode Gaussian measurements. An efficient certification protocol exists for verifying multimode Gaussian states [AGKE15] and thus instances of Gaussian Boson Sampling [HKS$^+$16] with single-mode Gaussian measurements under the i.i.d. assumption [AL18]. However, there is no efficient certification nor verification protocol using single-mode Gaussian measurements for Boson Sampling with input single photons: current methods used for validation of Boson sampling are either not scalable, e.g., computing the total variation distance with the ideal probability distribution, or else provide incomplete certificates, e.g., telling apart the tested distribution from classical mock-up distributions such as the uniform distribution [BGC$^+$19, WQD$^+$19].

When introducing the Boson Sampling model, Aaronson and Arkhipov importantly showed that even an approximate version of Boson Sampling is hard to sample for classical computers, provided two conjectures on the permanent of random Gaussian matrices hold true (see section 1.4.5 and [AA13]). More precisely, they showed under these conjectures that sampling from a probability distribution that has small constant total variation distance with an ideal Boson Sampling distribution is classically hard in the so-called antibunching regime $n = O(\sqrt{m})$. In particular, verifying a Boson Sampling quantum supremacy experiment amounts to verifying that the experimental quantum device samples from an ideal probability distribution, up to a constant error in total variation distance.

Our verification protocol derived in the previous section can be applied to check efficiently the fidelity of the output state of an experimental Boson Sampling interferometer with the ideal output state, using only balanced heterodyne detection. The fidelity witness gives in turn a certificate of the total variation distance with the ideal probability distribution for any observable by Eq. (4.89), therefore allowing for an experimental demonstration of quantum supremacy with Boson Sampling, with a verifier having minimal continuous variable quantum computational power, namely the ability to perform single-mode Gaussian measurements.

Performing verified Boson Sampling with our protocol, even under i.i.d. assumption, would already provide a convincing evidence of quantum supremacy with photonic quantum computing, as the verification without i.i.d. assumption only comes at the cost of an increased number of measurements, still polynomial in the number of modes m.

To that end, we optimise the bounds for the multimode certification protocol under i.i.d. assumption. In particular, the support estimation step in the protocol is no longer necessary, because the single-mode target pure states are either single-photon Fock states or vacuum states and thus have finite support over the Fock basis. We show that the number of copies needed for a constant additive precision in the antibunching regime $n = O(\sqrt{m})$—which is required for a demonstration of quantum supremacy—scales as $O(m^4 \log m)$, making reliable verification of Boson Sampling using single-mode Gaussian measurements within the reach of current experiments. Remarkably, this is only a logarithmic factor harder than verifying multimode Gaussian states [AGKE15].

Let $0 < \eta < 2/3$ and define, for all $z \in \mathbb{C}$,

$$f_0(z,\eta) = \frac{1}{\eta} \exp\left[\left(1 - \frac{1}{\eta}\right)|z|^2\right], \tag{4.178}$$

and

$$f_1(z,\eta) = \frac{1}{\eta^2}\left(\frac{|z|^2}{\eta} - 1\right)\exp\left[\left(1 - \frac{1}{\eta}\right)|z|^2\right]. \tag{4.179}$$

The verification protocol for Boson Sampling with n photons fed into a unitary interferometer U of size m under i.i.d. assumption reads:

1. The verifier chooses two precision parameters $0 < \eta_0, \eta_1 < 2/3$.

2. The verifier asks the prover for $N = O(m^4 \log m)$ copies of the target state $U|1\dots10\dots0\rangle$. Let $\boldsymbol{\rho}^{\otimes N}$ be the $(N \times m)$-mode (mixed) state sent by the prover, where $\boldsymbol{\rho}$ is an m-mode (mixed) state.

3. The verifier measures with balanced heterodyne detection all the m subsystems of all the N copies of $\boldsymbol{\rho}$, obtaining the N vectors of samples $\boldsymbol{\gamma}^{(1)}, \dots, \boldsymbol{\gamma}^{(N)} \in \mathbb{C}^m$.

4. For all $k \in \{1, \dots, N\}$, the verifier computes $\boldsymbol{\alpha}^{(k)} = U^\dagger \boldsymbol{\gamma}^{(k)}$. We write $\boldsymbol{\alpha}^{(k)} = (\alpha_1^{(k)}, \dots, \alpha_m^{(k)})$.

5. For all $i \in \{n+1, \dots, m\}$ the verifier computes the mean $\tilde{F}_i$ of the function $z \mapsto f_0(z, \eta_0)$ over the values $\alpha_i^{(1)}, \dots, \alpha_i^{(N)}$ and for all $j \in \{1, \dots, n\}$ the mean $\tilde{F}_j$ of the function $z \mapsto f_1(z, \eta_1)$.

6. The verifier computes $\tilde{W} = 1 - \sum_{i=1}^{m}(1 - \tilde{F}_i)$.

Theorem 4.6 (Certification of Boson Sampling using Gaussian measurements)**.** *$\tilde{W}$ is an estimate with constant precision of a tight lower bound on the fidelity with the ideal Boson Sampling output state, with probability exponentially close to* 1.

The estimate $\tilde{W}$ thus provides an efficient and reliable certificate of the total variation distance with the ideal probability distribution for any observable by Eq. (1.14).

Proof. Let $0 < \eta < 2/3$. By Lemma 4.1 we have, for any single-mode mixed state $\rho = \sum_{k,l\geq 0} \rho_{kl} |k\rangle\langle l|$,

$$\mathop{\mathbb{E}}_{\alpha\leftarrow Q_\rho(\alpha)}[f_0(\alpha,\eta)] = \mathrm{Tr}(\rho|0\rangle\langle 0|) + \eta \sum_{n=0}^{+\infty} \eta^n \rho_{n+1,n+1}, \tag{4.180}$$

and

$$\mathop{\mathbb{E}}_{\alpha\leftarrow Q_\rho}[f_1(\alpha,\eta)] = \mathrm{Tr}(\rho|1\rangle\langle 1|) + \eta \sum_{n=0}^{+\infty} \eta^n (n+2)\rho_{n+2,n+2}, \tag{4.181}$$

where $\mathop{\mathbb{E}}_{\alpha\leftarrow Q_\rho}[f]$ denotes the expected value of the function f for samples from single-mode balanced heterodyne detection of ρ.

Since $\eta \leq 2/3$ we have $\eta^{n+1} < \eta^n$ and $\eta^{n+1}(n+3) < \eta^n(n+2)$ for all $n \in \mathbb{N}$, so by a simple induction $\eta^n \leq 1$ and $\eta^n(n+2) \leq 2$, for all $n \in \mathbb{N}$. In particular, $\sum_{n=0}^{+\infty} \eta^n \rho_{n+1,n+1} \leq 1$ and $\sum_{n=0}^{+\infty} \eta^n (n+2)\rho_{n+2,n+2} \leq 2$, since $\mathrm{Tr}(\rho) = 1$. With Eqs. (4.180) and (4.181) we obtain

$$\left|\mathrm{Tr}(\rho|0\rangle\langle 0|) - \mathop{\mathbb{E}}_{\alpha\leftarrow Q_\rho}[f_0(\alpha,\eta)]\right| \leq \eta, \tag{4.182}$$

and

$$\left|\mathrm{Tr}(\rho|1\rangle\langle 1|) - \mathop{\mathbb{E}}_{\alpha\leftarrow Q_\rho}[f_1(\alpha,\eta)]\right| \leq 2\eta. \tag{4.183}$$

We also have, for all $z \in \mathbb{C}$,

$$0 < f_0(z,\eta) \leq \frac{1}{\eta}, \tag{4.184}$$

and

$$-\frac{1}{\eta^2} \leq f_1(z,\eta) \leq \frac{e^{\eta-2}}{\eta^2(1-\eta)}. \tag{4.185}$$

For $\eta < 2/3$, we have $\frac{e^{\eta-2}}{(1-\eta)} < 1$. In particular, the range of the function f_1 is less than $\frac{2}{\eta^2}$.

Let $N \in \mathbb{N}^*$, let $\alpha_1,\ldots,\alpha_N$ be i.i.d. samples from single-mode balanced heterodyne detection of a single mode state ρ_0 and let $\beta_1,\ldots,\beta_N$ be i.i.d. samples from single-mode balanced heterodyne detection of a single mode state ρ_1. Let $\epsilon_0,\eta_0,\epsilon_1,\eta_1 > 0$, by Hoeffding inequality,

$$\Pr\left[\left|\frac{1}{N}\sum_{p=1}^{N} f_0(\alpha_p,\eta_0) - \mathop{\mathbb{E}}_{\alpha\leftarrow Q_{\rho_0}}[f_0(\alpha,\eta_0)]\right| \geq \epsilon_0\right] \leq 2e^{-2N\epsilon_0^2\eta_0^2}, \tag{4.186}$$

and

$$\Pr\left[\left|\frac{1}{N}\sum_{p=1}^{N} f_1(\beta_p,\eta_1) - \mathop{\mathbb{E}}_{\beta\leftarrow Q_{\rho_1}}[f_1(\beta,\eta_1)]\right| \geq \epsilon_1\right] \leq 2e^{-\frac{N\epsilon_1^2\eta_1^4}{2}}. \tag{4.187}$$

Let $\boldsymbol{\sigma}$ be an m-mode state and let σ_i denote its k^{th} single-mode subsystem for all $k \in \{1,\ldots,m\}$. Let $\boldsymbol{\alpha}^{(1)},\ldots,\boldsymbol{\alpha}^{(N)} \in \mathbb{C}^m$ be samples from product balanced heterodyne detection of N identical copies of the state $\boldsymbol{\sigma}$. For all $p \in \{1,\ldots,N\}$, we write $\boldsymbol{\alpha}_p = (\alpha_1^{(p)},\ldots,\alpha_m^{(p)})$.

Combining Eqs. (4.182) and (4.186) we obtain, for any $i \in \{n+1,\dots,m\}$,

$$\left|\mathrm{Tr}(\sigma_i |0\rangle\langle 0|) - \frac{1}{N}\sum_{p=1}^{N} f_0\left(\alpha_i^{(p)}, \eta_0\right)\right| \leq \epsilon_0 + \eta_0, \tag{4.188}$$

with probability greater than $1-2\exp\left[-N\epsilon_0^2\eta_0^2\right]$. Similarly, combining Eqs. (4.183) and (4.187) we obtain, for any $j \in \{1,\dots,n\}$,

$$\left|\mathrm{Tr}(\sigma_j |1\rangle\langle 1|) - \frac{1}{N}\sum_{p=1}^{N} f_1\left(\alpha_j^{(p)}, \eta_1\right)\right| \leq \epsilon_1 + 2\eta_1, \tag{4.189}$$

with probability greater than $1-2\exp\left[-N\epsilon_1^2\eta_1^4/2\right]$.

We now choose $\epsilon_0, \eta_0, \epsilon_1, \eta_1$ in order to minimize the error probabilities for a given precision. Let $\lambda_0 > 0$. Setting $\epsilon_0 + \eta_0 = \lambda_0$, the optimal choice, which maximises $\epsilon_0^2\eta_0^2$, is $\epsilon_0 = \eta_0 = \frac{\lambda_0}{2}$ and with Eq. (4.188) we obtain, for any $i \in \{n+1,\dots,m\}$,

$$\left|\mathrm{Tr}(\sigma_i |0\rangle\langle 0|) - \frac{1}{N}\sum_{p=1}^{N} f_0\left(\alpha_i^{(p)}, \frac{\lambda_0}{2}\right)\right| \leq \lambda_0, \tag{4.190}$$

with probability greater than $1-2\exp\left[-\frac{N\lambda_0^4}{8}\right]$.

Let $\lambda_1 > 0$. Setting $\epsilon_1 + 2\eta_1 = \lambda_1$, the optimal choice, which maximises $\epsilon_1^2\eta_1^4$, is $\epsilon_1 = \eta_1 = \frac{\lambda_1}{3}$ and with Eq. (4.189) we obtain, for any $j \in \{1,\dots,n\}$,

$$\left|\mathrm{Tr}(\sigma_j |1\rangle\langle 1|) - \frac{1}{N}\sum_{p=1}^{N} f_1\left(\alpha_j^{(p)}, \frac{\lambda_1}{3m}\right)\right| \leq \lambda_1, \tag{4.191}$$

with probability greater than $1-2\exp\left[-\frac{N\lambda_1^6}{1458}\right]$.

Let us define the fidelity witness

$$W := 1 - \left(\sum_{k=n+1}^{m} 1 - F(\sigma_k, |0\rangle\langle 0|) + \sum_{l=1}^{n} 1 - F(\sigma_k, |1\rangle\langle 1|)\right), \tag{4.192}$$

as in Lemma 4.14, and the witness estimate

$$\tilde{W}(\lambda_0,\lambda_1) := 1 - \left(\sum_{i=n+1}^{m}\left[1 - \frac{1}{N}\sum_{p=1}^{N} f_0\left(\alpha_i^{(p)}, \frac{\lambda_0}{2}\right)\right] + \sum_{j=1}^{n}\left[1 - \frac{1}{N}\sum_{p=1}^{N} f_1\left(\alpha_j^{(p)}, \frac{\lambda_1}{3m}\right)\right]\right). \tag{4.193}$$

Taking the union bound of the failure probabilities for $i \in \{n+1,\dots,m\}$ and $j \in \{1,\dots,n\}$, we obtain with Eqs. (4.190) and (4.191),

$$\left|W - \tilde{W}(\lambda_0,\lambda_1)\right| \leq (m-n)\lambda_0 + n\lambda_1, \tag{4.194}$$

with probability greater than

$$1-2\left((m-n)\exp\left[-\frac{N\lambda_0^4}{8}\right]+n\exp\left[-\frac{N\lambda_1^6}{1458}\right]\right). \tag{4.195}$$

By Lemma 4.14 we have

$$F(\boldsymbol{\sigma},|1\ldots10\ldots0\rangle\langle1\ldots10\ldots0|)\geq W, \tag{4.196}$$

hence with Eq. (4.194) we obtain

$$F(\boldsymbol{\sigma},|1\ldots10\ldots0\rangle\langle1\ldots10\ldots0|)\geq\tilde{W}(\lambda_0,\lambda_1)-(m-n)\lambda_0-n\lambda_1, \tag{4.197}$$

with probability greater than

$$1-2\left((m-n)\exp\left[-\frac{N\lambda_0^4}{8}\right]+n\exp\left[-\frac{N\lambda_1^6}{1458}\right]\right). \tag{4.198}$$

Let $\boldsymbol{\rho}=\hat{U}^\dagger\boldsymbol{\sigma}\hat{U}$, where $\hat{U}$ is an m-mode passive linear transformation with $m\times m$ unitary matrix U. By Lemma 4.15 from the main text, the estimate $\tilde{W}$ can be computed using samples of product balanced heterodyne detection of $\boldsymbol{\rho}$ multiplied by the unitary matrix $U^\dagger$, rather than samples from the balanced heterodyne detection of $\boldsymbol{\sigma}$.
With Eqs. (4.197) and (4.198) we obtain

$$F(\boldsymbol{\rho},\hat{U}|1\ldots10\ldots0\rangle\langle1\ldots10\ldots0|\hat{U}^\dagger)\geq\tilde{W}(\lambda_0,\lambda_1)-(m-n)\lambda_0-n\lambda_1, \tag{4.199}$$

with probability greater than

$$1-2\left((m-n)\exp\left[-\frac{N\lambda_0^4}{8}\right]+n\exp\left[-\frac{N\lambda_1^6}{1458}\right]\right), \tag{4.200}$$

where $\tilde{W}(\lambda_0,\lambda_1)$ is computed using samples of product balanced heterodyne detection of N copies of the m-mode state ρ, each of the N vectors of samples being multiplied by the unitary matrix $U^\dagger$.

The values of λ_0 and λ_1 must be chosen by the verifier to maximise the above probability for a given precision of the witness. Equivalently, one needs to minimise:

$$(m-n)\exp\left[-\frac{N\lambda_0^4}{8}\right]+n\exp\left[-\frac{N\lambda_1^6}{1458}\right], \tag{4.201}$$

with the constraint $(m-n)\lambda_0+n\lambda_1=\epsilon$, for $\epsilon>0$. For a given experimental setup, Eq. (4.201) should be minimised depending on the values of m and n.

For example, setting $\lambda_0 = \frac{\epsilon}{2(m-n)}$ and $\lambda_1 = \frac{\epsilon}{2n}$ gives

$$F(\boldsymbol{\rho}, \hat{U}|1\dots10\dots0\rangle\langle1\dots10\dots0|\hat{U}^\dagger) \geq \tilde{W}\left(\frac{\epsilon}{2(m-n)}, \frac{\epsilon}{2n}\right) - \epsilon, \tag{4.202}$$

with probability greater than

$$1 - 2\left((m-n)\exp\left[-\frac{2N\epsilon^4}{[4(m-n)]^4}\right] + n\exp\left[-\frac{N\epsilon^6}{2(6n)^6}\right]\right), \tag{4.203}$$

and the estimate $\tilde{W}$ is ϵ-close to the actual fidelity witness, which by Lemma 4.14 is a tight witness of the fidelity. In particular, for a constant precision fidelity witness W, the estimate $\tilde{W}$ yields a constant precision fidelity witness with probability exponentially close (in m) to 1 for $N = O(\max\{(m-n)^4\log(m-n), n^6\log n\})$. In the antibunching regime $n = O(\sqrt{m})$, this means that the estimate has constant precision with exponentially small failure probability already for $N = O(m^4\log m)$.

■

A similar Boson Sampling verification protocol without i.i.d. assumption is obtained by using the version of the protocol which does not assume i.i.d. state preparation for estimating the single-mode fidelities in Theorem 4.4. This comes at the cost of an increased number of samples necessary for the same witness precision and confidence interval, which corresponds to a polynomial overhead in m and a slightly different classical post-processing: a small fraction of the measured subsystems have to be discarded at random and an additional support estimation step is necessary, for which the samples must be randomly chosen and cannot be used to compute the fidelity estimates. More precisely, changing the parameters of the protocol above to those of Theorem 4.4, we get that Theorem 4.6 holds without the i.i.d. assumption with polynomial confidence.

An interesting point is that by changing the unbalancing of heterodyne detection of the output modes of a Boson Sampling interferometer, one can switch between verification of Boson Sampling output states and demonstration of quantum sampling supremacy with continuous variable measurements. Indeed, CVS_{SP} circuits introduced in the previous chapter, which correspond to Boson Sampling with unbalanced heterodyne detection, are hard to sample classically when the unbalancing of the heterodyne detection is not too small (see section 3.3.3), but their output can be efficiently certified simply by switching to balanced heterodyne detection and computing a fidelity witness with the above method. This can be done within the same experimental setup using a reconfigurable beam splitter (Fig. 4.2) and showing the hardness of approximate CVS circuits sampling is an important step before an experimental demonstration.

Alternatively, by switching between balanced heterodyne detection and single-photon threshold detection, one can switch between verification of Boson Sampling output states and demon-

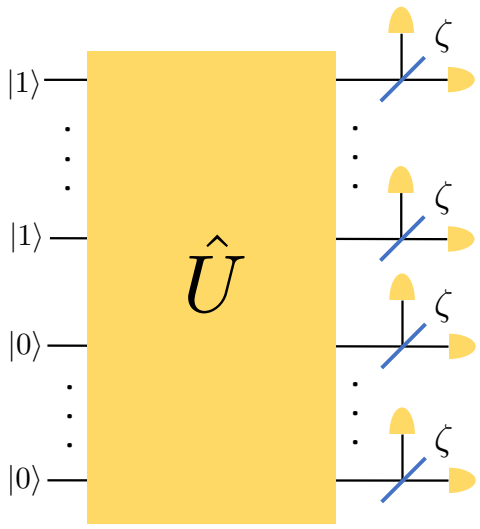

Figure 4.2: $\mathrm{CVS}_{\mathrm{SP}}$ circuit with interferometer U and heterodyne detection with reconfigurable unbalancing parameter ζ. The detectors represented are homodyne detectors. Setting $\zeta = 0$ (balanced detection) allows us to certify efficiently the multimode output state with a fidelity witness. Setting $\zeta \neq 0$, with $|\zeta| = \Omega(2^{-\mathrm{poly}\, m})$ allows us to perform efficiently a sampling task which is hard for classical computers, unless the polynomial hierarchy collapses.

stration of quantum sampling supremacy with discrete variable measurements, for which approximate sampling hardness is demonstrated, assuming two conjectures on the permanent of random Gaussian matrices and the fact that the polynomial hierarchy of complexity classes does not collapse [AA13].

4.8 Discussion and open problems

Existing methods for building trust for continuous variable quantum states like homodyne quantum state tomography require many different measurement settings, and heavy classical post-processing. For that purpose, we have introduced a reliable method for heterodyne quantum state tomography, which uses heterodyne detection as a single Gaussian measurement setting and allows for the retrieval of the density matrix of an unknown quantum state with analytical confidence intervals, without the need for data reconstruction nor binning of the sample space. For data reconstruction methods such as Maximum Likelihood, errors from the reconstruction procedure are usually indistinguishable from errors coming from the tested quantum device. For that reason, such methods do not extend well to the task of verification, unlike our method.

Building on these tomography techniques and with the addition of cryptographic techniques such as the de Finetti theorem, we have derived a protocol for verifying various copies of a continuous variable quantum state, without i.i.d. assumption, with Gaussian measurements.

This protocol is robust, as it directly gives a confidence interval on an estimate of the fidelity between the tested state and the target pure state. We emphasize that, while the target state is pure, the tested state is not required to be pure. The general protocol may be tailored to different uses and assumptions, from tomography to verification, simply by changing the classical post-processing.

Our verification protocol is complementary to the approach of [TMM+19], in which a verifier performs continuous variable quantum computing by delegating the preparation of Gaussian cluster states to a prover and has to perform non-Gaussian measurements. In our approach, the measurement-only verifier may perform continuous variable quantum computing by delegating the preparation of non-Gaussian states to the prover and has to perform Gaussian measurement, which are much easier to perform experimentally.

Importantly, we have promoted our single-mode protocols for fidelity estimation to multimode protocols yielding fidelity witnesses, showing in particular how to verify output states of a Boson Sampling interferometer efficiently, either under the i.i.d. assumption or with no assumption whatsoever. These protocols open the way for the most rigourous experimental demonstration of quantum computational supremacy so far, with Boson Sampling.

An exciting open problem is whether the technique employed in this chapter for promoting single-mode fidelity estimation protocols to protocols providing multimode fidelity witnesses can be applied in other contexts, for example discrete variable quantum computing. This technique crucially relies on being able to revert efficiently, at the stage of classical post-processing, specific quantum operations (passive linear transformations in this case) after a specific measurement (heterodyne detection, i.e., sampling from the Husimi Q function in this case).

CHAPTER 5

QUANTUM-PROGRAMMABLE MEASUREMENTS WITH LINEAR OPTICS

Distinguishing two unknown quantum states is central to many quantum applications [MdW13], notably for entanglement testing [MKB05, WRD+06, HM13], quantum communication [BCWDW01, dB04, KDK17] and quantum machine learning [EAO+02, LMR13]. This task is referred to as *unknown quantum state discrimination*.

The ability to program a fixed computer to perform a variety of computations is especially important: we do not want to build a new physical device for every different computation. In particular, quantum-programmable devices are quantum machines that take additional quantum states in input as a program, which dictates the rest of the computation. It is not possible to build a fixed quantum computer which can be programmed to perform any quantum computation [NC97], but we can design quantum-programmable devices for a restricted set of computations, such as projective measurements.

In this chapter, we show a correspondence between unknown quantum state discrimination and quantum-programmable measurements, by generalising the celebrated swap test [BCWDW01] for quantum state discrimination to an unbalanced setting where multiple copies of only one of the two tested states are available.

Next, we also generalise a known link between the Hong–Ou–Mandel effect for partially distinguishable photons and the swap test [GECP13]: we present the Hadamard interferometer and show that it provides a scheme for performing unknown quantum state discrimination and quantum-programmable measurements with linear optics and single photons.

In order to reduce the experimental requirements for implementation, we consider the case of projective measurements onto coherent states and simplify the previously derived scheme. In this case, we perform a simple analysis of the consequences of experimental imperfections.

This chapter is based on [CDM+18, KCK+20].

5.1 Testing quantum states

In the previous chapter, we discussed the efficient characterization of a continuous variable quantum state, either by full tomographic reconstruction, by fidelity estimation with a target state, or by obtaining a fidelity witness with a target state. In some cases, however, one merely wants to test simple properties of quantum systems. Given two unknown quantum states, one of the simplest questions one may ask is whether these states are equal or not. In this section, we make a connection between unknown quantum state discrimination schemes and quantum-programmable projective measurement devices.

5.1.1 Quantum state discrimination: the swap test

The swap test [BCWDW01] provides a simple probabilistic tool to compare two unknown quantum states. It takes as input two quantum states $|\phi\rangle$ and $|\psi\rangle$ that are not entangled and outputs 0 with probability $\frac{1}{2} + \frac{1}{2}|\langle\phi|\psi\rangle|^2$ and 1 with probability $\frac{1}{2} - \frac{1}{2}|\langle\phi|\psi\rangle|^2$, where $\langle\phi|\psi\rangle$ is the inner product of the states $|\phi\rangle$ and $|\psi\rangle$. When the measurement outcome is 0 (resp. 1), we conclude that the states were identical (resp. different), up to a global phase.

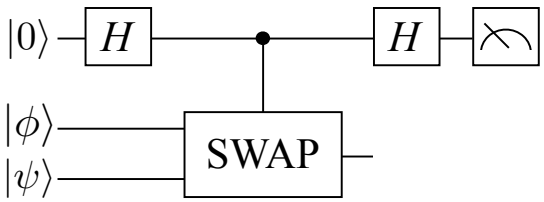

Figure 5.1: Circuit representation of a swap test. The ancilla qubit is measured in the computational basis.

A circuit implementing the swap test for qubits is represented in Fig. 5.1, where an ancilla is first prepared in the $|+\rangle$ state by a Hadamard gate

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{5.1}$$

and controls a swap between the two systems being tested.

The swap test meets the so-called *one-sided error requirement* [BF99], i.e., if the input states are identical, the test will always declare them as identical. On the other hand, if the input states are different, the test can obtain a wrong conclusion by declaring the states identical. The probability that this happens is strictly less than 1, hence by repeating the test various times, the probability that the sequence of tests never answers 1 can be brought down arbitrarily close to zero, exponentially fast. However, the swap test is destructive, in the sense that the output states of a previous test cannot be reused for a new test because they become maximally entangled

during the test [GECP13]. This means that in order to boost the correctness of the test in this manner, multiple copies of both states must be available.

5.1.2 Quantum state identity testing: generalised swap test

Let $m \geq 2$. We introduce the following generalisation of the swap test, in the context where one has access to various copies of a reference state $|\psi\rangle$ but to only a single copy of the other tested state $|\phi\rangle$:

Definition 5.1 (Swap test of order m)**.** The swap test of order m is a binary test that takes as input a state $|\phi\rangle$ and $m-1$ copies of a state $|\psi\rangle$, and outputs 0 with probability $\frac{1}{m} + \frac{m-1}{m}|\langle\phi|\psi\rangle|^2$ and 1 with probability $(\frac{m-1}{m})(1-|\langle\phi|\psi\rangle|^2)$. If the outcome 0 (resp. 1) is obtained, the test concludes that the states $|\phi\rangle$ and $|\psi\rangle$ were identical (resp. different).

Such a test clearly satisfies the one-sided error requirement.

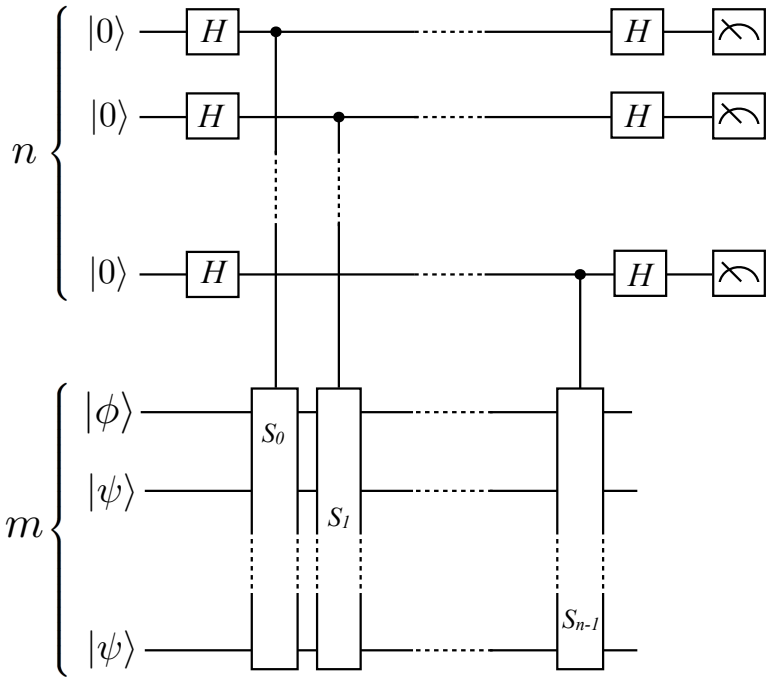

Figure 5.2: Swap circuit of order m. The unitaries S_k are tensor products of swap gates described in the main text (5.2). The $n = \log m$ ancilla qubits are measured in the computational basis at the end of the computation. The probability of obtaining 0 for all measurement outcomes is $\frac{1}{m} + \frac{m-1}{m}|\langle\phi|\psi\rangle|^2$.

In the following, we restrict to the swap test of order m when m is a power of 2, writing $n = \log m$. We introduce the swap circuit of order m (Fig. 5.2), that acts on m input qubits by applying n

consecutive layers of products of swap gates controlled by n ancilla qubits. These ancilla qubits are first initialised in the $|+\rangle$ state using Hadamard gates. Then, they are used as control qubits for the gates $S_0,\dots,S_{n-1}$, which can be applied in any order, where for all $k \in \{0,\dots,n-1\}$

$$S_k = \bigotimes_{\substack{i\in[0,2^k-1],\\ j\in[0,2^{n-k-1}-1]}} \mathrm{SWAP}\left[j2^{k+1}+i, j2^{k+1}+i+2^k\right], \tag{5.2}$$

with SWAP$[i,j]$ being the unitary operation that swaps the i^{th} and j^{th} qubits for $i,j \in \{0,\dots,m-1\}$. These controlled gates are applied to the input states $|\phi\rangle, |\psi\rangle, \dots, |\psi\rangle$ (one copy of a state $|\phi\rangle$ and $m-1$ copies of a state $|\psi\rangle$). Finally, a Hadamard gate is applied to each ancilla, which is then measured in the computational basis. By a simple induction, we obtain that the probability of obtaining the outcome 0 for all ancilla qubits is the squared norm of the following state:

$$\frac{1}{m}(|\phi\psi\dots\psi\rangle + |\psi\phi\dots\psi\rangle + \dots + |\psi\dots\psi\phi\rangle), \tag{5.3}$$

which only depends on the overlap between the states $|\phi\rangle$ and $|\psi\rangle$. More precisely,

$$\Pr[0,\dots,0] = \frac{1}{m} + \frac{m-1}{m}|\langle\phi|\psi\rangle|^2. \tag{5.4}$$

The swap circuit of order m thus implements the swap test of order m. Indeed, if the outcome $(0,\dots,0)$ is obtained, the test outputs 0 and we conclude that the states were identical, while for any other outcome the test outputs 1 and we conclude that the states were different. Note that in the case where $m=2$, the scheme reduces to the original swap test.

Because the $m-1$ last input states are identical, swapping them acts as the identity. This can be used to simplify the swap circuit of order m by replacing the $n = \log m$ layers of swap gates in Eq. (5.2) by the following n layers $S'_0,\dots,S'_{n-1}$, which have to be applied in this order:

$$S'_k = \bigotimes_{l=0}^{2^k-1} \mathrm{SWAP}\left[l, l+2^k\right]. \tag{5.5}$$

This reduces the total number of swap gates from $\frac{m\log m}{2}$ to $m-1$ without changing the number of ancilla qubits. This circuit has a simple structure of $n = \log m$ consecutive swap tests (Fig. 5.3).

For $k \in \{0,\dots,n-1\}$, conditioned on all the previous outputs being 0, the k^{th} swap test compares the output state of the previous test and the state $|\psi\rangle^{\otimes 2^k}$. Here, the swap test of two multipartite quantum states consists in applying a swap test to each of their corresponding subsystems. However, this multipartite swap test uses only a single ancilla qubit controlling the product of swap gates, as in Eq. (5.5), instead of an ancilla qubit for each pair of subsystems.

We now prove the optimality of the swap test of order m under the one-sided error requirement, i.e., we show that it achieves the lowest error probability in comparing states $|\phi\rangle$ and $|\psi\rangle$ given $m-1$ copies of $|\phi\rangle$ and one copy of $|\psi\rangle$ such that the one-sided error requirement is satisfied.

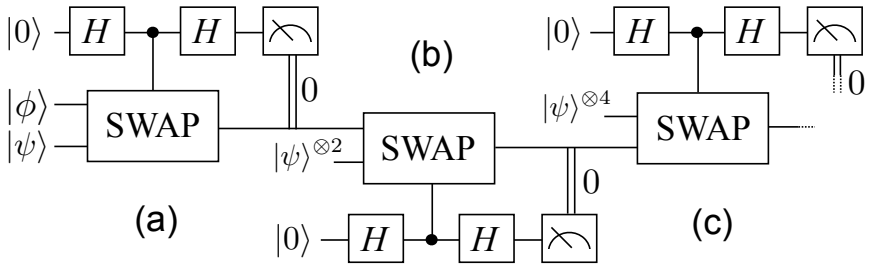

Figure 5.3: The simplified swap circuit of order m consisting in $n = \log m$ consecutive swap tests. (a) The first swap test compares the input states $|\phi\rangle$ and $|\psi\rangle$. (b) If this test is not able to tell apart the input states, i.e., if its outcome is 0, then the second swap test compares the bipartite output state of the first test with the state $|\psi\rangle^{\otimes 2}$. (c) If this test outcome is again 0, then the third swap test compares the quadripartite output state of the second test with the state $|\psi\rangle^{\otimes 4}$, and so on. If the n outcomes are 0, the test concludes that the states $|\phi\rangle$ and $|\psi\rangle$ were identical.

For this purpose, we first derive a more general result. In [KNY08], the authors consider the problem of testing if m quantum states are identical or not (the so-called *identity test*), with the promise that all the states are pairwise identical or orthogonal. In particular, they show that the optimal value for the error probability of any identity test with these assumptions satisfying the one-sided error requirement is $\frac{1}{m}$. We extend this result to the case where the states to be compared are no longer assumed pairwise identical or orthogonal:

Theorem 5.1. *Under the one-sided error requirement, any identity test of m unknown quantum states $|\psi_1\rangle, \ldots, |\psi_m\rangle$ has an error probability at least*

$$\frac{1}{m!} \sum_{\sigma \in S_m} \prod_{k=1}^{m} \langle \psi_k | \psi_{\sigma(k)} \rangle, \tag{5.6}$$

where S_m is the symmetric group over $\{1, \ldots, m\}$.

Proof. An identity test satisfying the one-sided error requirement can only be wrong when declaring identical (outputting 0) states that are not identical. Hence, to prove Theorem 5.1, it suffices to lower bound the probability of outputting 0 for any identity test. This is done by showing that the optimal identity test consists in a projection onto the symmetric subspace of the input states Hilbert space.

An identity test on a Hilbert space $\mathscr{H}$ is a binary test which can be written as a positive-operator valued measure $\{\Pi_0, \Pi_1\}$, with $\Pi_0 + \Pi_1 = \mathbb{1}$. Such a test takes as input a pure tensor product state $|\psi_1 \ldots \psi_m\rangle \in \mathscr{H}^{\otimes m}$ and outputs 0 with probability

$$\Pr[0] = \mathrm{Tr}[\Pi_0 |\psi_1 \ldots \psi_m\rangle \langle \psi_1 \ldots \psi_m|], \tag{5.7}$$

and 1 with probability

$$\Pr[1] = 1 - \Pr[0] = \mathrm{Tr}[\Pi_1 |\psi_1 \dots \psi_m\rangle \langle \psi_1 \dots \psi_m|]. \tag{5.8}$$

If the output 0 is obtained we conclude that we had $|\psi_1\rangle = \cdots = |\psi_m\rangle$, whereas if the output 1 is obtained we conclude that the states were not all identical. The one-sided error requirement can thus be written as

$$\forall |\psi\rangle, \mathrm{Tr}[\Pi_1 |\psi\rangle \langle\psi|^{\otimes m}] = 0. \tag{5.9}$$

Following [Har13], the symmetric subspace of $\mathscr{H}^{\otimes m}$ is characterised as

$$S = \mathrm{span}\{|\psi\rangle^{\otimes m} : |\psi\rangle \in \mathscr{H}\}, \tag{5.10}$$

and the orthogonal projector onto this space can be written as

$$P_S = \frac{1}{m!} \sum_{\sigma \in S_m} P_\sigma, \tag{5.11}$$

where for all $\sigma \in S_m$ and all $|\psi_1 \dots \psi_m\rangle \in \mathscr{H}^{\otimes m}$ we have $P_\sigma |\psi_1 \dots \psi_m\rangle = |\psi_{\sigma(1)} \dots \psi_{\sigma(m)}\rangle$. Given the characterisation of the symmetric subspace, the one-sided error requirement in Eq. (5.9) implies that the supports of P_S and Π_1 are disjoint. The support of P_S is thus included in the support of Π_0, given that $\Pi_0 + \Pi_1 = \mathbb{1}$ and this implies in turn that $\Pi_0 \geq P_S$ by positivity of Π_0.

The error probability of the identity test under the one-sided error requirement is given by the probability of outputting the result 0 while the states were not all identical:

$$\begin{aligned}
\Pr[0] &= \mathrm{Tr}[\Pi_0 |\psi_1 \dots \psi_m\rangle \langle \psi_1 \dots \psi_m|] \\
&\geq \mathrm{Tr}[P_S |\psi_1 \dots \psi_m\rangle \langle \psi_1 \dots \psi_m|] \\
&\geq \frac{1}{m!} \sum_{\sigma \in S_m} \mathrm{Tr}[P_\sigma |\psi_1 \dots \psi_m\rangle \langle \psi_1 \dots \psi_m|] \\
&\geq \frac{1}{m!} \sum_{\sigma \in S_m} \mathrm{Tr}[|\psi_{\sigma(1)} \dots \psi_{\sigma(m)}\rangle \langle \psi_1 \dots \psi_m|] \\
&\geq \frac{1}{m!} \sum_{\sigma \in S_m} \prod_{k=1}^{m} \langle \psi_k | \psi_{\sigma(k)} \rangle,
\end{aligned} \tag{5.12}$$

where in the third line we used the expression of the orthogonal projector P_S onto the symmetric subspace.

■

Applying Theorem 5.1 with $|\psi_1 \dots \psi_{k+l}\rangle = |\phi\rangle^{\otimes k} \otimes |\psi\rangle^{\otimes l}$, we obtain the following lower bound for

the error probability of any identity test of $k+l$ states $|\phi\rangle^{\otimes k} \otimes |\psi\rangle^{\otimes l}$:

$$\frac{1}{(k+l)!} \sum_{p=0}^{\min(k,l)} \binom{k}{p}\binom{l}{p} k!l! |\langle\phi|\psi\rangle|^{2p} = \sum_{p=0}^{\min(k,l)} \frac{\binom{k}{p}\binom{l}{p}}{\binom{k+l}{k}} |\langle\phi|\psi\rangle|^{2p}, \tag{5.13}$$

where $\binom{k}{p}\binom{l}{p}k!l!$ is the number of partitions of $\{1,\dots,k+l\}$ which map exactly $k-p$ elements of $\{1,\dots,k\}$ to elements of $\{1,\dots,k\}$. Testing quantum state identity with the input state $|\phi\rangle^{\otimes k} \otimes |\psi\rangle^{\otimes l}$ amounts to comparing the states $|\phi\rangle$ and $|\psi\rangle$ using k copies of $|\phi\rangle$ and l copies of $|\psi\rangle$.

In the case where $k=1$ and $l=m-1$, we have $|\psi_1 \dots \psi_m\rangle = |\phi\psi\dots\psi\rangle$ and Theorem 5.1 shows that the value $\frac{1}{m} + \frac{m-1}{m}|\langle\phi|\psi\rangle|^2$ is a lower bound for the error probability of any identity test of m states $|\phi\rangle, |\psi\rangle, \dots, |\psi\rangle$, i.e., one copy of a state $|\phi\rangle$ and $m-1$ copies of a state $|\psi\rangle$. With Definition 5.1 we directly obtain the following result:

Corollary 5.1. *The swap test of order m has optimal error probability $\frac{1}{m} + \frac{m-1}{m}|\langle\phi|\psi\rangle|^2$ under the one-sided error requirement.*

The swap circuit of order m is thus optimal for quantum state identity testing with an input $|\phi\rangle, |\psi\rangle, \dots, |\psi\rangle$, under the one-sided error requirement, since it implements the swap test of order m. In the next section, we show that the swap circuit of order m can be used to implement a programmable projective measurement.

5.1.3 Universal programmable measurements

In a typical experiment performing a quantum measurement, the choice of measurement is encoded in macroscopic, classical, information in the experimental setup. For example it can be encoded into the reflectance of a beam splitter, the phase in the branch of an interferometer or the spacial direction of a Stern Gerlach device. Often these choices are made beforehand and fixed. In some cases they can be programmed in a single set up (for example using thermo-optic phase shifters [CHS+15]). In all these cases, however, the choice of measurement basis is effectively programmed classically.

We consider the case where the choice of measurement is instead controlled by a quantum state. There are several reasons why one may consider a quantum state to control the choice of measurement. This state may be an output of a quantum computer, or a communication protocol, for example, which is not known before hand and only accessible as a quantum state. For example, in the cryptographic setting, non-orthogonal states can be used to remotely program a measurement which allows one to test the behaviour of a remote party. This is the essence behind the delegated blind verified quantum computation in [FK17]. At a fundamental level quantum programmable measurements separate as much as possible the choice of measurement basis and the bulk of the physical measurement apparatus, which could be interesting in probing foundational questions, for example in tests of contextuality where information about which measurements are being carried out leads to loopholes [Mey99, CK00, Win14].

A related and, in a sense, more general problem is that of a programmable quantum computer, where a quantum program state is used to encode a unitary to be run on a generic quantum computing device (gate array), first proposed by Nielsen and Chuang [NC97]. There it was shown that to do so deterministically requires orthogonal program states for every different unitary. To use the continuous parameters available in quantum states to encode more computations, the best one can do is probabilistic. In principle these techniques can be used to program quantum measurements. Indeed since the original proposal there have been several alternative schemes, extensions and applications, including programmable quantum state discriminators and measurements [VC00, DB02, RBCH03, ZB05, BBF+06]. These results, however, are either too general to consider the type of efficiency we show here, or specialized to tasks which are different from our setting.

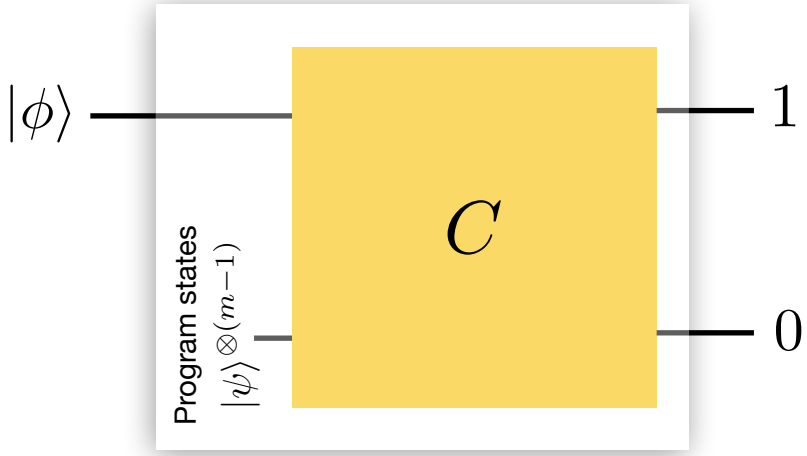

Figure 5.4: Programmable projective measurement. Given an input $|\phi\rangle$ and $m-1$ program registers $|\psi\rangle^{\otimes m-1}$ and allowing for possible ancillas (not pictured here), we apply some circuit C, independent of $|\psi\rangle$, and output a binary result where 0 is associated to projecting onto $|\psi\rangle$ and 1 to its complement.

We cast our problem as follows, illustrated in Fig. 5.4. One has $m-1$ program registers each prepared in the state $|\psi\rangle$ corresponding to the choice of measurement basis, and a single input register prepared in some state $|\phi\rangle$. Our aim is to output a classical bit corresponding to a projective measurement, where 0 represents the outcome $|\psi\rangle$ and 1 represents its complement. In an ideal measurement the result 0 would occur with probability $|\langle\phi|\psi\rangle|^2$. However, this is impossible for finite m. This follows from standard arguments based on the linearity of quantum mechanics, in analogy to necessity of orthogonal program states for computation mentioned above (see for example [NC97] for the case of programmable universal quantum computation, which easily extends to our case). We can thus only ever approximate perfect measurements. In our case

we parametrise this approximation by ϵ, requiring that the result 0 is returned with probability ϵ-close to $|\langle\phi|\psi\rangle|^2$.

We present a scheme which achieves this optimally in terms of how ϵ scales with m, under the condition that if the input is $|\psi\rangle$, the measurement always returns 0. This so-called one-sided error requirement [BF99] makes sense for various potential applications where it is important not to be wrong for this answer. One such example is the link between our scheme and the swap test [BCWDW01].

In the swap test, two unknown quantum states are compared using a controlled-swap operation. This test is especially relevant for the task of state discrimination. The general task of assessing if a set of m arbitrary states are identical has been addressed in [CAJ04, KNY08]. To solve this in generality requires controlled permutations for all possible permutations and therefore scales exponentially in circuit size. If one restricts oneself to the case where one has $m/2$ copies of one state and $m/2$ copies of the other, one can apply the construction in [KNY08] to get an optimal result. However, this scaling is not much better than simply doing the original swap test $m/2$ times, yet the corresponding test is much more difficult to implement.

From this point of view, the interesting cases of two states comparison is if one has an asymmetric number of one compared state compared to the other. In the most extreme case one would have just one copy of one state and $m-1$ copies of the other, which is exactly the case we consider for our programmable projective measurement, viewing the program state as the one we have many copies of. In particular, the $m=2$ case reduces to the swap test.

Given that a projective measurement with respect to a state $|\psi\rangle$ is a process that takes as input a state $|\phi\rangle$ and outputs 0 with probability $|\langle\phi|\psi\rangle|^2$ and 1 with probability $1-|\langle\phi|\psi\rangle|^2$, we introduce the notion of projective measurement with finite error:

Definition 5.2 (Approximate projective measurement)**.** Given a quantum state $|\psi\rangle$ and $\epsilon > 0$, a projective measurement with error ϵ with respect to the reference state $|\psi\rangle$ is a process that takes as input a quantum state $|\phi\rangle$ and outputs 0 with probability Pr[0] and 1 with probability Pr[1], such that $|\Pr[0]-(|\langle\phi|\psi\rangle|^2)| \leq \epsilon$ and $|\Pr[1]-(1-|\langle\phi|\psi\rangle|^2)| \leq \epsilon$.

Note that the two conditions in the previous definition are equivalent, since $\Pr[0]+\Pr[1]=1$. It will thus suffice to consider, e.g., the first condition. In this context, under the one-sided error requirement, a projective measurement with any error ϵ always outputs 0 if the input state is equal to the reference state.

Theorem 5.2. *A swap circuit of order m can be used to perform a projective measurement with error $\frac{1}{m}$ under the one-sided error requirement. Moreover, it is optimal in the sense that it uses the minimum number of copies of the reference state for achieving such an error.*

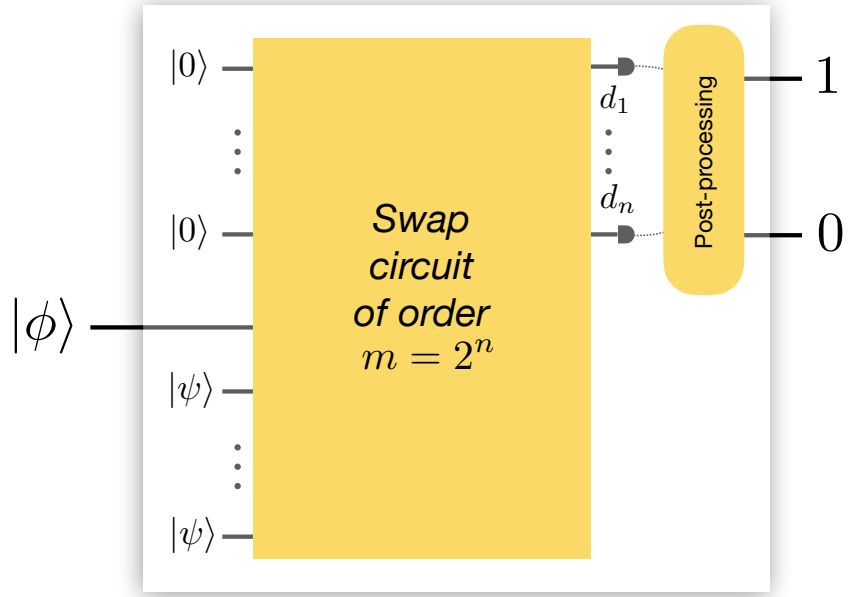

Figure 5.5: The swap circuit of order m used as a programmable projective measurement device. It takes as input a state $|\phi\rangle$ and the internal measurement outcomes are post-processed such that the device outputs 0 with probability $\frac{1}{m} + \frac{m-1}{m}|\langle\phi|\psi\rangle|^2$ and 1 with probability $\frac{m-1}{m}(1-|\langle\phi|\psi\rangle|^2)$. The programmable resource is the state $|\psi\rangle$ and the process uses $m-1$ copies of this state as well as $n = \log m$ ancillas.

Proof. For the swap circuit of order m, we have $\Pr[0,\ldots,0] = \frac{1}{m} + \frac{m-1}{m}|\langle\phi|\psi\rangle|^2$ by Eq. (5.4), so we can consider the whole circuit except the state $|\phi\rangle$ as a black box in Fig. 5.2, and post-process the measurement outcomes $\boldsymbol{d}$ as follows: if $\boldsymbol{d} = (0,\ldots,0)$, output 0, and output 1 otherwise. The setup now takes a single state $|\phi\rangle$ in input and outputs 0 with probability $\Pr[0] = \frac{1}{m} + \frac{m-1}{m}|\langle\phi|\psi\rangle|^2$, and 1 with probability $\Pr[1] = 1 - \Pr[0]$. We have $|\Pr[0] - (|\langle\phi|\psi\rangle|^2)| \leq \frac{1}{m}$ and when $|\phi\rangle = |\psi\rangle$, we have $\Pr[0] = 1 = |\langle\phi|\psi\rangle|^2$, hence this device performs a projective measurement with error $\frac{1}{m}$ and meets the one-sided error requirement.

We now prove the optimality of this device in terms of resources, i.e., we show that any device implementing a projective measurement with error $\frac{1}{m}$ and meeting the one-sided error requirement cannot use less than $m-1$ copies of the reference state.

We consider a device that implements a projective measurement with error ϵ, with respect to a reference state $|\psi\rangle$, using p copies of this reference state. This device takes as input a quantum state $|\phi\rangle$ and outputs 0 with probability $\Pr_\phi[0]$ and 1 with probability $\Pr_\phi[1] = 1 - \Pr_\phi[0]$. By Definition 5.2, the probability of outputting 0 satisfies $|\Pr_\phi[0] - (|\langle\phi|\psi\rangle|^2)| \leq \epsilon$. When the input state $|\phi\rangle$ is orthogonal to the reference state $|\psi\rangle$, the probability $\Pr_{\phi,\perp}[0]$ of

outputting 0 thus satisfies

$$\Pr_{\phi,\perp}[0] \leq \epsilon. \tag{5.14}$$

On the other hand, we can use this device to perform an identity test of $p+1$ states $|\phi\rangle, |\psi\rangle, \ldots, |\psi\rangle$ (one copy of the state $|\phi\rangle$ and p copies of the state $|\psi\rangle$): if the output 0 (resp. 1) is obtained we conclude that the states were identical (resp. different). This device meets the one-sided error requirement, so by Theorem 5.1 it has error probability at least $\frac{1}{p+1} + \frac{p}{p+1}|\langle\phi|\psi\rangle|^2$. This error probability corresponds to the probability of outputting 0 when the input states are different. In particular, when the input state $|\phi\rangle$ is orthogonal to the reference state $|\psi\rangle$, the probability $\Pr_{\phi,\perp}[0]$ of outputting 0 thus satisfies

$$\Pr_{\phi,\perp}[0] \geq \frac{1}{p+1}. \tag{5.15}$$

Combining both inequalities (5.14) and (5.15) we obtain $\frac{1}{p+1} \leq \epsilon$ or equivalently $p \geq \frac{1}{\epsilon} - 1$. For $\epsilon = \frac{1}{m}$, this amounts to $p \geq m - 1$, which completes the proof.

■

Theorem 5.2 implies that given a large enough swap circuit and the ability to produce many copies of a state $|\psi\rangle$, one can projectively measure any state with respect to the state $|\psi\rangle$ up to arbitrary small error. This error scales as the inverse of the number of copies. The circuit can thus be used as a programmable projective measurement device, where the programmable resource is the reference state $|\psi\rangle$ whose number of copies can be adjusted to control the precision of the measurement (Fig. 5.5).

The implementation of the swap circuit of order m is however challenging, due to the presence of many controlled-swap gates. In order to lower the implementation requirements, we study in the next section the linear optical Hadamard interferometer [Cre15, COR+16] and show that its statistics can be efficiently post-processed to reproduce those of a swap circuit of order m, without the need for ancillas. This comes at the cost that the device no longer has a quantum output, which however does not matter for most applications. In particular, we show that the Hadamard interferometer provides a simple linear optical platform for implementing the programmable projective measurement that we have described.

5.2 Universal programmable projective measurements with linear optics

The swap test has been shown equivalent to the linear optical Hong-Ou-Mandel effect [GECP13] (see section 1.4.4), in the sense that one can use Hong-Ou-Mandel effect to perform a state discrimination test between two partially distinguishable photons, whose statistics reproduce

those of a swap test. Generalising this equivalence, we present a practical solution to our problem with linear optics, using the Hadamard interferometer [Cre15, COR+16].

5.2.1 The Hadamard interferometer

In what follows, we consider optical unitary interferometers of size m which take as input one single photon in a quantum state $|\phi\rangle$ in the first mode and $m-1$ indistinguishable single photons in a state $|\psi\rangle$, one in each other spatial mode (the spatial modes of the interferometers are indexed from 1 to m). These states should be thought of as encoded in additional degrees of freedom of the photons (e.g., polarisation, time bins). The output modes are measured using photon number-resolving detection.

There exist complex amplitudes α and β and a state $|\psi^\perp\rangle$ with $\langle\psi|\psi^\perp\rangle = 0$ such that

$$|\phi\rangle = \alpha\,|\psi\rangle + \beta\,|\psi^\perp\rangle\,, \tag{5.16}$$

where $\alpha = \langle\psi|\phi\rangle$ and $|\alpha|^2 + |\beta|^2 = 1$. We have the following homomorphism property for single photon states:

$$|1_\phi\rangle = |1_{\alpha\psi+\beta\psi^\perp}\rangle = \alpha\,|1_\psi\rangle + \beta\,|1_{\psi^\perp}\rangle\,, \tag{5.17}$$

where for any state $|\chi\rangle$, $|1_\chi\rangle$ is the state of a single photon encoding the state $|\chi\rangle$. The single photon encoding maps identity of quantum states to distinguishability of single photons. In order to test the distinguishability of the photons, we look for detection events that do not occur when the photons are indistiguishable. In that case, it suffices to compute the output statistics separately when $|\phi\rangle = |\psi\rangle$ (*indistinguishable case*) and when $|\phi\rangle = |\psi^\perp\rangle$ (*distinguishable case*) to obtain the output statistics in the general case by linearity. The probability of detecting a photon number pattern $\boldsymbol{d} = (d_1,\ldots,d_m)$ which does not occur in the indistiguishable case, or equivalently that the k^{th} detector detects d_k photons for all $k \in \{1,\ldots,m\}$, is then

$$\begin{aligned}\Pr(\boldsymbol{d}) &= |\alpha|^2 \Pr_i(\boldsymbol{d}) + |\beta|^2 \Pr_d(\boldsymbol{d}) \\ &= \left(1 - |\langle\phi|\psi\rangle|^2\right) \Pr_d(\boldsymbol{d}),\end{aligned} \tag{5.18}$$

where $\Pr_i(\boldsymbol{d}) = 0$ is the probability in the indistinguishable case and $\Pr_d(\boldsymbol{d})$ is the probability in the distinguishable case. We thus have

$$\sum_{\substack{\boldsymbol{d} \\ \Pr_i(\boldsymbol{d})=0}} \Pr(\boldsymbol{d}) = \left(1 - |\langle\phi|\psi\rangle|^2\right) \sum_{\substack{\boldsymbol{d} \\ \Pr_i(\boldsymbol{d})=0}} \Pr_d(\boldsymbol{d}). \tag{5.19}$$

Note that for any measurement outcome $\boldsymbol{d} = (d_1,\ldots,d_m)$, we have $d_1 + \cdots + d_m = m$ since an interferometer is a passive device that does not change the total number of photons. For any interferometer of size m, we also obtain the following result:

Lemma 5.1. *For any detection pattern $\boldsymbol{d}$,*

$$Pr_d(\boldsymbol{d}) \geq \frac{Pr_i(\boldsymbol{d})}{m}, \tag{5.20}$$

Proof. We consider optical unitary interferometers of size m which take as input one single photon in a quantum state $|\phi\rangle$ and $m-1$ indistinguishable single photons in a state $|\psi\rangle$, one in each spatial mode, indexed from 1 to m. The output modes are measured using photon number detection. A measurement outcome thus has the form $\boldsymbol{d} = (d_1, \ldots, d_m)$, with $d_1 + \cdots + d_m = m$.

Recall that the permanent of an $m \times m$ matrix $A = (a_{ij})_{1 \leq i,j \leq m}$ is defined by

$$\mathrm{Per}(A) = \sum_{\sigma \in S_m} \prod_{k=1}^{m} a_{k\sigma(k)}, \tag{5.21}$$

where S_m is the symmetric group over $\{1, \ldots, m\}$. We now compute $\mathrm{Pr}_i(\boldsymbol{d})$ and $\mathrm{Pr}_d(\boldsymbol{d})$ for all detection patterns $\boldsymbol{d}$.

In the indistinguishable case, m indistinguishable photons, one in each mode, are sent through a linear optical network described by an $m \times m$ unitary matrix $U = (u_{ij})_{1 \leq i,j \leq m}$. The probability of a detection event $\boldsymbol{d}$ can be computed as (see, section 1.4.5 and [AA13])

$$\mathrm{Pr}_i(\boldsymbol{d}) = \frac{|\mathrm{Per}(U_{\boldsymbol{d}})|^2}{\boldsymbol{d}!}, \tag{5.22}$$

where $\boldsymbol{d}! = d_1! \ldots d_m!$ and where $U_{\boldsymbol{d}}$ is the matrix obtained from U by repeating d_k times the k^{th} column for $k \in \{1, \ldots, m\}$.

In the distinguishable case, $m-1$ indistinguishable photons are sent in modes $2, \ldots, m$ through a linear optical network described by an $m \times m$ unitary matrix $U = (u_{ij})_{1 \leq i,j \leq m}$, along with one additional photon in the first mode in an orthogonal state. Since it is fully distinguishable from the others, the additional photon behaves independently, hence the probability of detecting the photon number pattern $\boldsymbol{d}$ for one distinguishable photon and $m-1$ indistinguishable photons in input is

$$\mathrm{Pr}_d(\boldsymbol{d}) = \sum_{\substack{k=1 \\ d_k \neq 0}}^{m} \mathrm{Pr}_i(\boldsymbol{d} - \mathbf{1}_k) \cdot \mathrm{Pr}_i(\mathbf{1}_k), . \tag{5.23}$$

This last expression formalises the fact that the $m-1$ indistinguishable photons give a detection pattern $\boldsymbol{d} - \mathbf{1}_k$ which, completed by the additional distinguishable photon in the k^{th} output mode, forms the pattern $\boldsymbol{d}$. Developing this expression with Eq. (5.22) yields

$$\mathrm{Pr}_d(\boldsymbol{d}) = \frac{1}{\boldsymbol{d}!} \sum_{\substack{k=1 \\ d_k \neq 0}}^{m} d_k |u_{1k}\, \mathrm{Per}(U_{1,\boldsymbol{d}-\mathbf{1}_k})|^2 \tag{5.24}$$

where $U_{1,\boldsymbol{d}-\mathbf{1}_k}$ is the matrix obtained from U by removing the first row, then by repeating d_l times the l^{th} column for $l \neq k$ and by repeating $d_k - 1$ times the k^{th} column.

In order to obtain more readable expressions, we define for all $k \in \{1,\dots,m\}$ and for any detection pattern $\boldsymbol{d}$,

$$p_k(\boldsymbol{d}) = \begin{cases} \frac{u_{1k}\,\mathrm{Per}(U_{1,\boldsymbol{d}-\mathbf{1}_k})}{\sqrt{\boldsymbol{d}!}} \text{ if } d_k \neq 0, \\ 0 \text{ otherwise.} \end{cases} \tag{5.25}$$

Using the Laplace expansion of the permanent, the previous equations (5.22, 5.24) rewrite

$$\Pr_i(\boldsymbol{d}) = \left| \sum_{k=1}^{m} d_k p_k(\boldsymbol{d}) \right|^2, \tag{5.26}$$

and

$$\Pr_d(\boldsymbol{d}) = \sum_{k=1}^{m} d_k |p_k(\boldsymbol{d})|^2. \tag{5.27}$$

Since $\sum_{k=1}^{m} d_k = m$, we obtain, using Cauchy-Schwarz inequality with the complex vectors $\left\{\sqrt{d_k}\right\}_{1\leq k\leq m}$ and $\left\{\sqrt{d_k}\,p_k(\boldsymbol{d})\right\}_{1\leq k\leq m}$,

$$\Pr_d(\boldsymbol{d}) \geq \frac{\Pr_i(\boldsymbol{d})}{m}, \tag{5.28}$$

for any detection pattern $\boldsymbol{d}$.

■

For all $\boldsymbol{d}$ we have

$$\sum_{\substack{\boldsymbol{d} \\ \Pr_i(\boldsymbol{d})=0}} \Pr_d(\boldsymbol{d}) + \sum_{\substack{\boldsymbol{d} \\ \Pr_i(\boldsymbol{d})\neq 0}} \Pr_d(\boldsymbol{d}) = 1. \tag{5.29}$$

Combining Lemma 5.1 with Eqs. (5.19) and (5.29) yields

$$\sum_{\substack{\boldsymbol{d} \\ \Pr_i(\boldsymbol{d})\neq 0}} \Pr(\boldsymbol{d}) \geq \left(\frac{1}{m} + \frac{m-1}{m} |\langle \phi|\psi\rangle|^2 \right). \tag{5.30}$$

This last expression is valid for any interferometer and can be used it to retrieve, in the context of linear optics, the error probability bound for state identity testing under the one-sided error requirement obtained in Corollary 5.1. Indeed, assume that $\boldsymbol{d}$ is a detection event such that $\Pr_i(\boldsymbol{d}) \neq 0$, which could be a disjoint union of multiple detection events, used for an identity test: if $\boldsymbol{d}$ is obtained we conclude that the states were identical (or equivalently that the photons were indistinguishable), otherwise we assume that the states were different (or equivalently that the first photon was distinguishable from the others). The one-sided error requirement can thus be written as $\sum_{\boldsymbol{d},\Pr_i(\boldsymbol{d})\neq 0} \Pr_i(\boldsymbol{d}) = 1$: indistinguishable photons always pass the test. For different input states $|\phi\rangle$ and $|\psi\rangle$, the error probability of the corresponding test is then given by $\sum_{\boldsymbol{d},\Pr_i(\boldsymbol{d})\neq 0} \Pr(\boldsymbol{d})$, which by Eq. (5.30) is lower bounded by $\frac{1}{m} + \frac{m-1}{m} |\langle \phi|\psi\rangle|^2$.

We now study a particular unitary interferometer, when the size m is a power of 2: the Hadamard interferometer [Cre15, COR+16]. We show that it provides a simple implementation

of the swap test of order m. For $m = 4$ spatial modes (Fig. 5.6), this interferometer is described by the Hadamard-Walsh transform of order 2:

$$\frac{1}{\sqrt{2}}\begin{pmatrix} H & H \\ H & -H \end{pmatrix} \tag{5.31}$$

where H is a Hadamard matrix, see Eq. (5.1).

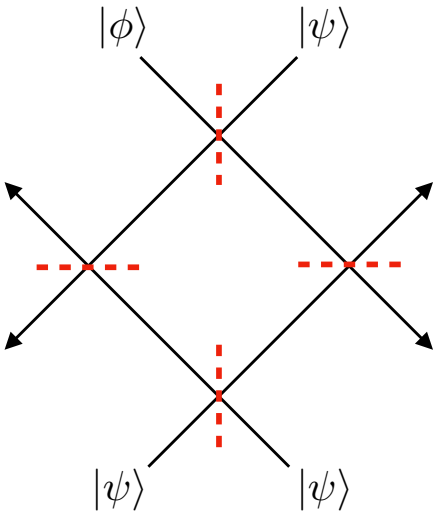

Figure 5.6: Hadamard interferometer with 4 input modes. The dashed red lines represent balanced beam splitters. The input states are one single photon in state $|\phi\rangle$ and three single photons in state $|\psi\rangle$, one in each mode.

In the general case, the Hadamard interferometer of order m is described by the Hadamard-Walsh transform of order $n = \log m$, which is defined by induction:

$$H_{k+1} = \frac{1}{\sqrt{2}}\begin{pmatrix} H_k & H_k \\ H_k & -H_k \end{pmatrix}, \tag{5.32}$$

with $H_0 = 1$ and $H_1 = H$. We can now state our main result linking the Hadamard interferometer and the swap test of order m.

Theorem 5.3. *The output statistics of the Hadamard interferometer of order m can be classically post-processed in time $O(m \log m)$ to reproduce those of the swap test of order m.*

Proof. Let us define

$$S = (s_{ij})_{1 \leq i,j \leq m} = \sqrt{m} H_n, \tag{5.33}$$

thus omitting the normalisation factor. We have

$$S = \underbrace{\sqrt{2}H \otimes \cdots \otimes \sqrt{2}H}_{n\ times}, \tag{5.34}$$

where H is a Hadamard matrix. The rows of $\sqrt{2}H$, together with the element-wise multiplication, form a group isomorphic to $\mathbb{Z}/2\mathbb{Z}$, thus the rows of S together with the element-wise multiplication form a group isomorphic to $(\mathbb{Z}/2\mathbb{Z})^n$. As a consequence, multiplying element-wise all the rows of S by its i^{th} row for a given i amounts to permuting the rows of S.

Let $\boldsymbol{d} = (d_1, \ldots, d_m)$ and $k \in \{1, \ldots, m\}$ such that $d_k \neq 0$. Let also $S_{\boldsymbol{d}-\mathbf{1}_k}$ be the matrix obtained from S by repeating d_l times the l^{th} column for $l \neq k$ and $d_k - 1$ the k^{th} column. For all $i \in \{1, \ldots, m\}$, one can obtain the matrix $S_{1,\boldsymbol{d}-\mathbf{1}_k}$ (with the first row removed) from the matrix $S_{i,\boldsymbol{d}-\mathbf{1}_k}$ (with the i^{th} row removed) by multiplying element-wise all rows by the i^{th} row and permuting the rows. Since the permanent is invariant by row permutation we obtain, for all $i \in \{1, \ldots, m\}$ and all $k \in \{1, \ldots, m1\}$ such that $d_k \neq 0$,

$$\mathrm{Per}(S_{i,\boldsymbol{d}-\mathbf{1}_k}) = \epsilon_{ik}(\boldsymbol{d})\mathrm{Per}(S_{1,\boldsymbol{d}-\mathbf{1}_k}), \tag{5.35}$$

where $\epsilon_{ik}(\boldsymbol{d}) = s_{ik}\prod_{j=1}^{m}\left(s_{ij}\right)^{d_j}$. Let us define for all $\boldsymbol{d} = (d_1, \ldots, d_m)$

$$\pi(\boldsymbol{d}) = \sum_{i=1}^{m}\prod_{j=1}^{m}\left(s_{ij}\right)^{d_j}. \tag{5.36}$$

We use the Laplace row expansion formula for the permanent of $S_{\boldsymbol{d}}$ to obtain, for all $\boldsymbol{d} = (d_1, \ldots, d_m)$ and all $k \in \{1, \ldots, m\}$ such that $d_k \neq 0$,

$$\begin{aligned}
\mathrm{Per}(S_{\boldsymbol{d}}) &= \sum_{i=1}^{m} s_{ik}\,\mathrm{Per}(S_{i,\boldsymbol{d}-\boldsymbol{d}1_k}) \\
&= \left(\sum_{i=1}^{m} s_{ik}\epsilon_{ik}(\boldsymbol{d})\right)\mathrm{Per}(S_{1,\boldsymbol{d}-\mathbf{1}_k}) \\
&= \left(\sum_{i=1}^{m}\prod_{j=1}^{m}\left(s_{ij}\right)^{d_j}\right)\mathrm{Per}(S_{1,\boldsymbol{d}-\mathbf{1}_k}) \\
&= \pi(\boldsymbol{d})\mathrm{Per}(S_{1,\boldsymbol{d}-\mathbf{1}_k}),
\end{aligned} \tag{5.37}$$

where we used Eq. (5.35) in the second line. With the general expressions of $\mathrm{Pr}_i(\boldsymbol{d})$ (5.22) and $\mathrm{Pr}_d(\boldsymbol{d})$ (5.24), this equation implies

$$m\mathrm{Pr}_i(\boldsymbol{d}) = \pi(\boldsymbol{d})^2\mathrm{Pr}_d(\boldsymbol{d}). \tag{5.38}$$

With the Laplace column expansion formula for the permanent of $S_{\boldsymbol{d}}$ and the last line of Eq. (5.37), we also obtain

$$m^2\mathrm{Pr}_i(\boldsymbol{d}) = \pi(\boldsymbol{d})^2\mathrm{Pr}_i(\boldsymbol{d}). \tag{5.39}$$

In particular, combining Eqs. (5.38) and (5.39),

$$m^2\pi(\boldsymbol{d})^2\mathrm{Pr}_d(\boldsymbol{d}) = \pi(\boldsymbol{d})^4\mathrm{Pr}_d(\boldsymbol{d}). \tag{5.40}$$

Now $\mathrm{Pr}_d(\boldsymbol{d})$ is non-zero for all $\boldsymbol{d}$, since by Eq. (5.24) it is a sum of moduli squared of permanents of $(2^n-1)\times(2^n-1)$ matrices, which in turn cannot vanish by a result of [SS83]. Hence the previous equation rewrites

$$m\pi(\boldsymbol{d}) = \pi(\boldsymbol{d})^2. \tag{5.41}$$

As a consequence, $\pi(\boldsymbol{d}) = m$ or $\pi(\boldsymbol{d}) = 0$ for all $\boldsymbol{d}$. Combining Eqs. (5.38) and (5.41) we obtain

$$\begin{aligned} \pi(\boldsymbol{d}) \neq 0 &\Leftrightarrow \pi(\boldsymbol{d}) = m \\ &\Leftrightarrow \mathrm{Pr}_i(\boldsymbol{d}) \neq 0 \\ &\Leftrightarrow \mathrm{Pr}_d(\boldsymbol{d}) = \frac{\mathrm{Pr}_i(\boldsymbol{d})}{m}, \end{aligned} \tag{5.42}$$

and thus

$$\begin{aligned} \mathrm{Pr}_i[\pi(\boldsymbol{d}) = m] &= \sum_{\pi(\boldsymbol{d})=m} \mathrm{Pr}_i(\boldsymbol{d}) \\ &= \sum_{\mathrm{Pr}_i(\boldsymbol{d})\neq 0} \mathrm{Pr}_i(\boldsymbol{d}) \\ &= 1. \end{aligned} \tag{5.43}$$

We also obtain

$$\begin{aligned} \mathrm{Pr}_d[\pi(\boldsymbol{d}) = m] &= \sum_{\pi(\boldsymbol{d})=m} \mathrm{Pr}_d(\boldsymbol{d}) \\ &= \frac{1}{m}\sum_{\pi(\boldsymbol{d})=} \mathrm{Pr}_i(\boldsymbol{d}) \\ &= \frac{1}{m}. \end{aligned} \tag{5.44}$$

We finally conclude by combining Eqs. (5.43), (5.44) and (5.19):

$$\begin{aligned} \mathrm{Pr}[\pi(\boldsymbol{d}) = m] &= \sum_{\pi(\boldsymbol{d})=m} \mathrm{Pr}(\boldsymbol{d}) \\ &= \frac{1}{m} + \frac{m-1}{m}\left|\langle\phi|\psi\rangle\right|^2. \end{aligned} \tag{5.45}$$

The post-processing (i.e., computing $\pi(\boldsymbol{d})$) can be done efficiently in time $O(m\log m)$ for any detection pattern $\boldsymbol{d} = (d_1,\ldots,d_m)$. Indeed, let $S_{\boldsymbol{d}}$ be the $m\times m$ matrix obtained from S by repeating d_k times the k^{th} column for $k\in\{1,\ldots,m\}$. The expression $\pi(\boldsymbol{d})$ in Eq. (5.36) is the sum of the product of the elements of each row of $S_{\boldsymbol{d}}$. Since the entries of the matrix S are only $+1$ and -1, $\pi(\boldsymbol{d}) = m$ if and only if the number of -1 on the rows of $S_{\boldsymbol{d}}$ is even for all rows. The condition $\pi(\boldsymbol{d}) = m$ can thus be written as a system of m linear equations modulo 2. Since $(\mathbb{Z}/2\mathbb{Z})^n$ is finitely generated by n elements, the m rows of $S_{\boldsymbol{d}}$ can be generated with

at most n rows using element-wise multiplication, for any measurement outcome $\boldsymbol{d}$. Hence, computing the parity of the number of -1 on each row of $S_{\boldsymbol{d}}$, which is equivalent to testing $\pi(\boldsymbol{d}) = m$, can be done by computing at most $n = \log m$ parity equations, with a number of terms in each equation which is at most m.

A simple induction shows that a possible choice for the rows whose parity has to be tested is the rows with index $2^k + 1$ for $k \in \{0, \ldots, n-1\}$ (the rows of the matrix being indexed from 1 to m).

■

Note that the group structure invoked in the proof is preserved under permutations, so Theorem 5.3 also applies to the unitary interferometers described by permutations of the Hadamard-Walsh transform.

The conclusion to be drawn from Theorem 5.3 is that as long as a state $|\psi\rangle$ can be encoded using single photons, then one can perform a swap test of order m with respect to the state $|\psi\rangle$ using the Hadamard interferometer of order m and an efficient classical post-processing of the measurement outcomes. The post-processing consists in the following parity test: given the measurement outcome $\boldsymbol{d} = (d_1, \ldots, d_m)$, where $d_1 + \cdots + d_m = m$, construct the matrix $S_{\boldsymbol{d}}$ from the matrix $S = \sqrt{m} H_n$ by keeping the k^{th} column only if d_k is odd. If the rows $(2, 3, 5, \ldots, 2^{n-1}+1)$ of $S_{\boldsymbol{d}}$ all have an even number of -1, output 0. Output 1 otherwise. This means that the post-processing only requires the parity of the photon number in each output mode.

In particular, the photon number-resolving detectors can be replaced by detecting the parity of the number of photons in each output mode. Detecting this parity can for example be achieved with microwave technology [HBR07, VKL$^+$13, SPL$^+$14]. Also only $m-1$ detectors are necessary, since the parity of the number of photon in the remaining mode can be deduced from the parities of the other modes, given that the total number of photons is m. If only the parity is measured, the discrimination test is non-destructive and the remaining single-mode state is a mixture of either even or odd photon-number states, depending on the measured parity and the total number of photons.

Using the argument developed in the proof of Theorem 5.2, by considering the $m-1$ photons and the interferometer as a black box (Fig. 5.7) whose outcomes are post-processed as described above, we also deduce the following result from Theorem 5.3:

Corollary 5.2. *The Hadamard interferometer of order m can be used to perform a projective measurement with error $\frac{1}{m}$, using a classical post-processing of its measurement outcomes that takes time $O(m \log m)$.*

Interestingly, the unitary interferometers described by the Hadamard-Walsh transform and its permutations are not the only unitary interferometers which can reproduce the statistics of a swap test with efficient post-processing, and indeed we present a generalisation in the next section. However, it is the simplicity of the Hadamard interferometer in terms of experimental

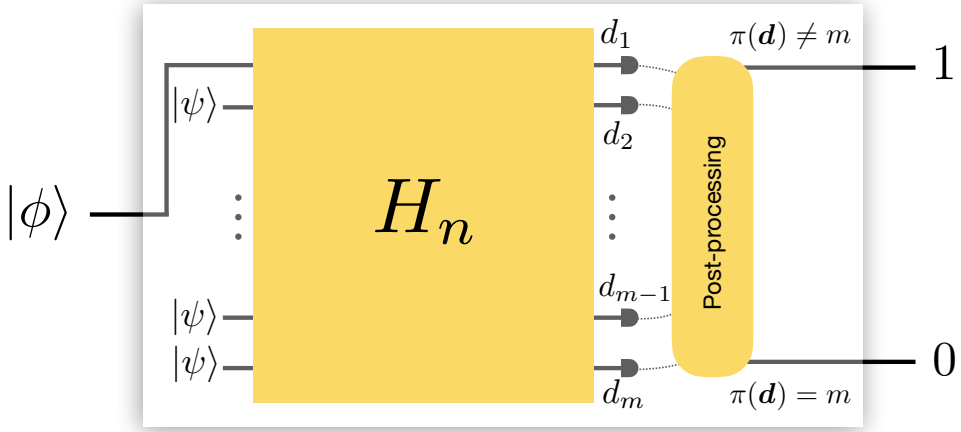

Figure 5.7: The Hadamard interferometer of order m used as a programmable projective measurement device. A single photon in the state $|\phi\rangle$ goes through a linear interferometer along with $m-1$ indistinguishable single photons in the state $|\psi\rangle$. The parity of the number of photons in each output mode is measured and efficiently post-processed, such that the device outputs 0 with probability $\frac{1}{m} + \frac{m-1}{m}|\langle\phi|\psi\rangle|^2$ and 1 with probability $\frac{m-1}{m}(1-|\langle\phi|\psi\rangle|^2)$.

implementation that motivates our interest towards this interferometer. In particular, this interferometer can be simply implemented with a few balanced beam splitters. A result by Reck *et al.* [RZBB94] states that any $m \times m$ unitary interferometer can be implemented using phase shifters and at most $\frac{m(m-1)}{2}$ beam splitters, possibly unbalanced. For the Hadamard interferometer, less beam splitters are needed and no phase shifters:

Lemma 5.2. *The $m \times m$ Hadamard interferometer can be implemented using $\frac{m \log m}{2}$ balanced beam splitters.*

Proof. The size m is a power of 2, with $n = \log m$. We prove by induction over n that there exist $P_0(n), \ldots, P_{n-1}(n)$ permutation matrices of order $m/2$, such that

$$H_n = \prod_{k=0}^{n-1} P_k(n)(\mathbb{1}_{m/2} \otimes H) P_k(n)^T. \tag{5.46}$$

Since multiplying matrices is equivalent to setting up experimental devices in sequence, and given that H is the matrix describing a balanced beam splitter, Eq. (5.46) implies the result we want to prove.

For $n = 1$, we have $m = 2$ and Eq. (5.46) is true with $P_0(1) = \mathbb{1}_1$. For brevity, we define for all k

$$H^{(k)} = \mathbb{1}_k \otimes H. \tag{5.47}$$

Assuming that Eq. (5.46) is true for n, we use the recursive definition of the Hadamard-Walsh

transform

$$H_{n+1} = H \otimes H_n, \tag{5.48}$$

along with properties of the tensor product of matrices in order to obtain

$$\begin{aligned} H_{n+1} &= (H_n \otimes \mathbb{1}_2) H^{(m)} \\ &= Q\,(\mathbb{1}_2 \otimes H_n) Q^T H^{(m)} \\ &= Q \left[\mathbb{1}_2 \otimes \prod_{k=0}^{n-1} P_k(n) H^{(m/2)} P_k(n)^T \right] Q^T H^{(m)} \\ &= Q \left[\prod_{k=0}^{n-1} (\mathbb{1}_2 \otimes P_k(n)) H^{(m)} \left(\mathbb{1}_2 \otimes P_k(n)^T \right) \right] Q^T H^{(m)} \\ &= \prod_{k=0}^{n-1} [Q\,(\mathbb{1}_2 \otimes P_k(n))] H^{(m)} [Q\,(\mathbb{1}_2 \otimes P_k(n))]^T H^{(m)}, \end{aligned} \tag{5.49}$$

where Q is a permutation matrix of order m and where in the third line we have used Eq. (5.46). Setting $P_k(n+1) = Q\,(\mathbb{1}_2 \otimes P_k(n))$ for $k \in \{0,\dots,n-1\}$ and $P_n(n+1) = \mathbb{1}_m$ proves Eq. (5.46) for $n+1$, since these matrices are permutation matrices of order m. This completes the induction and the proof of the result.

■

5.2.2 Group generalisation

Using the Hadamard interferometer requires the size parameter m to be a power of 2. This requirement can be relaxed, possibly raising the experimental requirements at the same time. Indeed, for any value of m, one can associate to any abelian group of order m an interferometer of size m which gives the desired statistics. This is the object of the following result that uses the invariant factor decomposition of an abelian group:

Theorem 5.4. *Let $\mathcal{G}$ be an abelian group of order m. Then, there exists $n \in \mathbb{N}^*$ and $a_1,\dots,a_n \in \mathbb{N}^*$, where $a_i | a_{i+1}$ for $i \in \{1,\dots,n-1\}$ and $a_1 \cdots a_n = m$, such that the interferometer described by the $m \times m$ unitary matrix*

$$U_{\mathcal{G}} := \frac{1}{\sqrt{m}} F_{a_1} \otimes \cdots \otimes F_{a_n}, \tag{5.50}$$

where $F_a = (e^{\frac{2i\pi}{a}(k-1)(l-1)})_{1\le k,l\le a}$ is the quantum Fourier transform of order a for all $a \in \mathbb{N}^*$, can perform a $\frac{1}{m}$-approximate projective measurement with a post-processing of its measurement outcomes that takes time at most mn. The rows of $F_{\mathcal{G}} := \sqrt{m} U_{\mathcal{G}}$ together with the element-wise multiplication form a group isomorphic to $\mathcal{G}$.

Proof. We use the notations of the theorem. The invariant factor decomposition of $\mathcal{G}$ gives

$$\mathcal{G} \simeq (\mathbb{Z}/a_1\mathbb{Z}) \otimes \cdots \otimes (\mathbb{Z}/a_n\mathbb{Z}), \tag{5.51}$$

where $n \in \mathbb{N}^*$ and $a_1, \ldots, a_n \in \mathbb{N}^*$ are unique, satisfying $a_i | a_{i+1}$ for $i \in \{1, \ldots, n-1\}$ and $a_1 \cdots a_n = m$. Given that the rows of F_a together with the element-wise multiplication form a group isomorphic to $(\mathbb{Z}/a\mathbb{Z})$ for all $a \in \mathbb{N}^*$, the rows of $F_{\mathscr{G}} = (f_{ij})_{1 \leq i,j \leq m} = \sqrt{m} U_{\mathscr{G}}$ together with the element-wise multiplication form a group isomorphic to $\mathscr{G}$.

Since the group structure was the only argument invoked in the proof of Theorem 5.3, the same conclusion can be drawn here, by following the same argument: for any detection event $\boldsymbol{d} = (d_1, \ldots, d_m)$,

$$\Pr[\pi(\boldsymbol{d}) = m] = \frac{1}{m} + \frac{m-1}{m} \left| \langle \phi | \psi \rangle \right|^2, \tag{5.52}$$

where

$$\pi(\boldsymbol{d}) = \sum_{i=1}^{m} \prod_{j=1}^{m} \left(f_{ij}\right)^{d_j}. \tag{5.53}$$

The group $\mathscr{G}$ is finitely generated by n elements, so n rows of $F_{\mathscr{G}}$ are sufficient to generate all its rows by element-wise multiplication. The condition $\pi(\boldsymbol{d}) = m$ can thus be checked in time $O(mn)$.

■

In particular, for $\mathscr{G} \simeq (\mathbb{Z}/m\mathbb{Z})$, the corresponding interferometer is described by the (normalised) quantum Fourier transform of order m, while for $\mathscr{G} \simeq (\mathbb{Z}/2\mathbb{Z})^n$, we retrieve Theorem 5.3 and the Hadamard interferometer.

5.3 Programmable projective measurements with coherent states

The previous scheme for performing programmable measurements with linear optics requires creation and manipulation of high-dimensional superposition states. In order to simplify the experimental requirements, we adapt this scheme to an encoding of quantum states in coherent states of light. Since coherent states are natural realisations of states produced by lasers, they can be efficiently produced and manipulated experimentally. The coherent state scheme takes as input a generic single-mode continuous variable quantum state, the test state, and $m-1$ copies of a coherent state in the program registers, and approximates the projective measurement $\{|\beta\rangle\langle\beta|, \mathbb{1} - |\beta\rangle\langle\beta|\}$ on the input state in a single run, using only threshold detectors. In particular, we obtain a more faithful projective measurement using coherent states than using a single-photon encoding.

In what follows, we introduce three different schemes for performing state discrimination and programmable projective measurement with coherent states: the *Hadamard scheme*, the *merger scheme*, and the *looped merger scheme*. Further, we give the proof for the optimality of the coherent state projective measurement performed by all three schemes, under the one-sided error requirement.

5.3.1 Coherent state discrimination

The swap test discriminates between two unknown states. If the unknown states are coherent states instead, then an analogous test can be performed by mixing the states on a balanced beam splitter and measuring the lower output branch with a single-photon threshold detector (Fig. 5.8).

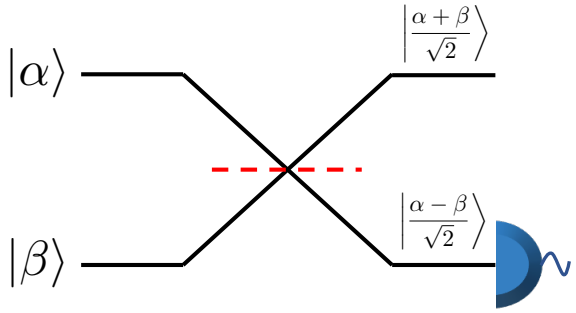

Figure 5.8: Balanced beam splitter operation on input states $|\alpha\rangle$ and $|\beta\rangle$. The second output mode is measured with a single-photon threshold detector. The probability of obtaining a click relates to the distinguishability of the two unknown coherent states.

A beam splitter maps the input modes creation operators $\{\hat{a}^\dagger, \hat{b}^\dagger\}$ to the output modes creation operators $\{\hat{c}^\dagger, \hat{d}^\dagger\}$ as

$$\begin{aligned} \hat{a}^\dagger &\to \frac{1}{\sqrt{2}}(\hat{c}^\dagger + \hat{d}^\dagger), \\ \hat{b}^\dagger &\to \frac{1}{\sqrt{2}}(\hat{c}^\dagger - \hat{d}^\dagger). \end{aligned} \tag{5.54}$$

The input state at the beam splitter is

$$|\alpha\rangle_a \otimes |\beta\rangle_b \,, \tag{5.55}$$

where the subscripts denote the mode in which the coherent states enter the beam splitter. In the absence of experimental imperfections, this yields the output state

$$\left|\frac{\alpha+\beta}{\sqrt{2}}\right\rangle_c \otimes \left|\frac{\alpha-\beta}{\sqrt{2}}\right\rangle_d , \tag{5.56}$$

for a balanced beam splitter. The probability of obtaining a click in the detector on the second mode is given by

$$1 - \exp\left(-\frac{|\alpha-\beta|^2}{2}\right) = 1 - |\langle\alpha|\beta\rangle|. \tag{5.57}$$

We now consider the scenario when one receives a single copy of an unknown coherent state $|\alpha\rangle$ and the objective is to check wether if the test state is equal to the reference coherent state $|\beta\rangle$. Here one has access to multiple copies of the reference state but is limited to just a single copy of the test state. In the simpler case, the state discrimination can be performed with a

single copy of the reference state, like in the previous section. This succeeds with a probability given by Eq. (5.57). In this section, we prove that having multiple copies of the reference state $|\beta\rangle$ increases the success probability of discriminating with the test state $|\alpha\rangle$. For this, we first provide a generalised interferometer construction, the *Hadamard scheme*, based on Hadamard transformations, following the previous section. We then show how this interferometer can be simplified, thanks to coherent state encoding and we introduce the *merger scheme* and the *looped merger scheme*.

5.3.2 The Hadamard scheme

We consider the Hadamard interferometer over m modes, where m is a power of 2. The input is now composed of coherent states:

$$|\alpha\rangle_1 \otimes |\beta\rangle_2 \otimes \cdots \otimes |\beta\rangle_m , \tag{5.58}$$

where the subscript denotes the mode in which the coherent state enters the generalised interferometer. For brevity, we address this state as $|\alpha\beta...\beta\rangle$. All the output modes but the first are measured with single-photon threshold detectors.

With Eq. (5.32), the Hadamard interferometer of order m is described by the Hadamard-Walsh transform of order $n = \log m$, which is defined by:

$$H_n = H^{\otimes n}, \tag{5.59}$$

with $H_0 = (1)$ and $H_1 = H$. The input coherent states $|\alpha\beta\dots\beta\rangle$ upon interaction with the interferometer of order m transform as,

$$|\alpha\beta\dots\beta\rangle \mapsto H_n |\alpha\beta\dots\beta\rangle = |\delta_1\delta_2\dots\delta_m\rangle , \tag{5.60}$$

where, with a simple induction, we obtain $\delta_1 = \frac{\alpha+(m-1)\beta}{\sqrt{m}}$ and $\delta_k = \frac{\alpha-\beta}{\sqrt{m}}$ for $k > 1$. Thus the last $m-1$ modes have the same probability of a click when measured with single-photon threshold detectors. The probability $\Pr[\varnothing]$ that none of the $m-1$ detectors clicks is,

$$\begin{aligned}
\Pr_{\alpha,\beta,m}[\varnothing] &= \prod_{k=1}^{m-1} (1 - \Pr[\text{click in } k^{\text{th}} \text{ mode}]) \\
&= \prod_{k=1}^{m-1} [1 - (1 - \exp(-|\delta_k|^2))] \\
&= \exp\left(-\frac{m-1}{m}|\alpha-\beta|^2\right) \\
&= (|\langle\alpha|\beta\rangle|^2)^{1-\frac{1}{m}}.
\end{aligned} \tag{5.61}$$

In particular, for all $\alpha, \beta \in \mathbb{C}$, $\Pr_{\alpha,\beta,+\infty}[\varnothing] = |\langle\alpha|\beta\rangle|^2$, which corresponds to a perfect projective measurement of the states $|\alpha\rangle$ and $|\beta\rangle$. Writing $x = |\langle\alpha|\beta\rangle|^2$ the overlap of the test and reference states, we obtain

$$\Pr_{x,m}[\varnothing] = x^{1-\frac{1}{m}}. \tag{5.62}$$

Assigning to the event 'none of the detectors clicks' the value 0 and to other detection events ('at least one of the $m-1$ detectors clicks') the value 1, we obtain a device whose statistics approach those of a projective measurement, with

$$\Pr_{x,m}[0] = 1 - \Pr_{x,m}[1] = x^{1-\frac{1}{m}}. \tag{5.63}$$

With Theorem 5.3 and Eq. (5.4), for an m-mode input state $|\phi\psi\ldots\psi\rangle$ the corresponding statistics with single-photon encoding are

$$\Pr_{x,m}[0] = 1 - \Pr_{x,m}[1] = \frac{1}{m} + \left(1 - \frac{1}{m}\right)x. \tag{5.64}$$

The single-photon encoding implies having $m-1$ number-resolving or parity detectors. On contrary, the encoding with coherent states requires $m-1$ single-photon threshold detectors. Experimentally, this is relatively easier to implement. The test based on coherent state also satisfies the one-sided error requirement: if the states are the same, then their trace distance is 0 and hence the probability of having the detection event 1 is 0. Moreover, coherent state encoding provides a more faithful projective measurement than single-photon encoding. Indeed, the statistics produced by coherent state encoding are closer to the ones of a perfect projective measurement. For any given value of the overlap x:

$$\forall x \in [0,1], \quad x \leq x^{1-\frac{1}{m}} \leq \frac{1}{m} + \left(1 - \frac{1}{m}\right)x. \tag{5.65}$$

In particular, for a given size m, the maximal statistical gap with a perfect projective measurement is,

$$\begin{aligned} e_{\mathrm{SP}}(m) &= \max_{x\in[0,1]} \left| \left[\frac{1}{m} + \left(1 - \frac{1}{m}\right)x \right] - x \right| \\ &= \frac{1}{m}, \end{aligned} \tag{5.66}$$

for the single-photon encoding, and

$$\begin{aligned} e_{\mathrm{CS}}(m) &= \max_{x\in[0,1]} \left| \left(x^{1-\frac{1}{m}}\right) - x \right| \\ &= \frac{(m-1)^{m-1}}{m^m} \\ &\sim \frac{1}{e} \cdot \frac{1}{m}, \end{aligned} \tag{5.67}$$

for the coherent state encoding, which is lower than the single-photon encoding gap. This happens because for the single-photon encoding no assumption is made about the states $|\phi\rangle$ and $|\psi\rangle$, while the states $|\alpha\rangle$ and $|\beta\rangle$ are assumed to be coherent states. This additional information about the states allows us to better approximate a perfect projective measurement with the same number of input states. We show in the next section that there exists a simpler measurement setting than the Hadamard interferometer, achieving the same performance in the test, due to coherent state encoding.

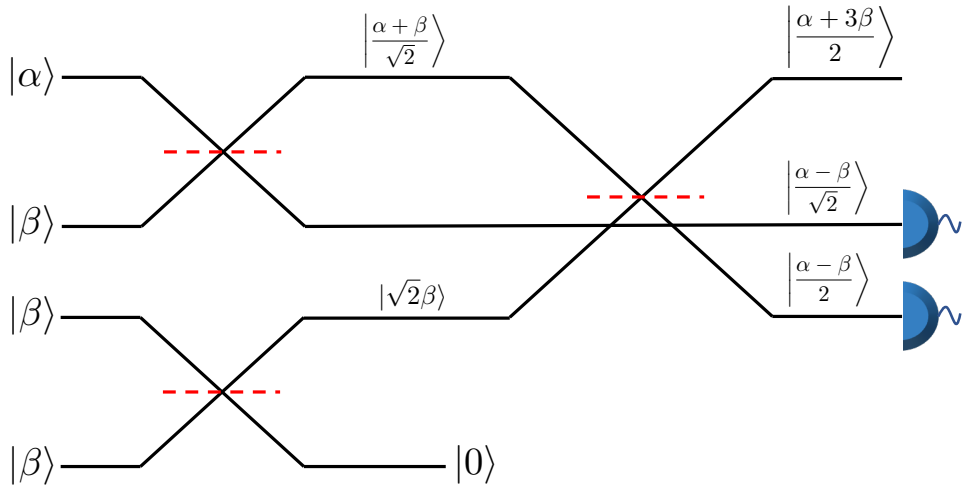

Figure 5.9: Merger scheme with 4 input modes. The input states are one tested state $|\alpha\rangle$ and three local states $|\beta\rangle$, one in each mode. The detectors are single-photon threshold detectors.

5.3.3 The merger scheme

By Lemma 5.2, the Hadamard scheme of size m uses $\frac{m \log m}{2}$ balanced beam splitters and $m-1$ single-photon threshold detectors. We introduce a simplified scheme over the same number of modes m, which only uses $m-1$ balanced beam splitters and $\log m$ detectors, and show that it achieves the same performance than the Hadamard scheme. We refer to this scheme as the *merger* scheme, since it merges identical input coherent states into an amplified coherent state in the first output mode and the vacuum in all other modes.

For $m = 4$ spatial modes, this interferometer acting on modes $\{1,2,3,4\}$ is described by the following unitary matrix:

$$\begin{aligned} U_4 &= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \\ &= H_{1,3} \times (H_{1,2} \oplus H_{3,4}), \end{aligned} \tag{5.68}$$

where $H_{i,j}$ corresponds to the balanced beam splitter operation acting on modes i and j (where the modes are indexed from 1 to m) and identity on the other modes (Fig. 5.9).

The generalised merger interferometer is defined by induction:

$$U_m = H_{1,m/2+1} \times (U_{m/2} \oplus U_{m/2}), \tag{5.69}$$

where $U_1 = H_{0,1} = H$ is a Hadamard matrix. This induction relation is illustrated in Fig. 5.10.

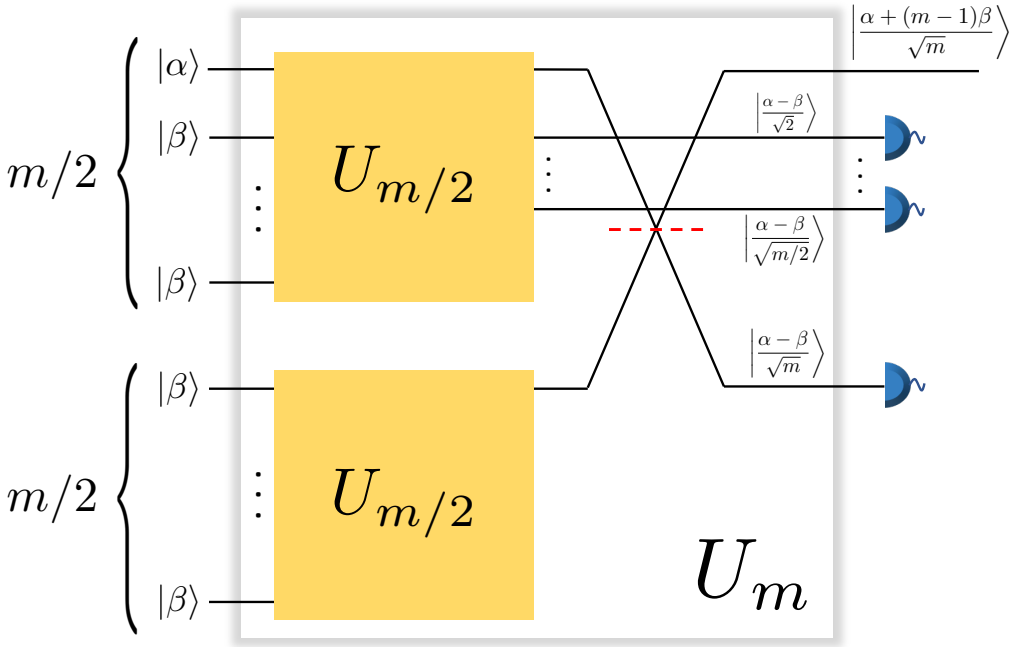

Figure 5.10: General merger scheme of size m, with one copy of $|\alpha\rangle$ and $m-1$ copies of $|\beta\rangle$: the first output modes of two interferometers described by $U_{m/2}$ are mixed on a balanced beam splitter. Indexing the spatial modes from 1 to m, the 2^k+1 output modes are measured with single-photon threshold detectors, for $k = 0 \ldots n-1$.

Indexing the spatial modes from 1 to m, the 2^k+1 output modes are measured with single-photon threshold detectors, for $k = 0 \ldots n-1$. A simple induction shows that the output state in the 2^k+1 output mode is $\left|\frac{\alpha-\beta}{\sqrt{2^{k+1}}}\right\rangle$. Hence, the probability that none of the $n = \log m$ detectors clicks is given by

$$\begin{aligned}
\Pr_{\alpha,\beta,m}[\varnothing] &= \prod_{k=0}^{n-1} (1 - \Pr[\text{click in the } 2^k th \text{ mode}]) \\
&= \prod_{k=0}^{n-1} \left[1 - \left(1 - \exp\left(-\left|\frac{\alpha-\beta}{2^{\frac{k+1}{2}}}\right|^2\right)\right)\right] \\
&= \exp\left(-\sum_{k=0}^{n-1} \left(\frac{1}{2}\right)^{k+1} |\alpha-\beta|^2\right) \\
&= \exp\left(-\frac{m-1}{m}|\alpha-\beta|^2\right) \\
&= (|\langle\alpha|\beta\rangle|^2)^{1-\frac{1}{m}},
\end{aligned} \tag{5.70}$$

thus retrieving the statistics obtained with the Hadamard scheme, using only $n = \log m$ detectors. Moreover, a simple induction shows that the merger interferometer can be implemented with only $m-1$ balanced beam splitters.

Noting the recursive character of the merger scheme, we present another possible implementation of the merger scheme using a looped beam splitter interaction, one single-photon threshold detector and an active optical element, namely an active amplitude modulator (Fig. 5.11).

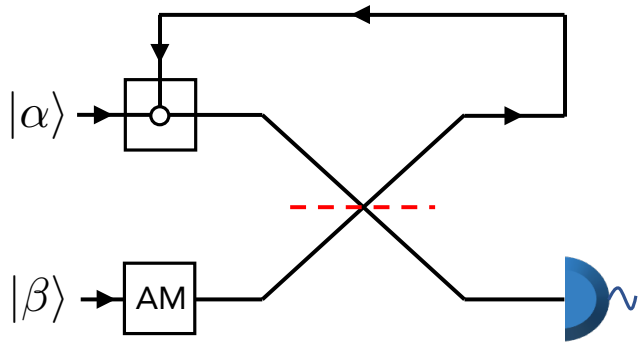

Figure 5.11: The looped merger scheme. The coherent pulse $|\alpha\rangle$ is sent once, while at each loop a new coherent pulse $|\beta\rangle$ is sent. An optical switch ensures a closed loop after the first pulse $|\alpha\rangle$ passes through. An amplitude modulator (AM) transforms the k-th pulse $|\beta\rangle$ to $|\sqrt{2^k}\beta\rangle$, k starting at 0.

This setup now uses an active optical element and a constant number of passive linear optical elements and approximates a perfect projective measurement up to arbitrary precision. By construction, the statistics of the setup after $m-1$ pulses $|\beta\rangle$ sent reproduce those of the merger scheme of size m.

The three schemes discussed provide experimentally-friendly devices to perform a variety of quantum information processing tasks using coherent states, ranging from state discrimination to programmable projective measurements, in a non-destructive manner. These schemes are also optimal for coherent state discrimination:

Theorem 5.5. *The Hadamard interferometer and the merger interferometer are optimal for coherent states discrimination, under the one-sided error requirement.*

> **Proof.** The proof extends results from [SZP+07]. We first start by deriving the optimal POVM for discriminating coherent states under the one-sided error requirement.
>
> Let $\{\Pi_0, \Pi_1\}$ be a POVM for discriminating coherent states $|\alpha\rangle$ and $|\beta\rangle$ under the one-sided error requirement, when provided a single copy of $|\alpha\rangle$ and $m-1$ copies of $|\beta\rangle$ (the proof of [SZP+07] assumes $m = 2$). The operator Π_0 corresponds to saying that the states $|\alpha\rangle$ and

$|\beta\rangle$ are the same, while the operator Π_1 corresponds to saying that they are different. These operators thus verify the following conditions:

$$\Pi_0, \Pi_1 \geq \mathbb{0},\ \Pi_0 + \Pi_1 = \mathbb{1}, \tag{5.71}$$

and

$$\forall \alpha \in \mathbb{C}, \mathrm{Tr}\left[\Pi_1 |\alpha\rangle \langle\alpha|^{\otimes m}\right] = 0, \tag{5.72}$$

where the last condition is the one-sided error requirement. Integrating this condition over $\mathbb{C}$ yields

$$0 = \int d^2\alpha \, \mathrm{Tr}\left[\Pi_1 |\alpha\rangle \langle\alpha|^{\otimes m}\right] = \mathrm{Tr}[\Pi_1 \Delta_m], \tag{5.73}$$

where we have defined

$$\Delta_m = \int d^2\alpha \, |\alpha\rangle \langle\alpha|^{\otimes m} \geq 0. \tag{5.74}$$

Note that the condition in (5.73) is equivalent to the one-sided requirement in (5.72) because the operators Π_1 and $|\alpha\rangle \langle\alpha|^{\otimes m}$ are positive.

The operator $\frac{m}{\pi}\Delta_m$ is actually a projector. This result can be obtained by writing the state $|\alpha\rangle$ in the Fock basis and an integration in polar coordinates, where $\alpha = re^{i\theta}$, as follows: writing

$$|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{k=0}^{+\infty} \frac{\alpha^k}{\sqrt{k!}} |k\rangle, \tag{5.75}$$

we obtain

$$\begin{aligned}
\Delta_m &= \int d^2\alpha \ \exp[-m|\alpha|^2] \sum_{\substack{k_j, l_j = 0 \\ \forall j \in [m]}}^{\infty} \frac{\alpha^{\sum_j k_j} (\alpha^*)^{\sum_j l_j}}{\sqrt{k_1! \ldots k_m! l_1! \ldots l_m!}} |k_1 \ldots k_m\rangle \langle l_1 \ldots l_m| \\
&= \sum_{k_j, l_j = 0}^{\infty} \frac{|k_1 \ldots k_m\rangle \langle l_1 \ldots l_m|}{\sqrt{k_1! \ldots k_m! l_1! \ldots l_m!}} \int_{r=0}^{\infty} dr \ \exp[-mr^2] r^{1 + \sum_j k_j + l_j} \int_{\theta=0}^{2\pi} d\theta \ \exp[i\theta \sum_j (k_j - l_j)] \\
&= \frac{\pi}{m} \sum_{k_j, l_j = 0}^{\infty} \frac{\delta_{\sum_j k_j, \sum_j l_j}}{m^{\frac{\sum_j k_j}{2}} m^{\frac{\sum_j l_j}{2}}} \sqrt{\frac{(\sum_j k_j)! (\sum_j l_j)!}{k_1! \ldots k_m! l_1! \ldots l_m!}} |k_1 \ldots k_m\rangle \langle l_1 \ldots l_m| \\
&= \frac{\pi}{m} \sum_{p=0}^{\infty} \sum_{\substack{\sum_j k_j = p \\ \sum_j l_j = p}} m^{-p} \sqrt{\frac{p!}{k_1! \ldots k_m!}} \sqrt{\frac{p!}{l_1! \ldots l_m!}} |k_1 \ldots k_m\rangle \langle l_1 \ldots l_m| \\
&= \frac{\pi}{m} \sum_{p=0}^{\infty} |\chi_p^m\rangle \langle \chi_p^m|,
\end{aligned} \tag{5.76}$$

where we have defined for all $p \geq 0$,

$$|\chi_p^m\rangle = m^{-p/2} \sum_{\sum_j k_j = p} \sqrt{\frac{p!}{k_1! \ldots k_m!}} |k_1 \ldots k_m\rangle. \tag{5.77}$$

With the multinomial formula, we obtain $\langle \chi_p^m | \chi_p^m \rangle = 1$ for all $p \geq 0$, and since the states $|\chi_p^m\rangle$ have exactly p photons, we have $\langle \chi_p^m | \chi_q^m \rangle = \delta_{pq}$ for all $p, q \geq 0$. The states $|\chi_p^m\rangle$ thus are orthonormal and with Eq. (5.76), the operator $\frac{m}{\pi}\Delta_m$ is a projector.

By Eq. (5.73), the supports of Π_1 and $\frac{m}{\pi}\Delta_m$ are disjoint, and by Eq. (5.72) we have $\Pi_0 + \Pi_1 = \mathbb{1}$, so the support of $\frac{m}{\pi}\Delta_m$ is included in the support of Π_0. The optimal POVM $\{\Pi_0^{opt}, \Pi_1^{opt}\}$ for state discrimination minimises the error probability, hence with the one-sided error requirement Π_0^{opt} must have minimal support, meaning that

$$\Pi_0^{opt} = \frac{m}{\pi}\Delta_m = \sum_{p=0}^{+\infty} |\chi_p^m\rangle \langle \chi_p^m| \quad \text{and} \quad \Pi_1^{opt} = \mathbb{1} - \Pi_0^{opt}. \tag{5.78}$$

Note that, with the same proof, this choice of POVM is also optimal in the generalised setting where one is given one unknown generic state and $m-1$ unknown coherent states, and is asked to test if all the states are identical or not.

We now show that the POVM $\{\Pi_0^h, \Pi_1^h\}$ corresponding to the Hadamard interferometer with a threshold detection of the last $m-1$ modes is optimal for coherent state discrimination under the one-sided error requirement, i.e., that

$$\Pi_0^h = \Pi_0^{opt}, \tag{5.79}$$

where Π_0^{opt} is defined in Eq. (5.78). We have

$$\Pi_0^h = \hat{H}_n^\dagger \Pi_0^d \hat{H}_n, \tag{5.80}$$

where $\hat{H}_n$ is the unitary evolution corresponding to the action of the interferometer of order m defined in Eq. (5.59), with $n = \log m$, and $\Pi_0^d = \mathbb{1} \otimes |0\rangle\langle 0|^{\otimes m-1}$ is the POVM operator corresponding to the event where none of the $m-1$ threshold detectors clicks. We obtain

$$\begin{aligned} \Pi_0^h &= \hat{H}_n^\dagger \left(\mathbb{1} \otimes |0\rangle\langle 0|^{\otimes m-1}\right) \hat{H}_n \\ &= \sum_{p=0}^{+\infty} \tilde{H}_n^\dagger \left(|p\rangle\langle p| \otimes |0\rangle\langle 0|^{\otimes m-1}\right) \hat{H}_n. \end{aligned} \tag{5.81}$$

For $k = 1 \dots m$, we write $a_k^\dagger$ the creation operator for the k^{th} mode. For all $p \geq 0$ we have

$$\begin{aligned}
\hat{H}_n^\dagger \left(|p\rangle \otimes |0\rangle^{\otimes m-1} \right) &= \frac{1}{\sqrt{p!}} \hat{H}_n^\dagger (\hat{a}_1^\dagger)^p |0\rangle^{\otimes m} \\
&= \frac{1}{\sqrt{p!}} (\hat{H}_n^\dagger \hat{a}_1^\dagger \hat{H}_n)^p |0\rangle^{\otimes m} \\
&= \frac{m^{-p/2}}{\sqrt{p!}} (\hat{a}_1^\dagger + \dots + \hat{a}_m^\dagger)^p |0\rangle^{\otimes m} \\
&= \frac{m^{-p/2}}{\sqrt{p!}} \sum_{k_1 + \dots + k_m = p} \frac{p!}{k_1! \dots k_m!} (\hat{a}_1^\dagger)^{k_1} \dots (\hat{a}_m^\dagger)^{k_m} |0\rangle^{\otimes m} \\
&= m^{-p/2} \sum_{k_1 + \dots + k_m = p} \sqrt{\frac{p!}{k_1! \dots k_m!}} |k_1 \dots k_m\rangle \\
&= |\chi_p^m\rangle ,
\end{aligned} \tag{5.82}$$

where we have used $\hat{H}_n |0\rangle^{\otimes m} = |0\rangle^{\otimes m}$, $\hat{H}_n^\dagger \hat{H}_n = \mathbb{1}$, $\hat{H}_n^\dagger \hat{a}_1^\dagger \hat{H}_n = \frac{\hat{a}_1^\dagger + \dots + \hat{a}_m^\dagger}{\sqrt{m}}$, the multinomial formula, and Eq. (5.77). With Eqs. (5.78) and (5.81), this concludes the proof.

Given that the statistics obtained with the merger scheme and the looped merger scheme mimic those of the Hadamard scheme, these schemes are also optimal for the same discrimination task.

■

While these device are relatively easy to implement, any implementation will suffer from experimental imperfections. In the next section, we investigate how such imperfections affect the performance of the merger scheme, for $m = 4$ modes.

5.3.4 Experimental imperfections

In this section, we analyse the performance of the merger scheme in presence of experimental imperfections. Our error model is the following, with three major sources of error: (i) the limited detector efficiency and channel transmission loss, characterized by a parameter $0 \leq \eta \leq 1$, which changes the coherent state $|\alpha\rangle$ to $|\sqrt{\eta}\alpha\rangle$ thus reducing the probability of obtaining a click using a single-photon threshold detector by a factor η; (ii) the limited beam-splitter visibility $0 \leq \nu \leq 1$, which may lead to a click in the wrong detector, and (iii) the dark count in the detectors characterized by a probability p_{dark}. For our analysis, the click probability due to the coherent states is of $O(1)$ and thus significantly larger than the dark count probability $p_{dark} \sim 10^{-8}$. The dark counts can thus be safely ignored.

For $m = 2$, when the input $|\alpha\rangle$, $|\beta\rangle$ is fed in an imperfect beam splitter, the transformation from

	η	ν	p_{dark}
Exp.	0.9	$(98.8 \pm 0.3)\%$	$(1 \pm 0.1) * 10^{-8}$

Table 5.1: Table illustrating the experimental parameters used in simulation of our results. The dark-count rate, p_{dark}, achievable with super-conducting detectors [SJZ+18]. The standard values are set-up efficiency, η, and beam splitter visibility ν are from [KKD19].

input modes $\{\hat{a}^\dagger, \hat{b}^\dagger\}$ into the output modes $\{\hat{c}^\dagger, \hat{d}^\dagger\}$, is the following:

$$|\alpha\rangle_a \otimes |\beta\rangle_b \mapsto \left|\sqrt{\nu}\frac{\alpha+\beta}{\sqrt{2}} + \sqrt{1-\nu}\frac{\alpha-\beta}{\sqrt{2}}\right\rangle_c \otimes \left|\sqrt{\nu}\frac{\alpha-\beta}{\sqrt{2}} + \sqrt{1-\nu}\frac{\alpha+\beta}{\sqrt{2}}\right\rangle_d . \tag{5.83}$$

The corresponding unitary transformation is

$$H' = \frac{1}{\sqrt{2}}\begin{pmatrix} A & B \\ A & -B \end{pmatrix}, \tag{5.84}$$

where $A = \sqrt{\nu} + \sqrt{1-\nu}$, and $B = \sqrt{\nu} - \sqrt{1-\nu}$.

We consider the case of $m = 4$ spatial modes (Fig 5.9), indexed from 1 to 4. We apply the imperfect transformation on the input $|\alpha\beta\beta\beta\rangle$. This results in

$$|\alpha\beta\beta\beta\rangle \mapsto U_4' |\alpha\beta\beta\beta\rangle = |\delta_1\delta_2\delta_3\delta_4\rangle , \tag{5.85}$$

where from Eq. (5.68) we derive

$$\begin{aligned} U_2' &= H_{1,3}' \times (H_{1,2}' \oplus H_{3,4}') \\ &= \begin{pmatrix} \frac{1}{2}A^2 & \frac{1}{2}AB & \frac{1}{2}AB & \frac{1}{2}B^2 \\ \frac{1}{\sqrt{2}}A & -\frac{1}{\sqrt{2}}B & 0 & 0 \\ \frac{1}{2}A^2 & \frac{1}{2}AB & -\frac{1}{2}AB & -\frac{1}{2}B^2 \\ 0 & 0 & \frac{1}{\sqrt{2}}A & -\frac{1}{\sqrt{2}}B \end{pmatrix}, \end{aligned} \tag{5.86}$$

with $A = \sqrt{\nu} + \sqrt{1-\nu}$ and $B = \sqrt{\nu} - \sqrt{1-\nu}$. We obtain $\delta_2 = \frac{A\alpha - B\beta}{\sqrt{2}}$, and $\delta_3 = \frac{A^2\alpha - B^2\beta}{2}$. Adding the channel and detector losses η, the output is mapped as $\delta_k \mapsto \sqrt{\eta}\delta_k$, for all k.

Similar to the analysis without experimental imperfection, we detect the output modes 1 and 2 of the imperfect merger interferometer, with the coherent state input being $|\alpha\beta\beta\beta\rangle$. The probability that none of the two detectors clicks is given by

$$\exp\left(-\eta(|\delta_2|^2 + |\delta_3|^2)\right). \tag{5.87}$$

Assigning to the detection event *no detector clicks* the value 0, and to other detection events, i.e., *at least one of the detectors clicks*, the value 1, we obtain a device whose statistics approximate those of a projective measurement.

When the states are the same, the completeness, which is the probability of not obtaining the detection event 1 is

$$c_4^{exp} = \exp(-2\eta(1-\nu)(1+2\nu)|\alpha|^2). \tag{5.88}$$

We observe that if $v = 1$ (no imperfections), then $c_4^{exp} = 1$, thus we obtain perfect completeness. For the imperfection values of Table 5.1, the value of c_4^{exp} is close to 1 for small $|\alpha|^2$ values.

The analogous completeness for $m = 2$ is

$$c_2^{exp} = \exp(-2\eta(1-v)|\alpha|^2). \tag{5.89}$$

From Eq. (5.89) and Eq. (5.88), we observe that $c_2^{exp} \leq c_4^{exp}$, which implies that the completeness for the $m = 4$ scheme is less than the completeness for the $m = 2$ scheme. The reduction in completeness probability for the $m = 4$ scheme is precisely what accounts for a lower failure probability when the local and reference states are different, which we detail in the next paragraph.

If the states are different, the probability of obtaining the detection event 1 (soundness) is given by:

Lemma 5.3.

$$\begin{aligned} s_4^{exp} = 1 - \exp\Big[&(4v^2-1)|\alpha-\beta|^2 + 4\left[(1+2v)(1-v) + 2\sqrt{v(1-v)}\right]|\alpha|^2 \\ &+ 4\left[(1+2v)(1-v) - 2\sqrt{v(1-v)}\right]|\beta|^2\Big)\Big]. \end{aligned} \tag{5.90}$$

Proof. When the states are different, the probability of obtaining the detection event 0 (failure probability) is

$$1 - s_4 = \exp\left(-\frac{\eta}{4}A\right), \tag{5.91}$$

where

$$\begin{aligned} A &= 2\left|\sqrt{v}(\alpha-\beta) + \sqrt{1-v}(\alpha+\beta)\right|^2 + \left|\alpha-\beta+2\sqrt{v(1-v)}(\alpha+\beta)\right|^2 \\ &= (1+2v)|\alpha-\beta|^2 + 2(1+2v)(1-v)|\alpha+\beta|^2 + 8\sqrt{v(1-v)}(|\alpha|^2 - |\beta|^2), \end{aligned} \tag{5.92}$$

where we used $(\alpha-\beta)(\alpha+\beta)^* + (\alpha-\beta)^*(\alpha+\beta) = 2|\alpha|^2 - 2|\beta|^2$. Using $|\alpha+\beta|^2 = 2|\alpha|^2 + 2|\beta|^2 - |\alpha-\beta|^2$ we obtain

$$A = (4v^2-1)|\alpha-\beta|^2 + 4\left[(1+2v)(1-v) + 2\sqrt{v(1-v)}\right]|\alpha|^2 + 4\left[(1+2v)(1-v) - 2\sqrt{v(1-v)}\right]|\beta|^2. \tag{5.93}$$

■

The analogous soundness in $m = 2$ experimental imperfection scheme is

$$s_2^{exp} = 1 - \exp\Big[-\eta\Big(v - \frac{1}{2}\Big)|\alpha-\beta|^2 - \eta\Big(1 - v + \sqrt{v(1-v)}\Big)|\alpha|^2 - \eta\Big(1 - v - \sqrt{v(1-v)}\Big)|\beta|^2\Big]. \tag{5.94}$$

We then obtain:

Lemma 5.4. *For all experimental parameters,*

$$s_2^{exp} \leq s_4^{exp}. \tag{5.95}$$

Proof. We have

$$\begin{aligned} s_2^{exp} &= 1-\exp\Big[-\eta\Big(\nu-\frac{1}{2}\Big)|\alpha-\beta|^2-\eta\Big(1-\nu+\sqrt{\nu(1-\nu)}\Big)|\alpha|^2 \\ &\qquad -\eta\Big(1-\nu-\sqrt{\nu(1-\nu)}\Big)|\beta|^2\Big] \\ &\equiv 1-\exp[-\eta A_2], \end{aligned} \tag{5.96}$$

and

$$\begin{aligned} s_4^{exp} &= 1-\exp\Big[-\eta\Big(\nu^2-\frac{1}{4}\Big)|\alpha-\beta|^2 \\ &\qquad -\eta\Big((1+2\nu)(1-\nu)+2\sqrt{\nu(1-\nu)}\Big)|\alpha|^2 \\ &\qquad -\eta\Big((1+2\nu)(1-\nu)-2\sqrt{\nu(1-\nu)}\Big)|\beta|^2\Big] \\ &\equiv 1-\exp[-\eta A_4]. \end{aligned} \tag{5.97}$$

Since the function $x \mapsto 1-e^{-x}$ is increasing, it is sufficient to show that $A_2 \le A_4$ for all α, β. Writing $\alpha = re^{i\phi}$ and $\beta = te^{i\psi}$, where $r,t \ge 0$ and $\phi,\psi \in [0,2\pi]$, we obtain

$$\begin{aligned} A_4 - A_2 &= \Big(\frac{1}{4}+\nu(1-\nu)+\sqrt{\nu(1-\nu)}\Big)r^2 \\ &\qquad +\Big(\frac{1}{4}+\nu(1-\nu)-\sqrt{\nu(1-\nu)}\Big)t^2 \\ &\qquad -2rt\Big(\frac{1}{4}-\nu(1-\nu)\Big)\cos(\phi-\psi). \end{aligned} \tag{5.98}$$

This last expression is a polynomial of degree 2 in r, with a positive leading coefficient. Thus if its discriminant is negative, then the expression is always positive. The discriminant is

$$\begin{aligned} \Delta &= 4t^2\Big[\Big(\frac{1}{4}-\nu(1-\nu)\Big)^2\cos(\phi-\psi)^2-\Big(\frac{1}{4}+\nu(1-\nu)+\sqrt{\nu(1-\nu)}\Big)\Big(\frac{1}{4}+\nu(1-\nu)-\sqrt{\nu(1-\nu)}\Big)\Big] \\ &\le -6t^2\nu(1-\nu) \\ &\le 0, \end{aligned} \tag{5.99}$$

where the second line is obtained by using $\cos(\phi-\psi) \le 1$. Hence for all experimental parameters within the error model we consider, we have $s_2^{exp} \le s_4^{exp}$.

■

Hence, the experimental $m = 4$ scheme outperforms the $m = 2$ scheme in soundness for all values of the noise parameters. On the other hand, the completeness of the scheme suffers from experimental imperfections.

5.4 Discussion and open problems

We have identified a connection between unknown quantum state discrimination and quantum-programmable measurements. We have presented an optimal scheme for a programmable projective measurement device, and a linear optical implementation, with the Hadamard interferometer and single-photon encoding, which is straightforward and efficient. This could for example be used to design a photonic circuit which would act as a universal projective measurement device for a broad range of potential applications from quantum information and cryptography to tests of contextuality.

Our scheme can also be interpreted as an optimal swap test when one has a single copy of one state, and $m-1$ of the other. We have chosen to phrase the problem in terms of $m-1$ copies of the state $|\psi\rangle$. In principle we could have chosen any other encoding of the quantum input into $m-1$ registers. The reason for this choice is twofold. Firstly it is part of the envisaged problem setting—we imagine a device producing states encoding our measurement, for example these could be the output of a computation. Secondly we do so in order to separate as much as possible the resource of $m-1$ program systems and the process of translating them into a measurement. In particular if one had any other encoding, for example into some entangled states, this encoding process could be incorporated into the circuit representing the generic measurement apparatus. In this sense the most quantum information that can be contained about the state $|\psi\rangle$ in $m-1$ systems is $m-1$ copies of the state $|\psi\rangle$—anything more can be done afterwards. This result also provides a natural interpretation of the notion of projective measurement in quantum mechanics, as a comparison between one state and several copies of another state using an interferometer: in the macroscopic limit, when many copies of a reference eigenstate are available, we retrieve a macroscopic classically programmable quantum measurement set up.

In order to reduce the experimental requirements, we have also presented an optimal programmable measurement scheme that projects the incoming single mode state in the test register into a local coherent state basis of the program registers. Our scheme is implemented using balanced beam splitters and single-photon threshold detectors. Threshold detectors with high efficiency and ultra low dark counts are commercially available [SJZ+18]. Additionally, the numbers of detectors needed is logarithmic in the size of the interferometer, which itself is composed only of a linear number of balanced beam splitters. This implementation using coherent states can act as a backbone in improving the performance of a range of quantum protocols in communication complexity [BCWDW01, dB04], cryptography and computational regimes [ABD+08, MKB05, WRD+06, HM13, EAO+02, LMR13].

For completeness, it would be interesting to characterise the full class of interferometers that are optimal for state identity testing under the one-sided error requirement, as we only gave a broad class of such interferometers using a group construction. It would be also interesting to consider the influence of real experimental conditions in a more general setting.

CHAPTER 6

QUANTUM WEAK COIN FLIPPING WITH LINEAR OPTICS

Weak coin flipping is among the fundamental cryptographic primitives which ensure the security of modern communication networks. It allows two mistrustful parties to remotely agree on a random bit when they favor opposite outcomes. Unlike other two-party computations, one can achieve information-theoretic security using quantum mechanics only: both parties are prevented from biasing the flip with probability higher than $1/2+\epsilon$, where ϵ is arbitrarily low. Classically, the dishonest party can always cheat with probability 1 unless computational assumptions are used. Despite its importance, no physical implementation has been proposed so far for quantum weak coin flipping.

In this chapter, we present a practical protocol for quantum weak coin flipping that requires a single photon and linear optics only. We show that it is secure even when threshold single-photon detectors are used, and reaches a bias as low as $\epsilon = 1/\sqrt{2} - 1/2 \approx 0.207$. We further show that the protocol may display quantum advantage over a few hundred meters with state-of-the-art technology.

This chapter is based on [BCKD20].

6.1 Weak coin flipping protocol with linear optics

Compared to weak coin flipping, where two mistrustful parties wish to remotely agree on the outcome of a coin flip when they favor different outcomes, the cryptographic task of strong coin flipping corresponds to the case where they want to agree on an unbiaised random bit when they do not necessarily favor a particular outcome. Despite its name, strong coin flipping is less general than weak coin flipping in the sense that optimal strong coin flipping protocols may be designed which use weak coin flipping protocols as a subroutine [CK09].

While quantum strong coin flipping protocols have been experimentally demonstrated [MTVUZ05, BBB$^+$11, PJL$^+$14], no implementation has been proposed for quantum weak coin flipping. This may be explained by two reasons. First, it is difficult to find an encoding and implementation which is robust to losses: a dishonest party may always declare an abort when they are not satisfied with the flip's outcome. Second, none of the proposed quantum weak coin flipping protocols [SR02, KN04, Moc04, Moc05, Moc07, ACG$^+$16, ARW19, ARV19] translate trivially into a simple experiment: they all involve performing single-shot generalized measurements or generating beyond-qubit states.

We introduce a family of quantum weak coin flipping protocols, inspired by [SR02], which achieve biases as low as $\epsilon = 1/\sqrt{2} - 1/2 \approx 0.207$. Our protocols involve simple projective measurements instead of generalized ones, require a single photon and linear optics only, and need at most three rounds of communication between the parties. The information is encoded by mixing a single photon with vacuum on an unbalanced beam splitter, which generates entanglement [MBH$^+$13]: both parties may then agree on a random bit, while the entanglement is simultaneously verified. This encoding is very robust to noise, as the single photon need not be pure or indistinguishable from other photons in any degree of freedom, save photon number. We also use a version of our schemes to construct a quantum strong coin flipping protocol with bias ≈ 0.31. We further derive a practical security proof for both number-resolving and threshold single-photon detectors, considering the extension to infinite-dimensional Hilbert spaces. Since the presence of losses may enable classical protocols to reach lower cheating probabilities than quantum protocols, we finally show that our quantum protocol bears no classical equivalent over a few hundred meters of lossy optical fiber and non-unit detection efficiency.

In the honest protocol, Alice and Bob wish to toss a fair coin, with a priori knowledge that they each favor opposite outcomes. Fig. 6.1 represents the implementation of the honest protocol, which follows five distinct steps. Defining $x \in [0, \frac{1}{2}]$ as a free protocol parameter, these read:

- Alice mixes a single photon with the vacuum on a beam splitter of reflectance x.
- Alice keeps the first half of the state obtained, and sends the second half to Bob.
- Bob mixes the half he receives with the vacuum on a beam splitter of reflectance $y = 1 - \frac{1}{2(1-x)}$.
- Bob measures the second register of his state with a single-photon detector, and broadcasts the outcome $c \in \{0, 1\}$.
- The last step is a verification step, which splits into two cases. If $c = 0$, Alice sends her half of the state to Bob, who mixes it with his half on a beam splitter of reflectance $z = 2x$. He then measures the two output modes with single-photon detectors. He declares Alice the winner if the outcome $(1, 0)$ is obtained. If $c = 1$: Bob discards his half, and Alice measures her half with a single-photon detector. If the outcome is (0), Bob is declared winner.

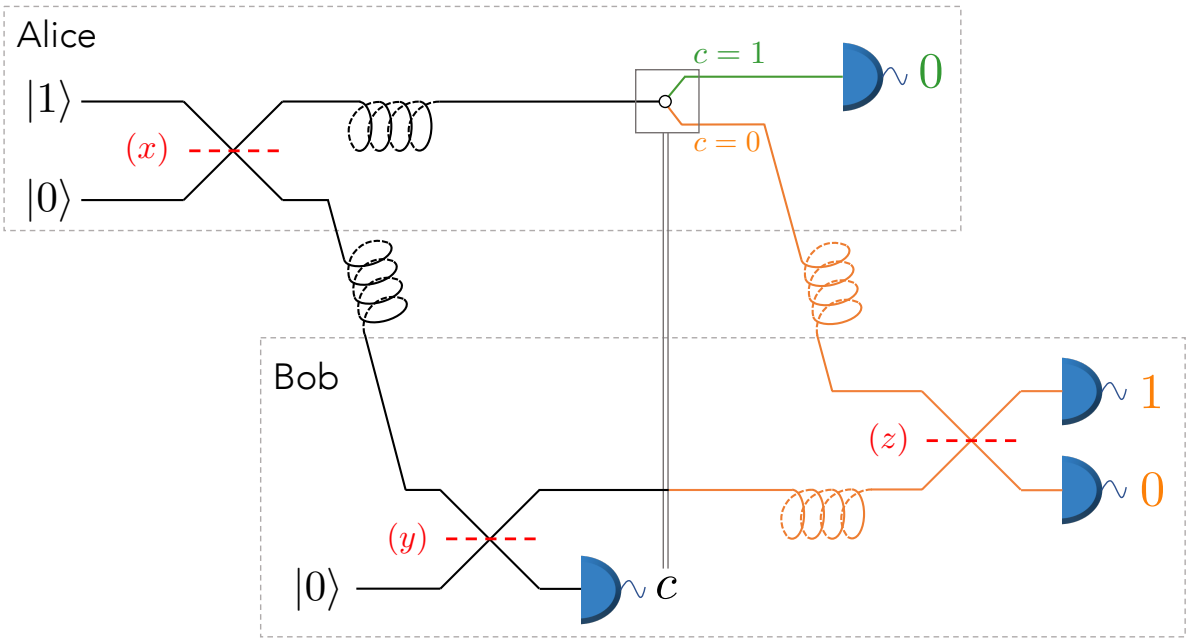

Figure 6.1: **Representation of the honest protocol.** The dashed boxes indicate Alice and Bob's laboratories, respectively. Dashed red lines represent beam splitters, with the reflectance indicated in red. $|0\rangle$ and $|1\rangle$ are the vacuum and single photon Fock states, respectively. Curly lines represent fiber used for quantum communication from Alice to Bob, or delay lines within Alice's or Bob's laboratory, when waiting for the other party's communication. Bob broadcasts the classical outcome c, which controls an optical switch on Alice's side. The protocol when Bob declares $c = 0/1$ is represented in orange/green. The final outcomes are the expected outcomes when both parties are honest.

6.1.1 Completeness

In what follows, we let the parameters x, y, z vary freely, and derive the relations these parameters need to satisfy to enforce a honest protocol without abort cases. We show that for the specific relations indicated above the protocol is fair, i.e., the probability of winning for each party is $\frac{1}{2}$ when they are both honest.

Single photons are quantized excitations of the electromagnetic field, which are described by the action of the creation operator onto the vacuum. Beam splitters act linearly on creation operators and leave invariant the vacuum. More precisely, a beam splitter of reflectance r acting on modes k, l maps the creation operators $\hat{a}_k^\dagger$ and $\hat{a}_l^\dagger$ of the input modes onto $\hat{b}_k^\dagger$ and $\hat{b}_l^\dagger$, where

$$\begin{pmatrix} \hat{b}_k^\dagger \\ \hat{b}_l^\dagger \end{pmatrix} = H_{kl}^{(r)} \begin{pmatrix} \hat{a}_k^\dagger \\ \hat{a}_l^\dagger \end{pmatrix}, \tag{6.1}$$

with

$$H_{kl}^{(r)} = \begin{pmatrix} \sqrt{r} & \sqrt{1-r} \\ \sqrt{1-r} & -\sqrt{r} \end{pmatrix}. \tag{6.2}$$

Hence, the evolution of the quantum state over the three modes up to Bob's first measurement reads:

$$\begin{aligned} |100\rangle &\underset{(x),12}{\rightarrow} \sqrt{x}\,|100\rangle + \sqrt{1-x}\,|010\rangle \\ &\underset{(y),23}{\rightarrow} \sqrt{x}\,|100\rangle + \sqrt{(1-x)y}\,|010\rangle \\ &\quad + \sqrt{(1-x)(1-y)}\,|001\rangle\,, \end{aligned} \tag{6.3}$$

where the notation $(r),kl$ indicates the reflectance of the beam splitter and the corresponding spatial modes. The probability that Bob obtains outcome $c=1$ when measuring the third register thus is $\Pr[1]=(1-x)(1-y)$, while the probability of outcome $c=0$ is $\Pr[0]=1-\Pr[1]$. Setting $y=1-\frac{1}{2(1-x)}$ ensures $\Pr[0]=\Pr[1]=\frac{1}{2}$.

When $c=1$, the state on modes 1 and 2 is projected onto $|00\rangle$, while $c=0$ projects the state onto $\sqrt{2x}\,|10\rangle+\sqrt{1-2x}\,|01\rangle$. In the first case, the measurement performed by Alice outputs (0) with probability 1. In the second case, the measurement performed by Bob outputs $(1,0)$ with probability 1 when

$$z=\frac{x}{1-(1-x)(1-y)}=2x. \tag{6.4}$$

In that case, the probability that Alice (resp. Bob) wins is directly given by $P_h^{(A)}=\Pr[0]$ (resp. $P_h^{(B)}=\Pr[1]$). This shows that the protocol is fair, since $\Pr[0]=\Pr[1]=\frac{1}{2}$.

In the following, we make use of a simple reduction which allows us to simplify calculations in the proofs:

Lemma 6.1. *Let $U=(H^{(z)}\otimes\mathbb{1})(\mathbb{1}\otimes H^{(y)})$, with $z>0$. For all density matrices τ,*

$$\mathrm{Tr}[(\tau\otimes|0\rangle\langle 0|)U^\dagger(\mathbb{1}\otimes|00\rangle\langle 00|)U]=\mathrm{Tr}[(\tau\otimes|0\rangle\langle 0|)V^\dagger(|0\rangle\langle 0|\otimes\mathbb{1}\otimes|0\rangle\langle 0|)V], \tag{6.5}$$

where $V=(\mathbb{1}\otimes H^{(b)})(H^{(a)}\otimes\mathbb{1})(\mathbb{1}\otimes R(\pi)\otimes\mathbb{1})$, with $a=\frac{y(1-z)}{1-(1-y)(1-z)}$ and $b=1-(1-y)(1-z)$, and $R(\pi)$ a phase shift of π acting on mode 2.

Proof. The action of U on the creation operators is given by

$$\begin{aligned} U &= \begin{pmatrix} \sqrt{z} & \sqrt{1-z} & 0 \\ \sqrt{1-z} & -\sqrt{z} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{y} & \sqrt{1-y} \\ 0 & \sqrt{1-y} & -\sqrt{y} \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{z} & \sqrt{y(1-z)} & \sqrt{(1-y)(1-z)} \\ \sqrt{1-z} & -\sqrt{yz} & -\sqrt{(1-y)z} \\ 0 & \sqrt{1-y} & -\sqrt{y} \end{pmatrix}. \end{aligned} \tag{6.6}$$

Linear interferometers map product coherent states onto product coherent states, and, for

all $\alpha \in \mathbb{C}$, we have that $U^\dagger |\alpha 00\rangle = |\beta_1\beta_2\beta_3\rangle$, where

$$\begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\end{pmatrix} = \begin{pmatrix}\alpha\sqrt{z}\\ \alpha\sqrt{y(1-z)}\\ \alpha\sqrt{(1-y)(1-z)}\end{pmatrix}. \tag{6.7}$$

We have $V = (\mathbb{1}\otimes H^{(b)})(H^{(a)}\otimes\mathbb{1})(\mathbb{1}\otimes R(\pi)\otimes\mathbb{1})$, with $a,b\in[0,1]$, and $R(\pi)$ a phase shift of π acting on mode 2. The action of V on the creation operators is given by

$$\begin{aligned} V &= \begin{pmatrix}1 & 0 & 0\\ 0 & \sqrt{b} & \sqrt{1-b}\\ 0 & \sqrt{1-b} & -\sqrt{b}\end{pmatrix}\begin{pmatrix}\sqrt{a} & \sqrt{1-a} & 0\\ \sqrt{1-a} & -\sqrt{a} & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 1\end{pmatrix}\\ &= \begin{pmatrix}\sqrt{a} & -\sqrt{1-a} & 0\\ \sqrt{b(1-a)} & \sqrt{ab} & \sqrt{1-b}\\ \sqrt{(1-a)(1-b)} & \sqrt{a(1-b)} & -\sqrt{b}\end{pmatrix}. \end{aligned} \tag{6.8}$$

For all $\alpha\in\mathbb{C}$, $V^\dagger|0\alpha 0\rangle = |\gamma_1\gamma_2\gamma_3\rangle$, where

$$\begin{pmatrix}\gamma_1\\ \gamma_2\\ \gamma_3\end{pmatrix} = \begin{pmatrix}\alpha\sqrt{b(1-a)}\\ \alpha\sqrt{ab}\\ \alpha\sqrt{1-b}\end{pmatrix}. \tag{6.9}$$

Since $a = \frac{y(1-z)}{1-(1-y)(1-z)}$ and $b = 1-(1-y)(1-z)$, we have $b(1-a)=z$, $ab = y(1-z)$, and $1-b = (1-y)(1-z)$, so $(\beta_1,\beta_2,\beta_3) = (\gamma_1,\gamma_2,\gamma_3)$.
Then,

$$\begin{aligned} \mathrm{Tr}[(\tau\otimes|0\rangle\langle 0|)U^\dagger(\mathbb{1}\otimes|00\rangle\langle 00|)U] &= \frac{1}{\pi}\int_{\mathbb{C}} d^2\alpha\,\mathrm{Tr}[(\tau\otimes|0\rangle\langle 0|)U^\dagger|\alpha 00\rangle\langle\alpha 00|U]\\ &= \frac{1}{\pi}\int_{\mathbb{C}} d^2\alpha\,\mathrm{Tr}[(\tau\otimes|0\rangle\langle 0|)V^\dagger|0\alpha 0\rangle\langle 0\alpha 0|V]\\ &= \mathrm{Tr}[(\tau\otimes|0\rangle\langle 0|)V^\dagger(|0\rangle\langle 0|\otimes\mathbb{1}\otimes|0\rangle\langle 0|)V], \end{aligned} \tag{6.10}$$

where we used the completeness relation of coherent states $\mathbb{1} = \frac{1}{\pi}\int_{\mathbb{C}}|\alpha\rangle\langle\alpha| d^2\alpha$.

■

6.1.2 Soundness

We now derive the soundness of the protocol. Namely, we obtain the maximal winning probabilities when Bob is dishonest and Alice is honest, and vice versa.

Lemma 6.2. *Bob's optimal cheating probability is given by*

$$P_d^{(B)} = 1-x. \tag{6.11}$$

Proof. Dishonest Bob should always declare the outcome $c = 1$ in order to maximize his winning probability. The outcome of the coin flip is then confirmed if Alice obtains the outcome 0 upon verification. Bob thus needs to maximize the probability of the outcome 0, applying a general quantum operation to his half of the state. However, the probability that the detector clicks is independent of Bob's action. It is given by x, so that Bob's winning probability is upper bounded by $(1-x)$. This upper bound is reached if Bob discards his half of the state and broadcasts $c = 1$. Bob's optimal cheating probability thus is $P_d^{(B)} = 1-x$.

■

Alice wins when Bob declares $c = 0$ and the outcome of his quantum measurement is $(1,0)$. The most general strategy of dishonest Alice is to send a (mixed) state σ, while Bob performs the rest of the protocol honestly. Assuming honest Bob has number-resolving detectors, we obtain the following result:

Lemma 6.3. *Alice's optimal cheating probability when Bob has number resolving detectors is given by*

$$P_d^{(A)} = 1-(1-y)(1-z). \tag{6.12}$$

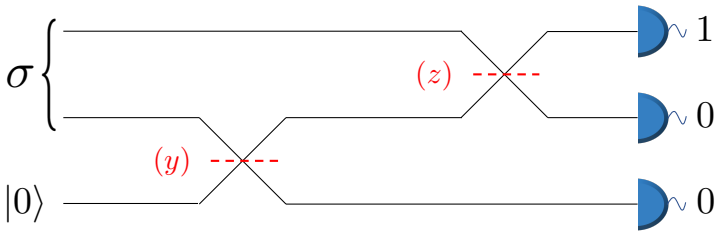

Figure 6.2: **Dishonest Alice.** Alice aims to maximize the outcome $(1,0,0)$: an outcome 0 on the third mode means that Bob declared Alice the winner, while an outcome $(1,0)$ for modes 1 and 2 means that Alice passed Bob's verification. The reflectances of the beam splitter are given by $y = 1-\frac{1}{2(1-x)}$ and $z = 2x$.

Proof. When using number-resolving single-photon detectors, any projection onto the $n > 1$ photon subspace leads to Alice getting caught cheating. Alice must therefore maximize the overlap with the projective measurement $|100\rangle\langle 100|$ only (Fig. 6.2).

Let σ be the state sent by Alice. Let $U = (H^{(z)} \otimes \mathbb{1})(\mathbb{1} \otimes H^{(y)})$, with $z = \frac{x}{1-(1-x)(1-y)}$. Alice needs to maximize the probability of the overall outcome $(1,0,0)$, which is given by

$$P_d^{(A)} = \mathrm{Tr}[U(\sigma \otimes |0\rangle\langle 0|)U^\dagger |100\rangle\langle 100|], \tag{6.13}$$

since Bob uses number-resolving detectors. By convexity of the probabilities, we may assume

without loss of generality that Alice sends a pure state $\sigma = |\psi\rangle\langle\psi|$, which allows us to write:

$$\begin{aligned} P_d^{(A)} &= \mathrm{Tr}[U(|\psi\rangle\langle\psi| \otimes |0\rangle\langle 0|)U^\dagger |100\rangle\langle 100|] \\ &= \mathrm{Tr}[(|\psi\rangle\langle\psi| \otimes |0\rangle\langle 0|)U^\dagger |100\rangle\langle 100|U] \\ &= \mathrm{Tr}[\langle\psi| \otimes \langle 0|U^\dagger |100\rangle\langle 100|U|\psi\rangle \otimes |0\rangle]. \end{aligned} \tag{6.14}$$

We have:

$$\begin{aligned} U^\dagger |100\rangle &= (\mathbb{1} \otimes H^{(y)})(H^{(z)} \otimes \mathbb{1})|100\rangle \\ &= (\mathbb{1} \otimes H^{(y)})(\sqrt{z}\,|100\rangle + \sqrt{1-z}\,|010\rangle) \\ &= \sqrt{z}\,|100\rangle + \sqrt{y(1-z)}\,|010\rangle + \sqrt{(1-y)(1-z)}\,|001\rangle, \end{aligned} \tag{6.15}$$

and therefore:

$$\begin{aligned} U^\dagger |100\rangle\langle 100|U =& z|100\rangle\langle 100| + y(1-z)|010\rangle\langle 010| + (1-y)(1-z)|001\rangle\langle 001| \\ &+ \sqrt{yz(1-z)}\,(|100\rangle\langle 010| + |010\rangle\langle 100|) \\ &+ \sqrt{z(1-y)(1-z)}\,(|100\rangle\langle 001| + |001\rangle\langle 100|) \\ &+ (1-z)\sqrt{y(1-y)}\,(|010\rangle\langle 001| + |001\rangle\langle 010|). \end{aligned} \tag{6.16}$$

Substituting back into Eq. (6.14) then reduces to:

$$\begin{aligned} P_d^{(A)} &= \langle\psi|\left(z|10\rangle\langle 10| + y(1-z)|01\rangle\langle 01| + \sqrt{yz(1-z)}(|10\rangle\langle 01| + |01\rangle\langle 10|)\right)|\psi\rangle \\ &= \langle\psi|\left(\sqrt{z}\,|10\rangle + \sqrt{y(1-z)}\,|01\rangle\right)\left(\sqrt{z}\,\langle 10| + \sqrt{y(1-z)}\,\langle 01|\right)|\psi\rangle \\ &= \left|\langle\psi|\left(\sqrt{z}\,|10\rangle + \sqrt{y(1-z)}\,|01\rangle\right)\right|^2. \end{aligned} \tag{6.17}$$

Using Cauchy-Schwarz inequality then allows us to upper bound $P_d^{(A)}$ as:

$$P_d^{(A)} \leq \|\psi\|^2 \left\|\left(\sqrt{z}\,|10\rangle + \sqrt{y(1-z)}\,|01\rangle\right)\right\|^2 \leq (1-(1-y)(1-z))\|\psi\|^2, \tag{6.18}$$

which is maximized for $\|\psi\| = 1$. Hence we finally get:

$$P_d^{(A)} \leq 1-(1-y)(1-z). \tag{6.19}$$

In order to find Alice's optimal cheating strategy (i.e., the optimal pure state $|\phi\rangle$ that she must send to achieve this bound), we remark that the unnormalized state $\sqrt{z}\,|10\rangle + \sqrt{y(1-z)}\,|01\rangle$ maximizes the expression in Eq. (6.18). Normalizing this state then provides Alice's optimal strategy, which is to prepare the state

$$|\phi\rangle := \sqrt{\frac{z}{1-(1-y)(1-z)}}\,|10\rangle + \sqrt{\frac{y(1-z)}{1-(1-y)(1-z)}}\,|01\rangle. \tag{6.20}$$

Hence,

$$P_d^{(A)} = 1-(1-y)(1-z). \tag{6.21}$$

In the case of a fair protocol, $y = 1-\frac{1}{2(1-x)}$ and $z = 2x$, so

$$P_d^{(A)} = \frac{1}{2(1-x)}, \tag{6.22}$$

and Alice's optimal strategy is to prepare the state

$$|\phi_x\rangle := 2\sqrt{x(1-x)}\,|10\rangle + (1-2x)|01\rangle\,. \tag{6.23}$$

■

Remarkably, the protocol is still secure even when Bob only uses single photon threshold detectors, which is essential to the practicality of the protocol. Moreover, Alice's optimal cheating probability remains the same:

Lemma 6.4. *Alice's optimal cheating probability when Bob has threshold detectors is given by*

$$P_d^{(A)} = 1-(1-y)(1-z). \tag{6.24}$$

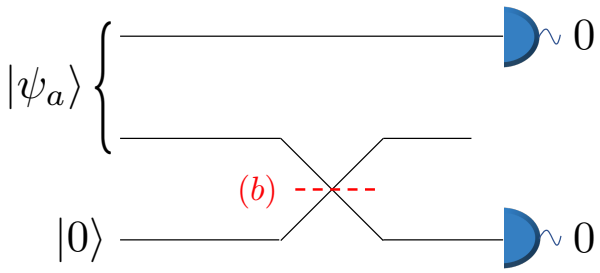

Figure 6.3: **Equivalent picture for dishonest Alice.** In the original dishonest setup of Fig. 6.2, Alice aims to maximize the outcome $(1,0,0)$. This is equivalent to Alice maximizing outcome 0 on spatial modes 1 and 3, independently of what is detected on mode 2. The outcomes indicated correspond to Alice winning. The reflectance is $b = 1-(1-y)(1-z)$.

Proof. Unlike the previous case, incorrect outcomes with higher photon number could still pass the test: for $n \geq 1$, the threshold detectors cannot discriminate between a $|100\rangle$ and $|n00\rangle$ projection. We show in the following that this doesn't help a dishonest Alice, and that the strategy described previously for the case of number resolving detectors is still optimal in the case of threshold detectors.

With the same notations as in the previous proof, Alice needs to maximize the probability of the overall outcome $(1,0,0)$, hence the overlap with the projector $\sum_{n=1}^{\infty} |n00\rangle\langle n00| =$

$(\mathbb{1} - |0\rangle\langle 0|) \otimes |00\rangle\langle 00|$. This allows us to write:

$$P_d^{(A)} = \mathrm{Tr}[U(|\psi\rangle\langle\psi| \otimes |0\rangle\langle 0|)U^\dagger((\mathbb{1} - |0\rangle\langle 0|) \otimes |00\rangle\langle 00|)], \tag{6.25}$$

since Bob uses threshold detectors, where $U = (H^{(z)} \otimes \mathbb{1})(\mathbb{1} \otimes H^{(y)})$, with $z = \frac{x}{1-(1-x)(1-y)}$.

Linear optical evolution conserves photon number. Hence if Alice sends the vacuum state, the detectors will never click. Removing the two-mode vacuum component of the state prepared by Alice and renormalizing therefore always increases her winning probability. Since we are looking for the maximum winning probability, we can assume without loss of generality that $\langle\psi|00\rangle = 0$, i.e.,

$$\mathrm{Tr}[U(|\psi\rangle\langle\psi| \otimes |0\rangle\langle 0|)U^\dagger |000\rangle\langle 000|] = |\langle\psi|00\rangle|^2, \tag{6.26}$$

So maximizing the winning probability in Eq. (6.25) is equivalent to maximizing

$$\tilde{P}_d^{(A)} = \mathrm{Tr}[U(|\psi\rangle\langle\psi| \otimes |0\rangle\langle 0|)U^\dagger(\mathbb{1} \otimes |00\rangle\langle 00|)], \tag{6.27}$$

given the constraint $\langle\psi|00\rangle = 0$. We have

$$\begin{aligned} \tilde{P}_d^{(A)} &= \mathrm{Tr}[U(|\psi\rangle\langle\psi| \otimes |0\rangle\langle 0|)U^\dagger(\mathbb{1} \otimes |00\rangle\langle 00|)] \\ &= \mathrm{Tr}[(|\psi\rangle\langle\psi| \otimes |0\rangle\langle 0|)U^\dagger(\mathbb{1} \otimes |00\rangle\langle 00|)U]. \end{aligned} \tag{6.28}$$

With Lemma 6.1 and Eq. (6.28), we may thus write:

$$\tilde{P}_d^{(A)} = \mathrm{Tr}[(|\psi\rangle\langle\psi| \otimes |0\rangle\langle 0|)V^\dagger(|0\rangle\langle 0| \otimes \mathbb{1} \otimes |0\rangle\langle 0|)V], \tag{6.29}$$

where $V = (\mathbb{1} \otimes H^{(b)})(H^{(a)} \otimes \mathbb{1})(\mathbb{1} \otimes R(\pi) \otimes \mathbb{1})$, with $a = \frac{y(1-z)}{1-(1-y)(1-z)}$ and $b = 1-(1-y)(1-z)$. Let us now define:

$$|\psi_a\rangle := H^{(a)}(\mathbb{1} \otimes R(\pi))|\psi\rangle\,. \tag{6.30}$$

The constraints $\langle\psi|00\rangle = 0$ and $\langle\psi_a|00\rangle = 0$ are equivalent, because the above transformation leaves the total number of photons invariant. With Eq. (6.29) we obtain

$$\tilde{P}_d^{(A)} = \mathrm{Tr}[(|\psi_a\rangle\langle\psi_a| \otimes |0\rangle\langle 0|)(\mathbb{1} \otimes H^{(b)})(|0\rangle\langle 0| \otimes \mathbb{1} \otimes |0\rangle\langle 0|)(\mathbb{1} \otimes H^{(b)})], \tag{6.31}$$

with the constraint $\langle\psi_a|00\rangle = 0$. Maximizing this expression thus corresponds to maximizing the probability of the outcome $(0,0)$ when measuring modes 1 and 3 of the state obtain by mixing the second half of $|\psi_a\rangle$ with the vacuum on a beam splitter of reflectance $b = 1-(1-y)(1-z)$ (Fig. 6.3).

We now show that an optimal strategy for Alice is to ensure that $|\psi_a\rangle = |01\rangle$. Let us write

$$|\psi_a\rangle = \sum_{p+q>0} \psi_{pq}\,|pq\rangle, \tag{6.32}$$

where we take into account the constraint $\langle\psi_x|00\rangle = 0$. Then, with Eq. (6.31) we obtain

$$\begin{aligned}
\tilde{P}_d^{(A)} &= \sum_{p+q>0,p'+q'>0} \psi_{pq}\psi^*_{p'q'} \mathrm{Tr}[|pq0\rangle\langle p'q'0|(|0\rangle\langle 0|\otimes H^{(b)}(\mathbb{1}\otimes|0\rangle\langle 0|)H^{(b)})] \\
&= \sum_{q>0,q'>0} \psi_{0q}\psi^*_{0q'} \mathrm{Tr}[|q0\rangle\langle q'0|H^{(b)}(\mathbb{1}\otimes|0\rangle\langle 0|)H^{(b)}] \\
&= \sum_{n\geq 0,q>0,q'>0} \psi_{0q}\psi^*_{0q'} \mathrm{Tr}[|q0\rangle\langle q'0|H^{(b)}|n0\rangle\langle n0|H^{(b)}] \\
&= \sum_{n>0} |\psi_{0n}|^2 |\langle n0|H^{(b)}|n0\rangle|^2 \\
&= \sum_{n>0} |\psi_{0n}|^2 b^n,
\end{aligned} \tag{6.33}$$

where we used in the fourth line the fact that $H^{(b)}$ doesn't change the number of photons. Since $b \in [0,1]$, this shows that

$$\begin{aligned}
\tilde{P}_d^{(A)} &\leq b \sum_{n>0} |\psi_{0n}|^2 \\
&= b,
\end{aligned} \tag{6.34}$$

since $|\psi_a\rangle$ is normalized, and this bound is reached for $|\psi_{01}|^2 = 1$, i.e., $|\psi_a\rangle = |01\rangle$. With Eq. (6.30), this implies that an optimal strategy for Alice is to prepare the state

$$\begin{aligned}
|\psi\rangle &= (\mathbb{1}\otimes R(\pi))H^{(a)}|01\rangle \\
&= \sqrt{1-a}\,|10\rangle + \sqrt{a}\,|01\rangle \\
&= \sqrt{\frac{z}{1-(1-y)(1-z)}}\,|10\rangle + \sqrt{\frac{y(1-z)}{1-(1-y)(1-z)}}\,|01\rangle \\
&= |\phi\rangle,
\end{aligned} \tag{6.35}$$

where $|\phi\rangle$ is the state that dishonest Alice needs to send to maximize her winning probability when Bob uses number-resolving detectors (Eq. (6.20)). Her winning probability is then

$$P_d^{(A)} = 1-(1-y)(1-z). \tag{6.36}$$

We therefore recover the same result as for number-resolving detectors. Once again, if the protocol is fair then $y = 1-\frac{1}{2(1-x)}$ and $z = 2x$, so

$$P_d^{(A)} = \frac{1}{2(1-x)}, \tag{6.37}$$

and an optimal strategy for Alice is to prepare the state

$$|\phi_x\rangle := 2\sqrt{x(1-x)}\,|10\rangle + (1-2x)|01\rangle. \tag{6.38}$$

■

Alice's cheating probability equals $\frac{1}{2(1-x)}$ for $y = 1-\frac{1}{2(1-x)}$ and $z = 2x$. In particular, for all values

of x, we retrieve the property shared by the protocols of [SR02]: $P_d^{(A)}P_d^{(B)} = \frac{1}{2}$. Setting $x = 1 - 1/\sqrt{2}$, we obtain a version of the protocol which is balanced, i.e., both players have the same cheating probability $1/\sqrt{2}$. The protocol bias is then $\epsilon = 1/\sqrt{2} - 1/2 \approx 0.207$.

6.1.3 Strong coin flipping protocol

Following [CK09], we show that our family of quantum weak coin flipping protocols allows us to construct a quantum strong coin flipping protocol:

Lemma 6.5. *There exists a quantum strong coin flipping protocol achieving bias $\epsilon \approx 0.31$ which uses an unbalanced linear optical weak coin flipping protocol as a subroutine.*

Proof. An unbalanced quantum weak coin flipping protocol can be turned into a quantum strong coin flipping protocol using an additional classical protocol, as described in [CK09]. In particular, let us consider a weak coin flipping protocol such that:

$$\begin{aligned} P_h^{(A)} &= p \\ P_h^{(B)} &= 1-p \\ P_d^{(A)} &= p+\epsilon \\ P_d^{(B)} &= 1-p+\epsilon, \end{aligned} \tag{6.39}$$

for $p \in [0,1]$ and $\epsilon > 0$. Then, the corresponding strong coin flipping protocol has bias [CK09]

$$\max\left(\frac{1}{2} - \frac{1}{2}(p-\epsilon), \frac{1}{2-(p+\epsilon)} - \frac{1}{2}\right). \tag{6.40}$$

For our weak coin flipping protocol, we have:

$$\begin{aligned} P_h^{(A)} &= 1-(1-x)(1-y) \\ P_h^{(B)} &= (1-x)(1-y) \\ P_d^{(A)} &= 1-(1-y)(1-z) \\ P_d^{(B)} &= 1-x, \end{aligned} \tag{6.41}$$

with the constraint $z = \frac{x}{1-(1-x)(1-y)}$ (so that the protocol does not abort in the honest case, Eq. (6.4)). Enforcing the conditions in Eq. (6.39), and optimizing over the corresponding strong coin flipping bias implies

$$\begin{aligned} x &= \frac{y^2}{(1-y)(1-2y)} \\ z &= \frac{y}{(1-y)^2} \\ 1-\frac{x}{2} &= \frac{1}{2-y-z+yz}, \end{aligned} \tag{6.42}$$

which in turn give the values

$$\begin{aligned} x &\approx 0.38 \\ y &\approx 0.31 \\ z &\approx 0.66, \end{aligned} \tag{6.43}$$

by enforcing $x, y, z \in [0,1]$, and a bias of ≈ 0.31, which is a lower bias than the best implemented strong coin flipping protocol so far [PJL+14].

■

6.2 Experimental imperfections

6.2.1 Noisy protocol

We investigate how imperfect state generation, non-ideal beam splitters and single-photon detector dark counts affect the correctness and security of the protocol. While we fixed the parameter values to $y = 1 - \frac{1}{2(1-x)}$ and $z = 2x$ in the ideal setting, we now allow the three parameters x, y, z to vary freely.

The vacuum/single-photon encoding is very robust to noise, in comparison to polarization or phase encoding for instance: the only property that must be preserved through propagation is photon number. This implies that photon indistinguishability and purity are not required in any degree of freedom other than photon number. In this case, Alice may simply produce a heralded single photon via spontaneous parametric down-conversion (SPDC) [Cou18], which generates a photon pair: one may be used for the flip, while the other may herald the presence of the first one. Given the photon-pair emission probability p, accidentally emitting two pairs at the same time using SPDC occurs with probability p^2. Since p may be arbitrarily tuned by changing the pump power, p^2—and therefore the probability of two photons being accidentally generated by Alice at once—may then be decreased to negligible values.

Note that, in the case where Alice's single photon source is probabilistic but heralded (as in SPDC), she may always inform Bob of a successful state generation prior to his announcement of c without compromising security. In what follows, we may therefore assume that both parties have agreed on the presence of an initial state, and hence know when the protocol occurs.

Noise will therefore stem from the non-ideal reflectances of the beam splitters, and the non-zero detector dark count probability p_{dc}. For each party, these may affect the protocol correctness in two ways: an undesired bias of the flip, and an added abort probability during the verification process.

Deviations on the beam splitter reflectances x, y, and z will first change the honest winning probabilities: these may be re-calculated by replacing the ideal reflectance $r \in \{x, y\}$ with an imperfect r'. As regards to honest aborts, a beam splitter with reflectance z' instead of z may be applied on the resulting state when $c = 0$. Noisy detectors may cause an unwanted abort

corresponding to a click because of dark counts. However, with superconducting nanowire single-photon detectors, this probability is typically very low, of the order of $p_{dc} < 10^{-8}$ [Had09].

We can therefore conclude that any source of noise may be incorporated in the security analysis by simply replacing parameters x, y, and z with x', y', and z'. Furthermore, this source of error will most likely be negligible with current technology. We therefore solely focus on the more consequential effects of losses.

6.2.2 Losses: completeness

Losses can be due to the channel transmission and to non-unit delay line transmission and detection efficiencies. We label η_t the transmission efficiency of the quantum channel from Alice to Bob. We also define as $\eta_f^{(i)}$ the transmission of party i's fiber delay, while $\eta_d^{(i)}$ denotes the detection efficiency of party i's single-photon detectors. Here, we assume the efficiencies of Bob's detectors to be the same, and that each party introduces a fiber delay whenever they are waiting for the other party's communication. The delay time therefore depends on the distance between the two parties. We give a representation of the honest protocol with losses, in Fig. 6.4.

We recall a useful simple property, which we will use extensively in the following:

Lemma 6.6. *Equal losses on various modes can be commuted through passive linear optical elements acting on these modes.*

This result was proven, e.g., in [BL10], and we give hereafter a quick proof.

Proof. One way to prove this statement is to use the fact that any interferometer may be decomposed as beam splitters and phase shifters [RZBB94]. Then, losses trivially commute with phase shifters, and are easily shown to commute with beam splitters. Indeed, consider a beam splitter of reflectance t acting on modes 1 and 2. Its action on the creation operators of the modes is given by

$$\hat{a}_1^\dagger, \hat{a}_2^\dagger \to \sqrt{t}\hat{a}_1^\dagger + \sqrt{1-t}\hat{a}_2^\dagger, \sqrt{1-t}\hat{a}_1^\dagger - \sqrt{t}\hat{a}_2^\dagger, \tag{6.44}$$

while equal losses η on both modes act as

$$\hat{a}_1^\dagger, \hat{a}_2^\dagger \to \sqrt{\eta}\hat{a}_1^\dagger, \sqrt{\eta}\hat{a}_2^\dagger. \tag{6.45}$$

Hence, the action of the beam splitter followed by losses is given by

$$\hat{a}_1^\dagger, \hat{a}_2^\dagger \to \sqrt{\eta}(\sqrt{t}\hat{a}_1^\dagger + \sqrt{1-t}\hat{a}_2^\dagger), \sqrt{\eta}(\sqrt{1-t}\hat{a}_1^\dagger - \sqrt{t}\hat{a}_2^\dagger), \tag{6.46}$$

while losses followed by the beam splitter act as

$$\hat{a}_1^\dagger, \hat{a}_2^\dagger \to \sqrt{t}(\sqrt{\eta}\hat{a}_1^\dagger) + \sqrt{1-t}(\sqrt{\eta}\hat{a}_2^\dagger), \sqrt{1-t}(\sqrt{\eta}\hat{a}_1^\dagger) - \sqrt{t}(\sqrt{\eta}\hat{a}_2^\dagger), \tag{6.47}$$

which is equal to the previous evolution.

■

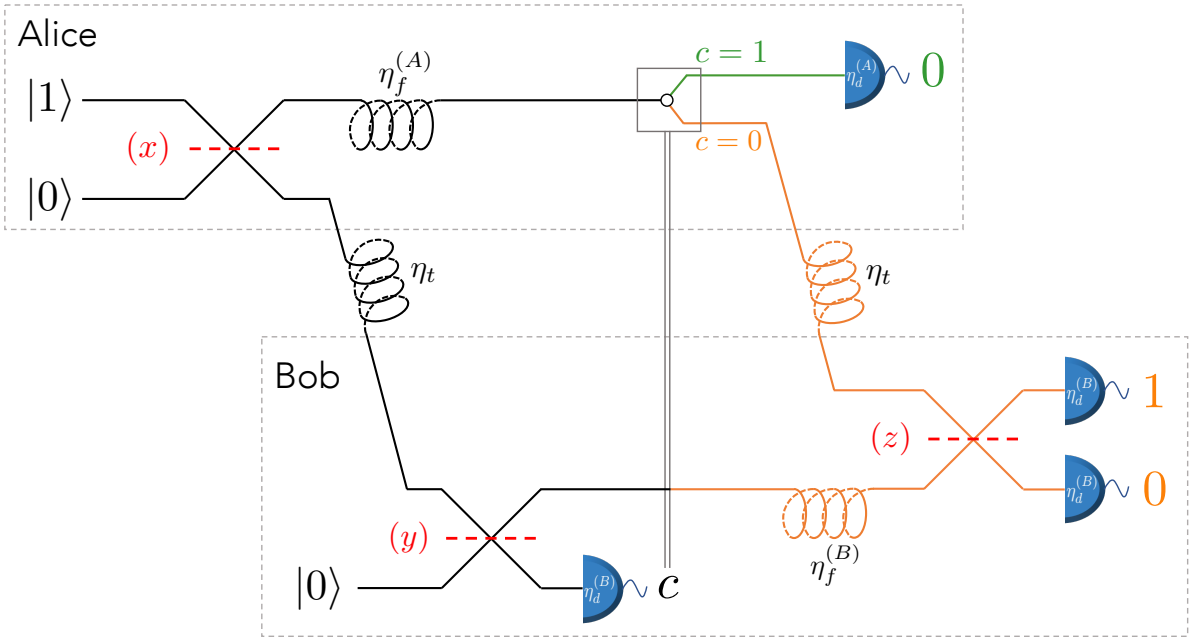

Figure 6.4: **Representation of the honest protocol with losses.** The dashed boxes indicate Alice and Bob's laboratories, respectively. Dashed red lines represent beam splitters, with the reflectance indicated in red. The efficiencies of the detectors, are indicated in white. Curly lines represent fiber used for quantum communication from Alice to Bob, or delay lines within Alice's or Bob's laboratory. $|0\rangle$ and $|1\rangle$ are the vacuum and single photon Fock states, respectively. Bob broadcasts the classical outcome c, which controls an optical switch on Alice's side. The protocol when Bob declares $c = 0/1$ is represented in orange/green. The final outcomes are the expected outcomes when both parties are honest.

In the presence of losses, the protocol may also abort when both parties are honest, when the photon is lost. We obtain the expressions for the honest winning probabilities $P_h^{(A)}$ and $P_h^{(B)}$, and hence the probability P_{ab} of abort, in the presence of losses:

Lemma 6.7.

$$\begin{aligned} P_h^{(A)} &= \eta_t \eta_d^{(B)} \left(\sqrt{xz\eta_f^{(A)}} + \sqrt{(1-x)y(1-z)\eta_f^{(B)}} \right)^2 \\ P_h^{(B)} &= \eta_t \eta_d^{(B)} (1-x)(1-y) \\ P_{ab} &= 1 - P_h^{(A)} - P_h^{(B)}. \end{aligned} \tag{6.48}$$

Proof. The honest winning probability for Bob is directly given by his chance of detecting the photon (the photon gets to his detector and doesn't get lost):

$$P_h^{(B)} = \eta_t \eta_d^{(B)} (1-x)(1-y). \tag{6.49}$$

On the other hand, Alice wins if the photon, starting from her first input mode, is detected by Bob in the last step.

The evolution of the creation operator of the first mode during the lossy honest protocol is given by:

$$\begin{aligned}
\hat{a}_1^\dagger &\to \sqrt{x}\hat{a}_1^\dagger + \sqrt{1-x}\hat{a}_2^\dagger \\
&\to \sqrt{x\eta_f^{(A)}}\hat{a}_1^\dagger + \sqrt{(1-x)\eta_t}\hat{a}_2^\dagger \\
&\to \sqrt{x\eta_f^{(A)}}\hat{a}_1^\dagger + \sqrt{(1-x)\eta_t y}\hat{a}_2^\dagger + \sqrt{(1-x)(1-y)\eta_t}\hat{a}_3^\dagger \\
&\to \sqrt{x\eta_f^{(A)}}\hat{a}_1^\dagger + \sqrt{(1-x)\eta_t y}\hat{a}_2^\dagger + \sqrt{(1-x)(1-y)\eta_t \eta_d^{(B)}}\hat{a}_3^\dagger \\
&\to \sqrt{x\eta_f^{(A)}\eta_t}\hat{a}_1^\dagger + \sqrt{(1-x)\eta_t y \eta_f^{(B)}}\hat{a}_2^\dagger + \sqrt{(1-x)(1-y)\eta_t \eta_d^{(B)}}\hat{a}_3^\dagger \\
&\to \left(\sqrt{x\eta_f^{(A)}\eta_t z} + \sqrt{(1-x)\eta_t y \eta_f^{(B)}(1-z)}\right)\hat{a}_1^\dagger + \left(\sqrt{x\eta_f^{(A)}\eta_t(1-z)} - \sqrt{(1-x)\eta_t y \eta_f^{(B)} z}\right)\hat{a}_2^\dagger \\
&\quad + \sqrt{(1-x)(1-y)\eta_t \eta_d^{(B)}}\hat{a}_3^\dagger \\
&\to \left(\sqrt{x\eta_f^{(A)}\eta_t z \eta_d^{(B)}} + \sqrt{(1-x)\eta_t y \eta_f^{(B)}(1-z)\eta_d^{(B)}}\right)\hat{a}_1^\dagger + \left(\sqrt{x\eta_f^{(A)}\eta_t(1-z)\eta_d^{(B)}} - \sqrt{(1-x)\eta_t y \eta_f^{(B)} z \eta_d^{(B)}}\right)\hat{a}_2^\dagger \\
&\quad + \sqrt{(1-x)(1-y)\eta_t \eta_d^{(B)}}\hat{a}_3^\dagger.
\end{aligned} \tag{6.50}$$

In particular, the photon reaches Bob's uppermost detector with probability

$$\begin{aligned}
P_h^{(A)} &= \left(\sqrt{x\eta_f^{(A)}\eta_t z \eta_d^{(B)}} + \sqrt{(1-x)\eta_t y \eta_f^{(B)}(1-z)\eta_d^{(B)}}\right)^2 \\
&= \eta_t \eta_d^{(B)} \left(\sqrt{xz\eta_f^{(A)}} + \sqrt{(1-x)y(1-z)\eta_f^{(B)}}\right)^2.
\end{aligned} \tag{6.51}$$

Finally, the protocol aborts for all other detection events:

$$P_{ab} = 1 - P_h^{(A)} - P_h^{(B)}. \tag{6.52}$$

■

Note that the overall correctness does not depend on Alice's detection efficiency $\eta_d^{(A)}$, since the declaration of outcome c depends solely on Bob's detector and the verification step on Alice's side involves detecting vacuum.

6.2.3 Losses: soundness

The soundness of the protocol is also affected by the presence of losses.

Dishonest Bob's best strategy is to perform the same attack as in the lossless case, because he has no control over Alice's half of the subsystem. His winning probability is then given by the following result:

Lemma 6.8. *Dishonest Bos's maximum winning probability is given by:*

$$P_d^{(B)} = 1 - x\eta_f^{(A)}\eta_d^{(A)}. \tag{6.53}$$

In a more general game-theoretic scenario, Bob's best strategy will in fact depend on the rewards and sanctions associated with honest aborts and 'getting caught cheating' aborts. In other words, Bob has to minimize his risk-to-reward ratio. Maximizing his winning probability makes him run the risk of getting caught cheating with probability $x\eta_f^{(A)}\eta_d^{(A)}$.

Dishonest Alice must still generate the state which maximizes the $(1,0,0)$ outcome on Bob's detectors after his honest transformations have been applied. However, the expression for Bob's corresponding projector now changes, as there is a finite probability $(1-\eta_d^{(B)})^n$ that the n-photon component is projected onto the vacuum. The 0 outcome on one spatial mode is therefore triggered by the projection $\Pi_0 = \sum_{n=0}^{\infty}(1-\eta_d^{(B)})^n |n\rangle\langle n|$. The total projector responsible for the $(1,0,0)$ outcome then reads $\Pi_{100} = (\mathbb{1}-\Pi_0)\otimes\Pi_0\otimes\Pi_0$.

Lemma 6.9. *Dishonest Alice's maximum winning probability is given by:*

$$\begin{aligned} P_d^{(A)} &= \max_{l>0}\left[\left(1-(1-y\eta_f^{(B)})(1-z)\eta_d^{(B)}\right)^l - \left(1-\eta_d^{(B)}\right)^l\right] \\ &\leq 1-(1-y)(1-z). \end{aligned} \tag{6.54}$$

The value of the upper bound in the second line is Alice's cheating probability in the lossless case. This shows that Alice cannot take advantage of Bob's imperfect detectors or his lossy delay line in order to increase her cheating probability.

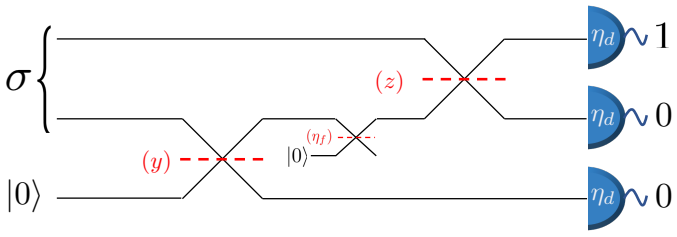

Figure 6.5: Alice aims to maximize the outcome $(1,0,0)$ by sending the state σ. The lossy delay line is represented by a mixing with the vacuum on a beam splitter of transmission amplitude η_f. The quantum efficiency of the detectors is indicated in white.

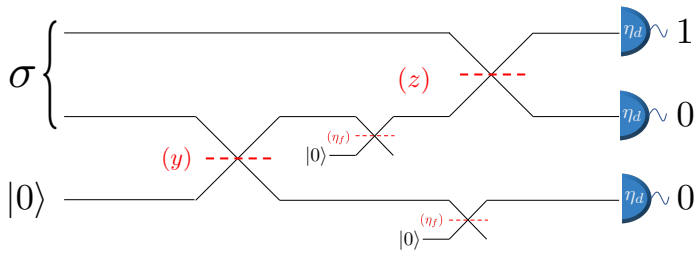

Figure 6.6: Adding losses on the third mode increases Alice's winning probability.

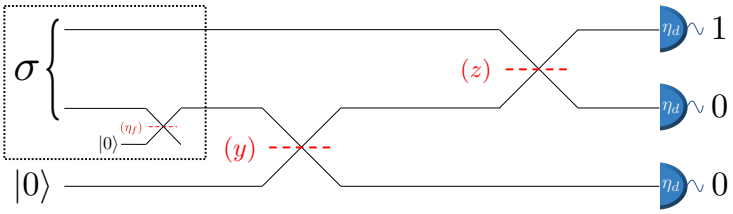

Figure 6.7: The losses η_f are commuted back to Alice's state preparation. The losses on input mode 3 can be omitted since the input state is the vacuum.

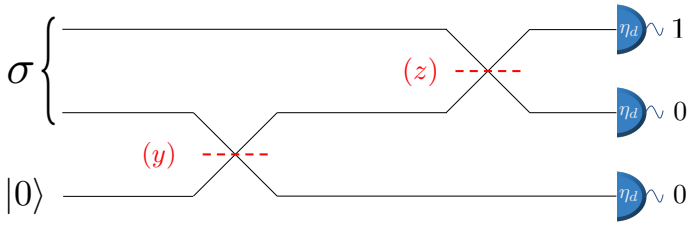

Figure 6.8: Alice aims to maximize the outcome $(1,0,0)$ by sending the state σ. The delay line efficiency η_f is equal to 1.

Proof. The losses η correspond to a probability $1-\eta$ of losing a photon. These can be modelled as a mixing with the vacuum on a beam splitter of reflectance η. We first show that we can obtain Alice's cheating probability by solving the case with perfect delay line, and replacing the parameter y by $y\eta_f$, independently of the efficiency η_d of his detectors.

The lossy delay line of efficiency η_f may be modelled as a mixing with the vacuum on a beam splitter of transmission η_f.

Alice prepares a state σ, which goes through the interferometer depicted in Fig. 6.5, and wins if the measurement outcome obtained by Bob is $(1,0,0)$.

In particular, note that the outcome 0 must be obtained for the third mode. Hence Alice's winning probability is always lower than if the third mode was mixed with the vacuum on

a beam splitter of transmission amplitude η_f just before the detection (Fig. 6.6), since this increases the probability of the outcome 0 for this mode.

Let us assume that this is the case. Then, by Lemma 6.6, the losses η_f on output modes 2 and 3 may be commuted back through the beam splitter of reflectance y, acting on modes 2 and 3.

Since the input state on mode 3 is the vacuum, the losses on this mode may then be removed (Fig. 6.7). In that case, the probability of winning is clearly lower than when the delay line is perfect (Fig. 6.8), because Alice is now restricted to lossy state preparation instead of ideal state preparation.

This reduction shows that Alice's maximum winning probability when Bob is using a lossy delay line is always lower than when Bob's delay line is perfect, independently of the efficiency η_d of his detectors.

Moreover, Alice's maximum cheating probability and optimal cheating strategy may be inferred from the case where Bob has a perfect delay line, as we show in what follows.

By convexity of the probabilities, Alice's best strategy is to send a pure state $|\psi\rangle = \sum_{k,l\geq 0} \psi_{kl} |kl\rangle$. Let us denote by W the interferometer depicted in Fig. 6.5, including the detection losses. Let us consider the evolution of Alice's state and the vacuum on the third input mode through the interferometer W. The creation operator for the first mode evolves as

$$\begin{aligned} \hat{a}_1^\dagger &\to \sqrt{z}\,\hat{a}_1^\dagger + \sqrt{1-z}\,\hat{a}_2^\dagger \\ &\to \sqrt{z\eta_d}\,\hat{a}_1^\dagger + \sqrt{(1-z)\eta_d}\,\hat{a}_2^\dagger \\ &= W\hat{a}_1^\dagger W^\dagger, \end{aligned} \tag{6.55}$$

while the creation operator for the second mode evolves as

$$\begin{aligned} \hat{a}_2^\dagger &\to \sqrt{y}\,\hat{a}_2^\dagger + \sqrt{1-y}\,\hat{a}_3^\dagger \\ &\to \sqrt{y\eta_f}\,\hat{a}_2^\dagger + \sqrt{1-y}\,\hat{a}_3^\dagger \\ &\to \sqrt{y(1-z)\eta_f}\,\hat{a}_1^\dagger - \sqrt{yz\eta_f}\,\hat{a}_2^\dagger + \sqrt{1-y}\,\hat{a}_3^\dagger \\ &\to \sqrt{y(1-z)\eta_f\eta_d}\,\hat{a}_1^\dagger - \sqrt{yz\eta_f\eta_d}\,\hat{a}_2^\dagger + \sqrt{(1-y)\eta_d}\,\hat{a}_3^\dagger \\ &= W\hat{a}_2^\dagger W^\dagger. \end{aligned} \tag{6.56}$$

Hence, the output state (before the ideal threshold detection) is given by

$$\begin{aligned} W|\psi 0\rangle &= W \sum_{k,l\geq 0} \psi_{kl} |kl0\rangle \\ &= W\left[\sum_{k,l\geq 0} \frac{\psi_{kl}}{\sqrt{k!l!}} (\hat{a}_1^\dagger)^k (\hat{a}_2^\dagger)^l\right] |000\rangle \end{aligned}$$

$$= \left[\sum_{k,l\geq 0} \frac{\psi_{kl}}{\sqrt{k!l!}} (W\hat{a}_1^\dagger W^\dagger)^k (W\hat{a}_2^\dagger W^\dagger)^l \right] |000\rangle \tag{6.57}$$

$$= \left[\sum_{k,l\geq 0} \frac{\psi_{kl}}{\sqrt{k!l!}} \left(\sqrt{z\eta_d}\,\hat{a}_1^\dagger + \sqrt{(1-z)\eta_d}\,\hat{a}_2^\dagger \right)^k \right.$$

$$\left. \times \left(\sqrt{y(1-z)\eta_f\eta_d}\,\hat{a}_1^\dagger - \sqrt{yz\eta_f\eta_d}\,\hat{a}_2^\dagger + \sqrt{(1-y)\eta_d}\,\hat{a}_3^\dagger \right)^l \right] |000\rangle .$$

Now Alice's maximum cheating probability is given by

$$P_d^{(A)} = \mathrm{Tr}[W|\psi 0\rangle\langle\psi 0|W^\dagger(\mathbb{1} - |0\rangle\langle 0|)|00\rangle\langle 00|]. \tag{6.58}$$

Hence, the state after a successful projection $(\mathbb{1} - |0\rangle\langle 0|)|00\rangle\langle 00|$, which has norm $P_d^{(A)}$, reads

$$\left[\sum_{k+l>0} \frac{\psi_{kl}}{\sqrt{k!l!}} (z\eta_d)^{k/2} [y(1-z)\eta_f\eta_d]^{l/2} (\hat{a}_1^\dagger)^{k+l} \right] |000\rangle . \tag{6.59}$$

When Bob has a perfect delay line ($\eta_f = 1$) this state reads

$$\left[\sum_{k+l>0} \frac{\psi_{kl}}{\sqrt{k!l!}} (z\eta_d)^{k/2} [y(1-z)\eta_d]^{l/2} (\hat{a}_1^\dagger)^{k+l} \right] |000\rangle , \tag{6.60}$$

and its norm is the winning probability of Alice in that case. Hence,

$$P_d^{(A)}[\eta_f, \eta_d, y, z] = P_d^{(A)}[1, \eta_d, y\eta_f, z], \tag{6.61}$$

i.e., we can obtain Alice's cheating probability by solving the case with perfect delay line, and replacing the parameter y by $y\eta_f$. In the following, we thus derive Alice's optimal strategy in that case.

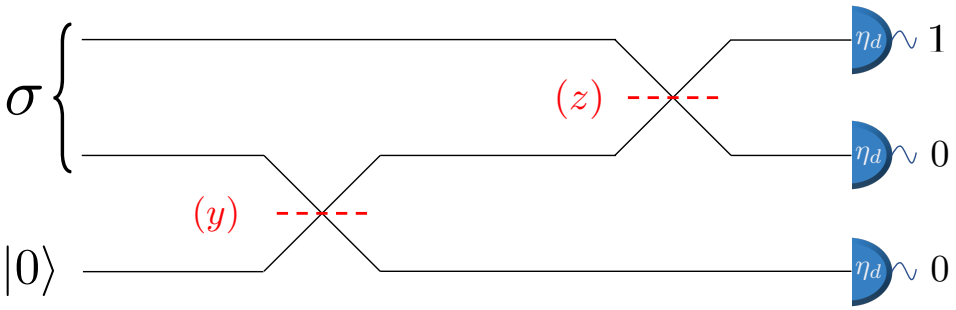

Figure 6.9: Alice aims to maximize the outcome $(1,0,0)$ by sending the state σ. The quantum efficiency of the detectors is indicated in white.

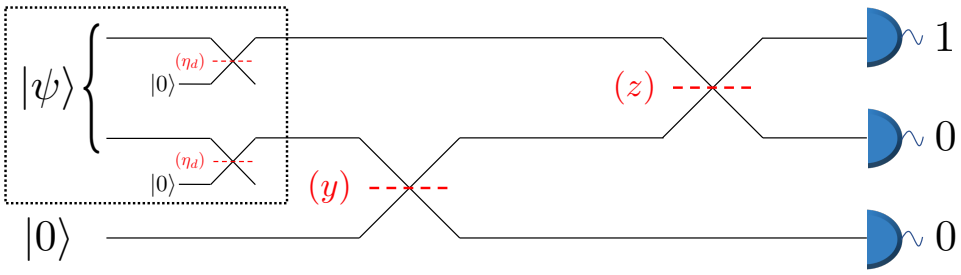

Figure 6.10: The quantum efficiency are modelled as losses η_d on modes 1, 2, and 3, which are then commuted through the interferometer, back to Alice's state preparation. The losses on input mode 3 can be omitted since the input state is the vacuum.

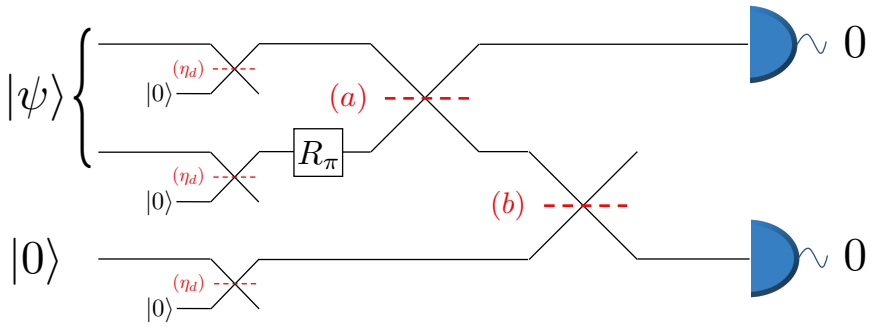

Figure 6.11: An equivalent picture for the first term P_1 of Eq. (6.66). The term P_1 is the probability of the simultaneous outcomes 0 for modes 1 and 3.

Let σ be the state sent by Alice, and η_d the detector efficiency. Alice needs to maximize the probability of the overall outcome $(1,0,0)$ at the output of the interferometer depicted in Fig. 6.9, hence the overlap with the projector:

$$\Pi^{\eta_d}_{(1,0,0)} = \left[\mathbb{1} - \sum_m (1-\eta_d)^m |m\rangle\langle m|\right] \otimes \left[\sum_{n,p} (1-\eta_d)^{n+p} |n\rangle\langle n| \otimes |p\rangle\langle p|\right]. \tag{6.62}$$

By convexity of the probabilities, we may assume without loss of generality that Alice sends a pure state $\sigma = |\psi\rangle\langle\psi|$. Moreover, the imperfect threshold detectors of quantum efficiency η_d can be modelled by mixing the state to be measured with the vacuum on a beam splitter of transmission amplitude η_d followed by an ideal threshold detection [FOP05]. In that case, this corresponds to losses η_d on modes 1, 2, and 3, followed by ideal threshold detections. By Lemma 6.6, commuting the losses back through the interferometer leads to the equivalent picture depicted in Fig. 6.10, where the losses on input mode 3 have been omitted, since the input state is the vacuum.

In that case, Alice's probability of winning is clearly lower than when the threshold

detectors are perfect (Fig. 6.2), because she is restricted to lossy state preparation instead of ideal state preparation. Let $|\tilde{\psi}\rangle$ be the lossy state obtained by applying losses η_d on both modes of Alice's prepared state $|\psi\rangle$. Alice's winning probability may then be written:

$$\begin{aligned} P_d^{(A)} &= \mathrm{Tr}[U(|\tilde{\psi}\rangle\langle\tilde{\psi}|\otimes|0\rangle\langle 0|)U^\dagger(\mathbb{1}-|0\rangle\langle 0|)\otimes|00\rangle\langle 00|] \\ &= \mathrm{Tr}[U(|\tilde{\psi}\rangle\langle\tilde{\psi}|\otimes|0\rangle\langle 0|)U^\dagger(\mathbb{1}\otimes|00\rangle\langle 00|)] - \mathrm{Tr}[U(|\tilde{\psi}\rangle\langle\tilde{\psi}|\otimes|0\rangle\langle 0|)U^\dagger|000\rangle\langle 000|], \end{aligned} \tag{6.63}$$

where $U = (H^{(z)}\otimes\mathbb{1})(\mathbb{1}\otimes H^{(y)})$ is the unitary corresponding to the general interferometer of the lossless protocol. By Lemma 6.1, we have

$$\mathrm{Tr}[(\tau\otimes|0\rangle\langle 0|)U^\dagger(\mathbb{1}\otimes|00\rangle\langle 00|)U] = \mathrm{Tr}[(\tau\otimes|0\rangle\langle 0|)V^\dagger(|0\rangle\langle 0|\otimes\mathbb{1}\otimes|0\rangle\langle 0|)V], \tag{6.64}$$

for any density matrix τ, where $V = (\mathbb{1}\otimes H^{(b)})(H^{(a)}\otimes\mathbb{1})(\mathbb{1}\otimes R(\pi)\otimes\mathbb{1})$, with $a = \frac{y(1-z)}{y+z-yz}$ and $b = y+z-yz$, and $R(\pi)$ a phase shift of π acting on mode 2. Hence,

$$P_d^{(A)} = \mathrm{Tr}[V(|\tilde{\psi}\rangle\langle\tilde{\psi}|\otimes|0\rangle\langle 0|)V^\dagger(|0\rangle\langle 0|\otimes\mathbb{1}\otimes|0\rangle\langle 0|)] - \mathrm{Tr}[|\tilde{\psi}\rangle\langle\tilde{\psi}||00\rangle\langle 00|], \tag{6.65}$$

where we used $U^\dagger|000\rangle = |000\rangle$ for the second term. Setting $|\tilde{\psi}_x\rangle = (H^{(a)}\otimes\mathbb{1})(\mathbb{1}\otimes R(\pi))|\tilde{\psi}\rangle$ yields

$$P_d^{(A)} = \underbrace{\mathrm{Tr}[(|\tilde{\psi}_x\rangle\langle\tilde{\psi}_x|\otimes|0\rangle\langle 0|)(\mathbb{1}\otimes H^{(b)})(|0\rangle\langle 0|\otimes\mathbb{1}\otimes|0\rangle\langle 0|)(\mathbb{1}\otimes H^{(b)})]}_{\equiv P_1} - \underbrace{\mathrm{Tr}[|\tilde{\psi}_x\rangle\langle\tilde{\psi}_x||00\rangle\langle 00|]}_{\equiv P_2}, \tag{6.66}$$

where we used $|00\rangle = (\mathbb{1}\otimes R(\pi))H^{(a)}|00\rangle$ for the second term P_2.

Let us consider the first term P_1. Since $|\tilde{\psi}\rangle$ is the state obtained by applying losses η_d on both modes of the state $|\psi\rangle$, we obtain the equivalent picture in Fig. 6.11, where we have added losses η_d also on mode 3, since the input state is the vacuum.

Let $|\psi_x\rangle = H^{(a)}(\mathbb{1}\otimes R(\pi))|\psi\rangle$. With Lemma 6.6, commuting the losses η_d to the output of the interferometer in Fig. 6.11, and combining the losses on mode 2 and 3 yields

$$P_1 = \mathrm{Tr}[|\psi_x\rangle\langle\psi_x|\,\Pi_{(0)}^{\eta_d}\otimes\Pi_{(0)}^{\eta_d(1-b)}], \tag{6.67}$$

where $\Pi_{(0)}^{\eta}$ is the POVM element corresponding to no click for a threshold detector of quantum efficiency η (recall that this is the same as an ideal detector preceded by a mixing with the vacuum on a beam splitter of transmission amplitude η). The same reasoning for the second term P_2 gives

$$P_2 = \mathrm{Tr}[|\psi_x\rangle\langle\psi_x|\,\Pi_{(0)}^{\eta_d}\otimes\Pi_{(0)}^{\eta_d}], \tag{6.68}$$

and we finally obtain with Eq. (6.66),

$$P_d^{(A)} = \mathrm{Tr}[|\psi_x\rangle\langle\psi_x|\,\Pi_{(0)}^{\eta_d}\otimes(\Pi_{(0)}^{\eta_d(1-b)} - \Pi_{(0)}^{\eta_d})]. \tag{6.69}$$

Let us write $|\psi_x\rangle = \sum_{k,l\geq 0}^{+\infty} \psi_{kl} |kl\rangle$. With the expression of the POVM in Eq. (6.62) the last equation reads

$$\begin{aligned} P_d^{(A)} &= \sum_{k,l\geq 0} |\psi_{kl}|^2 (1-\eta_d)^k [(1-\eta_d(1-b))^l - (1-\eta_d)^l] \\ &\leq \max_{k,l\geq 0} (1-\eta_d)^k [(1-\eta_d(1-b))^l - (1-\eta_d)^l] \sum_{k,l\geq 0} |\psi_{kl}|^2 \\ &= \max_{k,l\geq 0} (1-\eta_d)^k [(1-\eta_d(1-b))^l - (1-\eta_d)^l] \\ &= \max_{l\geq 1} [(1-\eta_d(1-b))^l - (1-\eta_d)^l] \\ &= \max_{l\geq 1} [(1-\eta_d(1-y)(1-z))^l - (1-\eta_d)^l], \end{aligned} \tag{6.70}$$

where we used $b = y + z - yz$. Let $l_0 \in \mathbb{N}^*$ such that $\max_{l\geq 1}[(1-\eta_d(1-b))^l - (1-\eta_d)^l] = (1-\eta_d(1-b))^{l_0} - (1-\eta_d)^{l_0}$. This last expression is an upperbound for $P_d^{(A)}$, which is attained for $\psi_{kl} = \delta_{k,0}\delta_{l,l_0}$, i.e., $|\psi_x\rangle = |0 l_0\rangle$. Thus, the best strategy for Alice is to send the state

$$\begin{aligned} |\psi\rangle &= (\mathbb{1} \otimes R(\pi)) H^{(a)} |\psi_x\rangle \\ &= (\mathbb{1} \otimes R(\pi)) H^{(a)} |0 l_0\rangle, \end{aligned} \tag{6.71}$$

where $a = \frac{y(1-z)}{y+z-yz}$, and her winning probability is then

$$P_d^{(A)} = (1-\eta_d(1-y)(1-z))^{l_0} - (1-\eta_d)^{l_0}, \tag{6.72}$$

when Bob has a perfect delay line. Recalling Eq. (6.61), the best strategy for Alice when Bob has a lossy delay line of efficiency η_f is to send the state

$$\begin{aligned} |\psi\rangle &= (\mathbb{1} \otimes R(\pi)) H^{(a)} |\psi_x\rangle \\ &= (\mathbb{1} \otimes R(\pi)) H^{(a)} |0 l_1\rangle, \end{aligned} \tag{6.73}$$

where $a = \frac{y(1-z)\eta_f}{y\eta_f + z - yz\eta_f}$, and $l_1 \in \mathbb{N}^*$ maximizes $(1-\eta_d(1-y\eta_f)(1-z))^l - (1-\eta_d)^l$. Her winning probability is then

$$\begin{aligned} P_d^{(A)} &= \max_{l>0} \left[\left(1-(1-y\eta_f)(1-z)\eta_d\right)^l - \left(1-\eta_d\right)^l \right] \\ &= (1-\eta_d(1-y\eta_f)(1-z))^{l_1} - (1-\eta_d)^{l_1} \\ &= \eta_d [1-(1-y\eta_f)(1-z)] \sum_{j=0}^{l_1-1} (1-\eta_d)^j (1-\eta_d(1-y\eta_f)(1-z))^{l_1-j-1} \\ &\leq \eta_d [1-(1-y\eta_f)(1-z)] \sum_{j=0}^{l_1-1} (1-\eta_d)^j \\ &= \eta_d [1-(1-y\eta_f)(1-z)] \frac{1-(1-\eta_d)^{l_1}}{1-(1-\eta_d)} \end{aligned} \tag{6.74}$$

$$
\begin{aligned}
&= [1-(1-y\eta_f)(1-z)][1-(1-\eta_d)^{l_1}] \\
&\leq 1-(1-y\eta_f)(1-z) \\
&\leq 1-(1-y)(1-z),
\end{aligned}
$$

and this last expression is her winning probability when there are no losses.

■

Let us derive the value of l_1 for which the maximum is achieved in Eq. (6.54). For this, we define:

$$
\begin{aligned}
r &= 1-\eta_d(1-y\eta_f)(1-z) \\
s &= 1-\eta_d.
\end{aligned}
\tag{6.75}
$$

We then consider a $\lambda_1 \in \mathbb{R}^{*+}$ which maximizes $(r^\lambda - s^\lambda)$ for $\lambda \in \mathbb{R}^{*+}$. We have that:

$$
\frac{d}{d\lambda_1}(r^{\lambda_1}-s^{\lambda_1}) = 0 \Leftrightarrow \lambda_1 = \frac{\log\log s - \log\log r}{\log r - \log s},
\tag{6.76}
$$

for strictly non-zero r and s. This allows us to deduce:

$$
l_1 = \begin{cases} \lfloor \lambda_1 \rfloor & \text{if} \quad r^{\lfloor\lambda_1\rfloor} - s^{\lfloor\lambda_1\rfloor} \geq r^{\lceil\lambda_1\rceil} - s^{\lceil\lambda_1\rceil} \\ \lceil \lambda_1 \rceil & \text{if} \quad r^{\lceil\lambda_1\rceil} - s^{\lceil\lambda_1\rceil} \geq r^{\lfloor\lambda_1\rfloor} - s^{\lfloor\lambda_1\rfloor}. \end{cases}
\tag{6.77}
$$

6.2.4 Quantum advantage

We now analyze the performance of our protocol in a practical setting, by enforcing three conditions on the free parameters: the protocol must be fair, balanced, and perform strictly better than any classical protocol. The latter condition is not required in an ideal implementation, since quantum weak coin flipping always provides a security advantage over classical weak coin flipping. Allowing for abort cases, however, may enable some classical protocols to perform better than quantum ones. This is because increasing the abort probability effectively decreases Alice and Bob's cheating probabilities. We say that the protocol allows for quantum advantage when it provides a strictly lower cheating probability than any classical protocol with the same abort probability. This is obtained using the bounds from [HW11], which yield the best classical cheating probability $P_d^C = 1-\sqrt{P_{ab}}$ for our protocol.

Condition (i): the first condition enforces a fair protocol, i.e., $P_h^{(A)} = P_h^{(B)}$. With Eq. (6.48), we aim to solve for y as a function of x and z:

$$
\begin{aligned}
(i) &\Leftrightarrow \eta_t\eta_d^{(B)}\left(\sqrt{xz\eta_f^{(A)}} + \sqrt{(1-x)y(1-z)\eta_f^{(B)}}\right)^2 = \eta_t\eta_d^{(B)}(1-x)(1-y) \\
&\Leftrightarrow (1-x)\left[(1-z)\eta_f^{(B)}+1\right]y + 2\sqrt{x(1-x)z(1-z)\eta_f^{(A)}\eta_f^{(B)}}\sqrt{y} + xz\eta_f^{(A)} - (1-x) = 0.
\end{aligned}
\tag{6.78}
$$

We make the substitution $Y = \sqrt{y}$ in order to transform Eq. (6.78) into a second-order polynomial equation. We then take only the positive solution (since y must be positive) which reads:

$$Y = \frac{\sqrt{xz(1-z)\eta_f^{(A)}\eta_f^{(B)} - \left[(1-z)\eta_f^{(B)}+1\right]\left[xz\eta_f^{(A)}-(1-x)\right]} - \sqrt{xz(1-z)\eta_f^{(A)}\eta_f^{(B)}}}{\sqrt{1-x}\left[(1-z)\eta_f^{(B)}+1\right]}. \tag{6.79}$$

We may finally write:

$$(i) \Leftrightarrow \; y = f\left(x,z,\eta_f^{(i)},\eta_d,\eta_t\right), \tag{6.80}$$

where

$$f\left(x,z,\eta_f^{(i)},\eta_d,\eta_t\right) = \frac{\left(\sqrt{(1-x)\left[(1-z)\eta_f^{(B)}+1\right]-xz\eta_f^{(A)}} - \sqrt{xz(1-z)\eta_f^{(A)}\eta_f^{(B)}}\right)^2}{(1-x)\left[(1-z)\eta_f^{(B)}+1\right]^2}. \tag{6.81}$$

Note that y should be a real number, and hence we require that the expression under the first square root of $f\left(x,z,\eta_f^{(i)},\eta_d,\eta_t\right)$ is positive, i.e.,

$$z \leq \frac{(1-x)(1+\eta_f^{(B)})}{x\eta_f^{(A)}+(1-x)\eta_f^{(B)}}. \tag{6.82}$$

Furthermore, note that, for $\eta_f^{(A)} = \eta_f^{(B)} = \eta_f$, y should be an increasing function of η_f, and therefore a decreasing function of d when $\eta_f = 10^{-\frac{0.2}{10}2d}$. Mathematically speaking, this is to prevent $y'(d) \to \infty$ and $y(d) > 1$. Physically speaking, this condition ensures that, as the probability of transmitting the photon (and of preserving it for verification) gets smaller, Bob should encourage a detection on the third mode, which evens out the honest probabilities of winning.

Condition (ii): the second condition enforces a balanced protocol, i.e., $P_d^{(A)} = P_d^{(B)}$. With Eqs. (6.53) and (6.54), this translates into the following expression for x:

$$(ii) \Leftrightarrow x = g\left(y,z,\eta_f^{(i)},\eta_d^{(i)}\right), \tag{6.83}$$

where

$$g\left(y,z,\eta_f^{(i)},\eta_d^{(i)}\right) = \frac{1}{\eta_f^{(A)}\eta_d^{(A)}}\left[1 - \max_{l\geq 1}[(1-\eta_d^{(B)}(1-y\eta_f^{(B)})(1-z))^l - (1-\eta_d^{(B)})^l]\right]. \tag{6.84}$$

Condition (iii): we recall the general coin flipping formalism from [HW11], in which any classical or quantum coin flipping protocol may be expressed as:

$$CF(p_{00},p_{11},p_{*0},p_{*1},p_{0*},p_{1*}), \tag{6.85}$$

where p_{ii} is the probability that two honest players output value $i \in \{0,1\}$, p_{*i} is the probability that Dishonest Alice forces Honest Bob to declare outcome i, and p_{i*} is the probability that

Dishonest Bob forces Honest Alice to declare outcome i. In this formalism, a perfect strong coin flipping protocol can then be expressed as $CF\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right)$, while a perfect weak coin flipping may be expressed as $CF\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},1,1,\frac{1}{2}\right)$. We may now express our quantum weak coin flipping protocol in the lossless setting as:

$$CF\left(\frac{1}{2},\frac{1}{2},\left[\frac{1}{2(1-x)}\right],1,1,[1-x]\right). \tag{6.86}$$

In the lossy setting, note that the probabilities that Alice and Bob each choose to lose (i.e., p_{*1} and p_{0*}, respectively), both remain 1. When Dishonest Bob chooses to lose, he may always declare outcome 0 regardless of what he detects, which yields $p_{0*}=1$. When Dishonest Alice chooses to lose, she may send a state $|n\rangle$ to Bob, and so:

$$\begin{aligned} p_{*1} &= \mathrm{Tr}\left[H^{(y)}|n0\rangle\langle n0|H^{(y)}I\otimes(I-\Pi_0)\right] \\ &= 1-\mathrm{Tr}\left[H^{(y)}|n0\rangle\langle n0|H^{(y)}(I\otimes\Pi_0)\right], \end{aligned} \tag{6.87}$$

where $\Pi_0=\sum_{l\geq 0}(1-\eta)^l|l\rangle\langle l|$ and $H^{(y)}=\begin{pmatrix}\sqrt{y} & \sqrt{1-y}\\ \sqrt{1-y} & -\sqrt{y}\end{pmatrix}$.

Now,

$$\begin{aligned} H^{(y)}|n0\rangle &= H^{(y)}\frac{(\hat{a}_1^\dagger)^n}{\sqrt{n!}}|00\rangle \\ &= \frac{1}{\sqrt{n!}}(\sqrt{y}\hat{a}_1^\dagger+\sqrt{1-y}\hat{a}_2^\dagger)^n|00\rangle \\ &= \frac{1}{\sqrt{n!}}\sum_{k=0}^{n}\binom{n}{k}y^{\frac{k}{2}}(1-y)^{\frac{n-k}{2}}\hat{a}_1^{\dagger k}\hat{a}_2^{\dagger(n-k)}|00\rangle \\ &= \sum_{k=0}^{n}\sqrt{\binom{n}{k}y^k(1-y)^{n-k}}\,|k\ (n-k)\rangle. \end{aligned} \tag{6.88}$$

We thus obtain, by linearity of the trace:

$$\begin{aligned} p_{*1} &= 1-\sum_{l,l'\geq 0}(1-\eta)^l\sum_{k,k'=0}^{n}\sqrt{\binom{n}{k}y^k(1-y)^{n-k}}\sqrt{\binom{n}{k'}y^{k'}(1-y)^{n-k'}}\,\mathrm{Tr}\left[|k\ (n-k)\rangle\langle k'\ (n-k')||l'l\rangle\langle l'l|\right] \\ &= 1-\sum_{k=0}^{n}(1-\eta)^{n-k}\binom{n}{k}y^k(1-y)^{n-k} \\ &= 1-\left[y+(1-\eta)(1-y)\right]^n, \end{aligned} \tag{6.89}$$

which goes to 1 when n goes to infinity, for $y<1$. Hence, in the lossy setting, the protocol becomes a:

$$CF\left(P_h^{(A)},P_h^{(B)},P_d^{(A)},1,1,P_d^{(B)}\right), \tag{6.90}$$

where $P_d^{(A)}=\max_{l>0}\left(1-(1-y\eta_f^{(A)})(1-z)\eta_d^{(B)}\right)^l-\left(1-\eta_d^{(B)}\right)^l$ and $P_d^{(B)}=1-x\eta_f^{(A)}\eta_d^{(A)}$.

Using Theorem 1 from [HW11], there exists a classical protocol that implements an information-theoretically secure coin flip with our parameters if and only if the following conditions hold:

$$\begin{cases} P_h^{(A)} \leq P_d^{(A)} \\ P_h^{(B)} \leq P_d^{(B)} \\ P_{ab} = 1 - P_h^{(A)} - P_h^{(B)} \geq (1 - P_d^{(A)})(1 - P_d^{(B)}). \end{cases} \tag{6.91}$$

Our quantum protocol therefore presents an advantage over classical protocols if at least one of these conditions *cannot* be satisfied. Since we are interested in fair and balanced protocols, setting

$$P_h = P_h^{(A)} = P_h^{(B)}, \quad \text{and} \quad P_d = P_d^{(A)} = P_d^{(B)} \tag{6.92}$$

allows us to rewrite (6.91) as:

$$\begin{cases} P_h \leq P_d \\ P_{ab} = 1 - 2P_h \geq (1 - P_d)^2 \Leftrightarrow P_h \leq \frac{1}{2}[1 - (1 - P_d)^2]. \end{cases} \tag{6.93}$$

Let us finally remark that for all x we have $\frac{1}{2}[1-(1-x)^2] = x - \frac{x^2}{2} \leq x$, so the first inequality above is implied by the second. The system is thus equivalent to the second inequality:

$$P_{ab} = 1 - 2P_h \geq (1 - P_d)^2, \tag{6.94}$$

provided that $P_h^{(A)} = P_h^{(B)} = P_h$ and $P_d^{(A)} = P_d^{(B)} = P_d$.

In order to get a clearer insight into the meaning of quantum advantage, we express this condition in terms of cheating probability: our protocol displays quantum advantage if and only if the lowest classical cheating probability

$$P_d^C = 1 - \sqrt{1 - 2P_h} = 1 - \sqrt{P_{ab}} \tag{6.95}$$

exceeds our quantum cheating probability P_d^Q.

The three conditions may then be translated into the following system of equations, where we define $P_d^Q = P_d^{(A)} = P_d^{(B)}$:

$$\begin{cases} (i) & P_h^{(A)} = P_h^{(B)} \text{ fairness} \\ (ii) & P_d^{(A)} = P_d^{(B)} \text{ balance} \\ (iii) & P_d^Q < P_d^C \text{ quantum advantage} \end{cases} \tag{6.96}$$

Fig. 6.12 shows a choice of parameters obtained numerically for which the system in Eq. (6.96) is satisfied, up to a distance of d km.

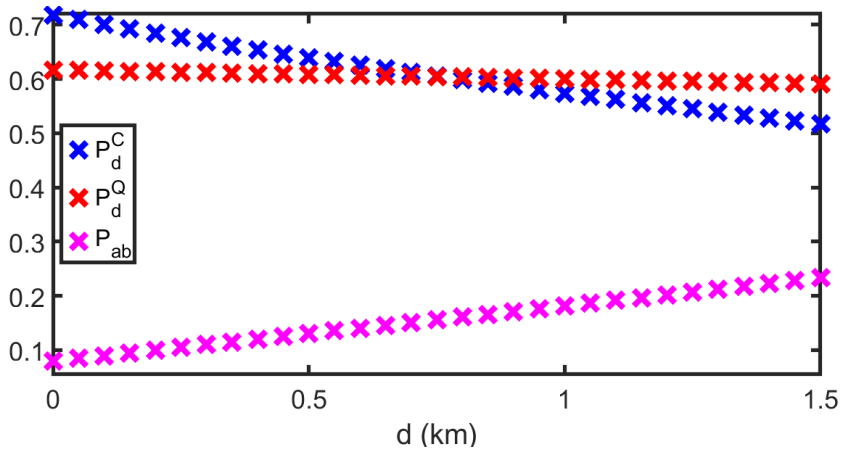

Figure 6.12: Practical quantum advantage for a fair and balanced protocol: numerical values for the lowest classical and quantum cheating probabilities, P_d^C and P_d^Q, are plotted as a function of distance d in blue and red, respectively. Honest abort probability P_{ab} (responsible for P_d^Q being lower than our ideal quantum cheating probability $1/\sqrt{2}$) is plotted in magenta. Our quantum protocol performs strictly better than any classical protocol when $P_d^Q < P_d^C$. We set $\eta_f = \eta_s \eta_t^2$, where η_s is the fiber delay transmission corresponding to 500ns of optical switching time, and $\eta_t^2 = (10^{-\frac{0.2}{10}d})^2$ is the fiber delay transmission associated with travelling distance d twice (once for quantum, once for classical) in single-mode fibers with attenuation 0.2 dB/km. We have $\eta_d = 0.95$ and $z = 0.57$.

6.3 Discussion and open problems

By noticing a non-trivial connection between the early protocol from [SR02] and linear optical transformations, we answer the question of the implementability of quantum weak coin flipping, and show that it is achievable with current technology over a few hundred meters. As the distance increases, the issue of stability of the interferometric setup should also be taken into account. Both parties require a set of beam splitters and single photon threshold detectors. State generation on Alice's side can be performed with any heralded probabilistic single-photon source, for which photon indistinguishability and state purity do not matter. Only Alice requires an optical switch, which is commercially available. Although short-term quantum storage is needed, a spool of optical fiber with twice the length of the quantum channel suffices, and provides the required storage/retrieval efficiency.

On the fundamental level, our results also raise the question of a potentially deeper connection between the large family of protocols from [Moc04, Moc05, Moc07]—which achieves biases as low as 1/6—and linear optics. Recalling that the protocol from [SR02], and hence our protocol, is conjectured optimal for this family, its extension to many rounds should be necessary in order to lower the bias. The optimality of the one-round protocol is crucial, as a recent result shows that the weak coin flipping bias decreases very inefficiently with the number of rounds [Mil20].

CONCLUSION AND OUTLOOK

Guided by three general questions about the use of quantum information in existing and upcoming technologies, this thesis has provided some answers in the context of continuous variable quantum information theory and linear quantum optics.

Firstly, *what leads to a quantum advantage?*

We have considered the case of non-Gaussian states as a resource for outperforming classical computing capabilities. Introducing the stellar formalism, we have characterised single-mode non-Gaussian states by the number of elementary non-Gaussian operations needed to engineer them [CMG20b]. Apart from providing insights about the structure of these states, we have seen direct consequences of the use of our formalism for Gaussian convertibility of states, comparing photon addition and photon subtraction, and cat state engineering [CRW$^+$20].

We have studied classical simulation regimes for a variety of continuous variable and optical quantum models [CMS20, CMG20a]. Our conclusions provide the minimum requirements necessary for the development of beyond-classical quantum applications with these models. Bridging the gap between classically simulable models and models universal for quantum computing, we have shown that CVS circuits, which form a subuniversal family of optical interferometers relating to Boson Sampling with Gaussian measurements, are hard to simulate classically [CDM$^+$17].

Secondly, *how do we check the correct functioning of a quantum device?*

We have developped a variety of certification and verification protocols for continuous variable quantum states with single-mode Gaussian measurements. As a first step, we showed how to perform efficient reliable tomography and fidelity estimation for any single-mode continuous variable quantum state, under the assumption of identical copies, and with no assumption whatsoever [CDG$^+$19]. Next, we showed how to obtain tight fidelity witnesses for a large class of multimode continuous variable quantum states with analytical confidence intervals [CGKM20]. These fidelity witnesses in turn allowed us to derive a verification protocol for multimode states, including the output states of Boson Sampling experiments, which was missing so far [EHW$^+$20], thus enabling an experimental demonstration of quantum supremacy with photonic quantum computing.

Thirdly, *what useful advantages can we obtain from the use of quantum information?*

We have considered various applications of quantum information and their implementation with linear optics. We have analysed the task of discriminating two unknown quantum states in an unbalanced setting, showing a connection with the concept of universal quantum-programmable measurement. For these two tasks, we have introduced an optimal implementation with linear optics and single-photon encoding [CDM$^+$18]. To obtain a more practical setup, we have discussed coherent state encoding which simplifies considerably the corresponding scheme [KCK$^+$20].

Turning to quantum cryptography, we have proposed, using once again linear optics, the first practical implementation of quantum weak coin flipping with information-theoretic security [BCKD20], a building block for a variety of cryptographic applications. We have analysed the robustness of our proposal to experimental imperfections and losses and showed that an experimental implementation could display quantum advantage already with current technology.

We have given various open problems at the end of each chapter. Let us finish by considering more general perspectives.

We believe that demonstrating verified quantum supremacy with photonic quantum computing using our fidelity witness protocol is a fascinating prospect, either for Boson Sampling or CVS circuits. Given its efficiency, the verification protocol is already within experimental reach. This would represent a milestone in the development of quantum technologies and fundamentally demonstrate the different nature of quantum and classical computation.

Other immediate perspectives are the ongoing implementations of proof-of-concept experiments for demonstrating the stellar rank with Gaussian measurements, performing quantum-programmable measurements with coherent states and performing quantum weak coin flipping with a single photon.

A natural outlook is to extend the answers to the three general questions above to contexts other than continuous variable quantum information theory and to experimental platforms other than linear quantum optics. The first half of the thesis makes extensive use of phase space formalism. It would be interesting to see if similar results can be obtained for discrete variable quantum information theory by considering analogous discrete phase space methods.

Ultimately, the interaction of various research topics relating to quantum information theory leads to a deeper understanding of quantum advantages and enables the development of quantum technologies.

BIBLIOGRAPHY

[AA13] S. Aaronson and A. Arkhipov.
The computational complexity of linear optics.
Theory of Computing, 9:143, 2013.

[AAB+19] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando GSL Brandao, David A Buell, et al.
Quantum supremacy using a programmable superconducting processor.
Nature, 574(7779):505–510, 2019.

[ABB96] FJ Arranz, F Borondo, and RM Benito.
Distribution of zeros of the Husimi function in a realistic Hamiltonian molecular system.
Physical Review E, 54(3):2458, 1996.

[ABD+08] Scott Aaronson, Salman Beigi, Andrew Drucker, Bill Fefferman, and Peter Shor.
The power of unentanglement.
In *Computational Complexity, 2008. CCC'08. 23rd Annual IEEE Conference on*, pages 223–236. IEEE, 2008.

[AC16] Scott Aaronson and Lijie Chen.
Complexity-theoretic foundations of quantum supremacy experiments.
arXiv preprint arXiv:1612.05903, 2016.

[ACG+16] Dorit Aharonov, André Chailloux, Maor Ganz, Iordanis Kerenidis, and Loïck Magnin.
A simpler proof of the existence of quantum weak coin flipping with arbitrarily small bias.
SIAM J. Comput., 45(3):633–679, 2016.

[ADDS+09] G Adesso, F Dell'Anno, S De Siena, F Illuminati, and LAM Souza.
Optimal estimation of losses at the ultimate quantum limit with non-Gaussian states.
Physical Review A, 79(4):040305, 2009.

[AGKE15] Leandro Aolita, Christian Gogolin, Martin Kliesch, and Jens Eisert.

Reliable quantum certification of photonic state preparations.
Nature communications, 6:8498, 2015.

[AGPF18] Francesco Albarelli, Marco G Genoni, Matteo GA Paris, and Alessandro Ferraro.
Resource theory of quantum non-Gaussianity and Wigner negativity.
Physical Review A, 98(5):052350, 2018.

[AL18] Raphael A Abrahao and Austin P Lund.
Continuous-variables boson sampling: Scaling and verification.
arXiv preprint arXiv:1812.08978, 2018.

[ARL14] Gerardo Adesso, Sammy Ragy, and Antony R Lee.
Continuous variable quantum information: Gaussian states and beyond.
Open Systems & Information Dynamics, 21(01n02):1440001, 2014.

[ARV19] Atul Singh Arora, Jérémie Roland, and Chrysoula Vlachou.
Explicit quantum weak coin flipping protocols with arbitrarily small bias.
arXiv, 1911.13283, 2019.

[ARW19] Atul Singh Arora, Jérémie Roland, and Stephan Weis.
Quantum weak coin flipping.
Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, pages 205–216, 2019.

[AS65] Milton Abramowitz and Irene A Stegun.
Handbook of mathematical functions: with formulas, graphs, and mathematical tables, volume 55.
Courier Corporation, 1965.

[Aul11] Martin Aulbach.
Symmetric entanglement classes for n qubits.
arXiv preprint arXiv:1103.0271, 2011.

[AWH78] Nijenhuis Albert and S Wilf Herbert.
Combinatorial algorithms: for computers and calculators.
Academic Press, 1978.

[Bar61] Valentine Bargmann.
On a Hilbert space of analytic functions and an associated integral transform part i.
Communications on pure and applied mathematics, 14(3):187–214, 1961.

[Bar16] Alexander Barvinok.
Combinatorics and complexity of partition functions, volume 9.
Springer, 2016.

[BB84a] C. H. Bennett and G. Brassard.

Quantum cryptography: public key distribution and coin tossing.
In *Proc. IEEE International Conference on Computers, Systems and Signal Processing*, volume 1, pages 175–179, Bangalore, India, 1984.

[BB84b] Charles H Bennett and Gilles Brassard.
Quantum cryptography: public key distribution and coin tossing.
Theor. Comput. Sci., 560(12):7–11, 1984.

[BBB+11] Guido Berlín, Gilles Brassard, Félix Bussières, Nicolas Godbout, Joshua A. Slater, and Wolfgang Tittel.
Experimental loss-tolerant quantum coin flipping.
Nat. Commun., 2:561, 2011.

[BBC+93] Charles H Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K Wootters.
Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels.
Physical review letters, 70(13):1895, 1993.

[BBF+06] János A Bergou, Vladimír Bužek, Edgar Feldman, Ulrike Herzog, and Mark Hillery.
Programmable quantum-state discriminators with simple programs.
Physical Review A, 73(6):062334, 2006.

[BCF+96] Howard Barnum, Carlton M Caves, Christopher A Fuchs, Richard Jozsa, and Benjamin Schumacher.
Noncommuting mixed states cannot be broadcast.
Physical Review Letters, 76(15):2818, 1996.

[BCKD20] Mathieu Bozzio, Ulysse Chabaud, Iordanis Kerenidis, and Eleni Diamanti.
Quantum weak coin flipping with a single photon.
arXiv preprint arXiv:2002.09005, 2020.

[BCWDW01] Harry Buhrman, Richard Cleve, John Watrous, and Ronald De Wolf.
Quantum fingerprinting.
Physical Review Letters, 87(16):167902, 2001.

[BDE+19] Rui Soares Barbosa, Tom Douce, Pierre-Emmanuel Emeriau, Elham Kashefi, and Shane Mansfield.
Continuous-variable nonlocality and contextuality.
arXiv preprint arXiv:1905.08267, 2019.

[Bec00] M. Beck.
Quantum state measurement with array detectors.
Phys. Rev. Lett., 84:5748, 2000.

[Bel64] John S Bell.

On the Einstein–Podolsky–Rosen paradox.
Physics Physique Fizika, 1(3):195, 1964.

[BF99] Harry Buhrman and Lance Fortnow.
One-sided versus two-sided error in probabilistic computation.
In *Annual Symposium on Theoretical Aspects of Computer Science*, pages 100–109. Springer, 1999.

[BGC+19] Daniel J Brod, Ernesto F Galvão, Andrea Crespi, Roberto Osellame, Nicolò Spagnolo, and Fabio Sciarrino.
Photonic implementation of boson sampling: a review.
Advanced Photonics, 1(3):034001, 2019.

[BGM19] Sergey Bravyi, David Gosset, and Ramis Movassagh.
Classical algorithms for quantum mean values.
arXiv preprint arXiv:1909.11485, 2019.

[BGQ19] Andreas Björklund, Brajesh Gupt, and Nicolás Quesada.
A faster hafnian formula for complex matrices and its benchmarking on a supercomputer.
Journal of Experimental Algorithmics (JEA), 24(1):1–17, 2019.

[BGZ75] Henri Bacry, A Grossmann, and J Zak.
Proof of completeness of lattice states in the k q representation.
Physical Review B, 12(4):1118, 1975.

[BIS+18] Sergio Boixo, Sergei V Isakov, Vadim N Smelyanskiy, Ryan Babbush, Nan Ding, Zhang Jiang, Michael J Bremner, John M Martinis, and Hartmut Neven.
Characterizing quantum supremacy in near-term devices.
Nature Physics, 14(6):595, 2018.

[BJS10] M. J. Bremner, R. Josza, and D. Shepherd.
Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy.
Proc. R. Soc. A, 459:459, 2010.

[BL10] DW Berry and AI Lvovsky.
Linear-optical processing cannot increase photon efficiency.
Phys. Rev. Lett., 105(20):203601, 2010.

[Blu83] Manuel Blum.
Coin flipping by telephone a protocol for solving impossible problems.
ACM SIGACT News, 15(1):23–27, 1983.

[BMS16] Michael J. Bremner, Ashley Montanaro, and Dan J. Shepherd.

Average-case complexity versus approximate simulation of commuting quantum computations.
Phys. Rev. Lett., 117:080501, Aug 2016.

[Boa54] RP Boas.
Entire functions.
New York: Academic, 1954.

[BS99] Debabrata Biswas and Sudeshna Sinha.
Distribution of Husimi zeros in polygonal billiards.
Physical Review E, 60(1):408, 1999.

[BSBN02] Stephen D. Bartlett, Barry C. Sanders, Samuel L. Braunstein, and Kae Nemoto.
Efficient classical simulation of continuous variable quantum information processes.
Phys. Rev. Lett., 88:097904, Feb 2002.

[BvL05] Samuel L. Braunstein and Peter van Loock.
Quantum information with continuous variables.
Rev. Mod. Phys., 77:513–577, Jun 2005.

[BW92] Charles H Bennett and Stephen J Wiesner.
Communication via one-and two-particle operators on einstein-podolsky-rosen states.
Physical review letters, 69(20):2881, 1992.

[BZ78] M Boon and J Zak.
Discrete coherent states on the von Neumann lattice.
Physical Review B, 18(12):6744, 1978.

[Cai53] Eduardo R Caianiello.
On quantum field theory—i: explicit solution of Dyson's equation in electrodynamics without use of Feynman graphs.
Il Nuovo Cimento (1943-1954), 10(12):1634–1652, 1953.

[CAJ04] Anthony Chefles, Erika Andersson, and Igor Jex.
Unambiguous comparison of the states of multiple quantum systems.
Journal of Physics A: Mathematical and General, 37(29):7315, 2004.

[CC17] L. Chakhmakhchyan and N. Cerf.
Boson sampling with Gaussian measurements.
arXiv:1705.05299, 2017.

[CDG+19] Ulysse Chabaud, Tom Douce, Frédéric Grosshans, Elham Kashefi, and Damian Markham.
Building trust for continuous variable quantum states.

arXiv preprint arXiv:1905.12700, 2019.

[CDM+17] Ulysse Chabaud, Tom Douce, Damian Markham, Peter Van Loock, Elham Kashefi, and Giulia Ferrini.
Continuous-variable sampling from photon-added or photon-subtracted squeezed states.
Physical Review A, 96(6):062307, 2017.

[CDM+18] Ulysse Chabaud, Eleni Diamanti, Damian Markham, Elham Kashefi, and Antoine Joux.
Optimal quantum-programmable projective measurement with linear optics.
Physical Review A, 98(6):062318, 2018.

[CG69a] Kevin E Cahill and Roy J Glauber.
Density operators and quasiprobability distributions.
Physical Review, 177(5):1882, 1969.

[CG69b] Kevin E Cahill and Roy J Glauber.
Ordered expansions in boson amplitude operators.
Physical Review, 177(5):1857, 1969.

[CGKM20] Ulysse Chabaud, Frédéric Grosshans, Elham Kashefi, and Damian Markham.
Efficient verification of boson sampling.
arXiv preprint arXiv:2006.03520, 2020.

[CHS+15] Jacques Carolan, Christopher Harrold, Chris Sparrow, Enrique Martín-López, Nicholas J Russell, Joshua W Silverstone, Peter J Shadbolt, Nobuyuki Matsuda, Manabu Oguma, Mikitaka Itoh, et al.
Universal linear optics.
Science, 349(6249):711–716, 2015.

[CK00] Rob Clifton and Adrian Kent.
Simulating quantum mechanics by non-contextual hidden variables.
In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 456, pages 2101–2114. The Royal Society, 2000.

[CK09] André Chailloux and Iordanis Kerenidis.
Optimal quantum strong coin flipping.
50th Annual IEEE Symposium on Foundations of Computer Science, pages 527–533, 2009.

[CMG20a] Ulysse Chabaud, Damian Markham, and Frédéric Grosshans.
Classical simulation of Gaussian circuits with non-Gaussian input states.
In preparation, 2020.

[CMG20b] Ulysse Chabaud, Damian Markham, and Frédéric Grosshans.

Stellar representation of non-Gaussian quantum states.
Physical Review Letters, 124(6):063605, 2020.

[CMM99] Paul T Cochrane, Gerard J Milburn, and William J Munro.
Macroscopically distinct quantum-superposition states as a bosonic code for amplitude damping.
Physical Review A, 59(4):2631, 1999.

[CMS20] Ulysse Chabaud, Damian Markham, and Adel Sohbi.
On the possibility of quantum machine learning with linear optics.
In preparation, 2020.

[COR+16] Andrea Crespi, Roberto Osellame, Roberta Ramponi, Marco Bentivegna, Fulvio Flamini, Nicolò Spagnolo, Niko Viggianiello, Luca Innocenti, Paolo Mataloni, and Fabio Sciarrino.
Suppression law of quantum states in a 3d photonic fast Fourier transform chip.
Nature communications, 7:10469, 2016.

[Cou18] Christophe Couteau.
Spontaneous parametric down-conversion.
Contemporary Physics, 59(3):291–304, 2018.

[CR12] Matthias Christandl and Renato Renner.
Reliable quantum state tomography.
Physical Review Letters, 109(12):120403, 2012.

[Cre15] Andrea Crespi.
Suppression laws for multiparticle interference in sylvester interferometers.
Physical Review A, 91(1):013811, 2015.

[CRW+20] Ulysse Chabaud, Ganaël Roeland, Mattia Walschaers, Valentina Parigi, Frédéric Grosshans, Damian Markham, and Nicolas Treps.
Ranking non-Gaussian states with operational measurements.
In preparation, 2020.

[DB02] Miloslav Dušek and Vladimír Bužek.
Quantum-controlled measurement device for quantum-state discrimination.
Physical Review A, 66(2):022112, 2002.

[dB04] J Niel de Beaudrap.
One-qubit fingerprinting schemes.
Physical Review A, 69(2):022307, 2004.

[Die88] Dennis Dieks.
Overlap and distinguishability of quantum states.
Physics Letters A, 126(5-6):303–306, 1988.

[Dir81] Paul Adrien Maurice Dirac.
The principles of quantum mechanics.
Number 27. Oxford university press, 1981.

[DlM05] Rafael De la Madrid.
The role of the rigged hilbert space in quantum mechanics.
European journal of physics, 26(2):287, 2005.

[DMB+08] Aleksander Divochiy, Francesco Marsili, David Bitauld, Alessandro Gaggero, Roberto Leoni, Francesco Mattioli, Alexander Korneev, Vitaliy Seleznev, Nataliya Kaurova, Olga Minaeva, et al.
Superconducting nanowire photon-number-resolving detector at telecommunication wavelengths.
Nature Photonics, 2(5):302, 2008.

[DMK+17] T. Douce, D. Markham, E. Kashefi, E. Diamanti, T. Coudreau, P. Milman, P. van Loock, and G. Ferrini.
Continuous-variable instantaneous quantum computing is hard to sample.
Phys. Rev. Lett., 118:070503, 2017.

[DMK+19] Tom Douce, Damian Markham, Elham Kashefi, Peter Van Loock, and Giulia Ferrini.
Probabilistic fault-tolerant universal quantum computation and sampling problems in continuous variables.
Physical Review A, 99(1):012344, 2019.

[DPS03] G Mauro D'Ariano, Matteo GA Paris, and Massimiliano F Sacchi.
Quantum tomography.
Advances in Imaging and Electron Physics, 128:206–309, 2003.

[DY96] G Mauro D'Ariano and HP Yuen.
Impossibility of measuring the wave function of a single quantum system.
Physical review letters, 76(16):2832, 1996.

[EAO+02] Artur K Ekert, Carolina Moura Alves, Daniel K L Oi, Michał Horodecki, Paweł Horodecki, and Leong Chuan Kwek.
Direct estimations of linear and nonlinear functionals of a quantum state.
Physical review letters, 88(21):217901, 2002.

[EHW+20] Jens Eisert, Dominik Hangleiter, Nathan Walk, Ingo Roth, Damian Markham, Rhea Parekh, Ulysse Chabaud, and Elham Kashefi.
Quantum certification and benchmarking.
Nature Reviews Physics, pages 1–9, 2020.

[EP03] Jens Eisert and MB Plenio.

Introduction to the basics of entanglement theory in continuous-variable systems.
International Journal of Quantum Information, 1(04):479–506, 2003.

[ESP02] Jens Eisert, Stefan Scheel, and Martin B Plenio.
Distilling Gaussian states with Gaussian operations is impossible.
Physical review letters, 89(13):137903, 2002.

[Fey82] Richard P Feynman.
Simulating physics with computers.
International journal of theoretical physics, 21(6-7):467–488, 1982.

[FGC+13] G Ferrini, J P Gazeau, T Coudreau, C Fabre, and N Treps.
Compact Gaussian quantum computation by multi-pixel homodyne detection.
New J. Phys., 15(9):093015, 2013.

[FH16] Edward Farhi and Aram W Harrow.
Quantum supremacy through the quantum approximate optimization algorithm.
arXiv preprint arXiv:1602.07674, 2016.

[Fiu02] Jaromír Fiurášek.
Gaussian transformations and distillation of entangled Gaussian states.
Physical review letters, 89(13):137904, 2002.

[FK17] Joseph F Fitzsimons and Elham Kashefi.
Unconditionally verifiable blind quantum computation.
Physical Review A, 96(1):012303, 2017.

[FMJ11] Radim Filip and Ladislav Mišta Jr.
Detecting quantum states with a positive Wigner function beyond mixtures of Gaussian states.
Physical Review Letters, 106(20):200401, 2011.

[FOP05] Alessandro Ferraro, Stefano Olivares, and Matteo GA Paris.
Gaussian states in continuous variable quantum information.
arXiv preprint quant-ph/0503237, 2005.

[FRS+20] Marco Fanizza, Matteo Rosati, Michalis Skotiniotis, John Calsamiglia, and Vittorio Giovannetti.
Beyond the swap test: optimal estimation of quantum state overlap.
Physical Review Letters, 124(6):060503, 2020.

[FVDG99] Christopher A Fuchs and Jeroen Van De Graaf.
Cryptographic distinguishability measures for quantum-mechanical states.
IEEE Transactions on Information Theory, 45(4):1216–1227, 1999.

[GC02] Géza Giedke and J Ignacio Cirac.
Characterization of Gaussian operations and distillation of Gaussian states.

Physical Review A, 66(3):032316, 2002.

[GCP07] Leuchs Gerd, Nicolas J Cerf, and Eugene S Polzik.
Quantum information with continuous variables of atoms and light.
Imperial College Press, London, 2007.

[GECP13] Juan Carlos Garcia-Escartin and Pedro Chamorro-Posada.
Swap test and Hong–Ou–Mandel effect are equivalent.
Physical Review A, 87(5):052330, 2013.

[GG02] Frédéric Grosshans and Philippe Grangier.
Continuous variable quantum cryptography using coherent states.
Physical review letters, 88(5):057902, 2002.

[GG19] Christos N. Gagatsos and Saikat Guha.
Efficient representation of Gaussian states for multimode non-Gaussian quantum state engineering via subtraction of arbitrary number of photons.
Phys. Rev. A, 99:053816, May 2019.

[GKK19] Alexandru Gheorghiu, Theodoros Kapourniotis, and Elham Kashefi.
Verification of quantum computation: an overview of existing approaches.
Theory of Computing Systems, 4:715–808, 2019.

[GKP01] D. Gottesman, A. Kitaev, and J. Preskill.
Encoding a qubit in an oscillator.
Phys. Rev. A, 64:012310, 2001.

[Gla63] Roy J Glauber.
Coherent and incoherent states of the radiation field.
Physical Review, 131(6):2766, 1963.

[GLM11] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone.
Advances in quantum metrology.
Nature photonics, 5(4):222, 2011.

[Gou72] Henry Wadsworth Gould.
Combinatorial Identities: a standardized set of tables listing 500 binomial coefficient summations.
Morgantown, W Va, 1972.

[GPB07] Marco G Genoni, Matteo GA Paris, and Konrad Banaszek.
Measure of the non-Gaussian character of a quantum state.
Physical Review A, 76(4):042327, 2007.

[GPFC+04] Raul García-Patrón, Jaromır Fiurášek, Nicolas J Cerf, Jérôme Wenger, Rosa Tualle-Brouri, and Ph Grangier.
Proposal for a loophole-free Bell test using homodyne detection.

Physical Review Letters, 93(13):130409, 2004.

[GPT+13] Marco G Genoni, Mattia L Palma, Tommaso Tufarelli, Stefano Olivares, MS Kim, and Matteo GA Paris.
Detecting quantum non-Gaussianity via the Wigner function.
Physical Review A, 87(6):062104, 2013.

[Gro98] Lov K. Grover.
Quantum computers can search rapidly by using almost any transformation.
Phys. Rev. Lett., 80:4329–4332, 1998.

[GS07] Shohini Ghose and Barry C Sanders.
Non-Gaussian ancilla states for continuous variable quantum computation via Gaussian maps.
Journal of Modern Optics, 54(6):855–869, 2007.

[Gur05] Leonid Gurvits.
On the complexity of mixed discriminants and related problems.
In *International Symposium on Mathematical Foundations of Computer Science*, pages 447–458. Springer, 2005.

[Had09] R. H. Hadfield.
Single-photon detectors for optical quantum information applications.
Nat. Photonics, 3:696–705, 2009.

[Har06] Michael Hardy.
Combinatorics of partial derivatives.
the electronic journal of combinatorics, pages R1–R1, 2006.

[Har13] Aram W Harrow.
The church of the symmetric subspace.
arXiv preprint arXiv:1308.6595, 2013.

[HBR07] S Haroche, M Brune, and J-M Raimond.
Measuring the photon number parity in a cavity: from light quantum jumps to the tomography of non-classical field states.
Journal of Modern Optics, 54(13-15):2101–2114, 2007.

[HCT+19] Vojtěch Havlíček, Antonio D Córcoles, Kristan Temme, Aram W Harrow, Abhinav Kandala, Jerry M Chow, and Jay M Gambetta.
Supervised learning with quantum-enhanced feature spaces.
Nature, 567(7747):209, 2019.

[Hei85] Werner Heisenberg.
Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik.

In *Original Scientific Papers Wissenschaftliche Originalarbeiten*, pages 478–504. Springer, 1985.

[HGT+14] Catherine Hughes, Marco G Genoni, Tommaso Tufarelli, Matteo GA Paris, and MS Kim.
Quantum non-Gaussianity witnesses in phase space.
Physical Review A, 90(1):013810, 2014.

[HKEG19] Dominik Hangleiter, Martin Kliesch, Jens Eisert, and Christian Gogolin.
Sample complexity of device-independently certified "quantum supremacy".
Physical review letters, 122(21):210502, 2019.

[HKS+16] Craig S. Hamilton, Regina Kruse, Linda Sansoni, Sonja Barkhofen, Christine Silberhorn, and Igor Jex.
Gaussian Boson Sampling.
arXiv:1612.01199v1, 2016.

[HKSE16] D. Hangleiter, M. Kliesch, M. Schwarz, and J. Eisert.
Direct certification of a class of quantum simulations.
arXiv:1602.00703, 2016.

[HM13] Aram W Harrow and Ashley Montanaro.
Testing product states, quantum Merlin–Arthur games and tensor optimization.
Journal of the ACM (JACM), 60(1):3, 2013.

[HM17] Aram W Harrow and Ashley Montanaro.
Quantum computational supremacy.
Nature, 549(7671):203, 2017.

[Hoe63] Wassily Hoeffding.
Probability inequalities for sums of bounded random variables.
Journal of the American statistical association, 58(301):13–30, 1963.

[HOM87] Chong-Ki Hong, Zhe-Yu Ou, and Leonard Mandel.
Measurement of subpicosecond time intervals between two photons by interference.
Physical review letters, 59(18):2044, 1987.

[Hud74] Robin L Hudson.
When is the Wigner quasi-probability density non-negative?
Reports on Mathematical Physics, 6(2):249–252, 1974.

[Hus40] Kôdi Husimi.
Some formal properties of the density matrix.
Proceedings of the Physico-Mathematical Society of Japan. 3rd Series, 22(4):264–314, 1940.

[HW11] Esther Hänggi and Jürg Wullschleger.

Tight bounds for classical and quantum coin flipping.
Proceedings of TCC, pages 468–485, 2011.

[JKJL+13] Paul Jouguet, Sébastien Kunz-Jacques, Anthony Leverrier, Philippe Grangier, and Eleni Diamanti.
Experimental demonstration of long-distance continuous-variable quantum key distribution.
Nature Photon., 7(5):378–381, 2013.

[KCK+20] Niraj Kumar, Ulysse Chabaud, Elham Kashefi, Damian Markham, and Eleni Diamanti.
Quantum information processing with coherent states: optimal and programmable toolkit schemes.
In preparation, 2020.

[KD19] Theodoros Kapourniotis and Animesh Datta.
Nonadaptive fault-tolerant verification of quantum supremacy with noise.
Quantum, 3:164, 2019.

[KDK17] Niraj Kumar, Eleni Diamanti, and Iordanis Kerenidis.
Efficient quantum communications with coherent state fingerprints over multiple channels.
Physical Review A, 95(3):032337, 2017.

[KHS+19] Regina Kruse, Craig S Hamilton, Linda Sansoni, Sonja Barkhofen, Christine Silberhorn, and Igor Jex.
Detailed study of Gaussian boson sampling.
Physical Review A, 100(3):032326, 2019.

[KKD19] Niraj Kumar, Iordanis Kerenidis, and Eleni Diamanti.
Experimental demonstration of quantum advantage for one-way communication complexity surpassing best-known classical protocol.
Nature Communications, 10(1):1–10, 2019.

[KLM01] Emanuel Knill, Raymond Laflamme, and Gerald J Milburn.
A scheme for efficient quantum computation with linear optics.
Nature, 409(6816):46–52, 2001.

[KMW97] HJ Korsch, C Müller, and H Wiescher.
On the zeros of the Husimi distribution.
Journal of Physics A: Mathematical and General, 30(20):L677, 1997.

[KN04] Iordanis Kerenidis and A. Nayak.
Weak coin flipping with small bias.
Inf. Proc. Lett., 89:131–135, 2004.

[KNY08] Masaru Kada, Harumichi Nishimura, and Tomoyuki Yamakami.
The efficiency of quantum identity testing of multiple states.
Journal of Physics A: Mathematical and Theoretical, 41(39):395309, 2008.

[KŻ04] Anatole Kenfack and Karol Życzkowski.
Negativity of the Wigner function as an indicator of non-classicality.
Journal of Optics B: Quantum and Semiclassical Optics, 6(10):396, 2004.

[Lau83] Clemens Lautemann.
BPP and the polynomial hierarchy.
Information Processing Letters, 17(4):215–217, 1983.

[LB95] N Lütkenhaus and Stephen M Barnett.
Nonclassical effects in phase space.
Physical Review A, 51(4):3340, 1995.

[LDT+18] Nana Liu, Tommaso F Demarie, Si-Hui Tan, Leandro Aolita, and Joseph F Fitzsimons.
Client-friendly continuous-variable blind and verifiable quantum computing.
arXiv:1806.09137, 2018.

[LGPRC13] Anthony Leverrier, Raúl García-Patrón, Renato Renner, and Nicolas J Cerf.
Security of continuous-variable quantum key distribution against general attacks.
Physical review letters, 110(3):030502, 2013.

[LLRK+14] A. P. Lund, A. Laing, S. Rahimi-Keshari, T. Rudolph, J. L. O'Brien, and T. C. Ralph.
Boson Sampling from a Gaussian state.
Phys. Rev. Lett., 113:100502, Sep 2014.

[LMR13] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost.
Quantum algorithms for supervised and unsupervised machine learning.
arXiv preprint arXiv:1307.0411, 2013.

[LR09] Alexander I Lvovsky and Michael G Raymer.
Continuous-variable optical quantum-state tomography.
Reviews of Modern Physics, 81(1):299, 2009.

[LRKR17] A. P. Lund, S. Rahimi-Keshari, and T. C. Ralph.
Exact Boson Sampling using Gaussian continuous variable measurements.
Phys. Rev. A, 96:022301, 2017.

[LRW+18] Ludovico Lami, Bartosz Regula, Xin Wang, Rosanna Nichols, Andreas Winter, and Gerardo Adesso.
Gaussian quantum resource theories.
Physical Review A, 98(2):022335, 2018.

[LSH+18] Lukáš Lachman, Ivo Straka, Josef Hloušek, Miroslav Ježek, and Radim Filip.

Faithful hierarchy of genuine n-photon quantum non-Gaussian light.
arXiv preprint arXiv:1810.02546, 2018.

[LV90] P Leboeuf and André Voros.
Chaos-revealing multiplicative representation of quantum eigenstates.
Journal of Physics A: Mathematical and General, 23(10):1765, 1990.

[MA10] Bellini Marco and Zavatta Alessandro.
Manipulating light states by single-photon addition and subtraction.
In *Progress in Optics*, volume 55, pages 41–83. Elsevier, 2010.

[Mah18] Urmila Mahadev.
Classical verification of quantum computations.
arXiv preprint arXiv:1804.01082, 2018.

[Man01] Steven M Manson.
Simplifying complexity: a review of complexity theory.
Geoforum, 32(3):405–414, 2001.

[MBH+13] Olivier Morin, Jean-Daniel Bancal, Melvyn Ho, Pavel Sekatski, Virginia D'Auria, Nicolas Gisin, Julien Laurat, and Nicolas Sangouard.
Witnessing trustworthy single-photon entanglement with local homodyne measurements.
Phys. Rev. Lett., 110:130401, 2013.

[MdW13] Ashley Montanaro and Ronald de Wolf.
A survey of quantum property testing.
arXiv preprint arXiv:1310.2035, 2013.

[ME12] Andrea Mari and Jens Eisert.
Positive wigner functions render classical simulation of quantum computation efficient.
Physical review letters, 109(23):230503, 2012.

[Meh67] CL Mehta.
Diagonal coherent-state representation of quantum operators.
Physical Review Letters, 18(18):752, 1967.

[Mey99] David A Meyer.
Finite precision measurement nullifies the Kochen-Specker theorem.
Physical Review Letters, 83(19):3751, 1999.

[MF09] David Menzies and Radim Filip.
Gaussian-optimized preparation of non-Gaussian pure states.
Physical Review A, 79(1):012313, 2009.

[MFF14] T. Morimae, K. Fujii, and J. F. Fitzsimons.

Hardness of classically simulating the one-clean-qubit model.
Phys. Rev. Lett., 112:130502, 2014.

[Mil20] Carl A. Miller.
The impossibility of efficient quantum weak coin flipping.
arXiv preprint quant-ph/1909.10103v2, 2020.

[MKB05] Florian Mintert, Marek Kuś, and Andreas Buchleitner.
Concurrence of mixed multipartite quantum states.
Physical Review Letters, 95(26):260502, 2005.

[Moc04] Carlos Mochon.
Quantum weak coin flipping with bias of 0.192.
45th Symposium on Foundations of Computer Science, pages CALT–68–2486, 2004.

[Moc05] C. Mochon.
Large family of quantum weak coin-flipping protocols.
Phys. Rev. A, 72(2):022341, 2005.

[Moc07] Carlos Mochon.
Quantum weak coin flipping with arbitrarily small bias.
arXiv, 0711.4114, 2007.

[MPKK17] Daniel Mills, Anna Pappa, Theodoros Kapourniotis, and Elham Kashefi.
Information theoretically secure hypothesis test for temporally unstructured quantum computation.
arXiv preprint arXiv:1704.01998, 2017.

[MTVUZ05] G. Molina-Terriza, A. Vaziri, R. Ursin, and A. Zeilinger.
Experimental quantum coin tossing.
Phys. Rev. Lett., 94:040501, 2005.

[NC97] Michael A Nielsen and Isaac L Chuang.
Programmable quantum gate arrays.
Physical Review Letters, 79(2):321, 1997.

[NC02] Michael A Nielsen and Isaac Chuang.
Quantum computation and quantum information, 2002.

[NFC09] Julien Niset, Jaromír Fiurášek, and Nicolas J Cerf.
No-go theorem for Gaussian quantum error correction.
Physical review letters, 102(12):120501, 2009.

[NT97] Michael Martin Nieto and D Rodney Truax.
Holstein–Primakoff/Bogoliubov transformations and the multiboson system.
Fortschritte der Physik/Progress of Physics, 45(2):145–156, 1997.

[OSM+15] J. P. Olson, K. P. Seshadreesan, K. R. Motes, P. P. Rohde, and J. P. Dowling.
Sampling arbitrary photon-added or photon-subtracted squeezed states is in the same complexity class as boson sampling.
Phys. Rev. A, 91:022317, 2015.

[PBG20] Hakop Pashayan, Stephen D Bartlett, and David Gross.
From estimation of quantum probabilities to simulation of quantum circuits.
Quantum, 4:223, 2020.

[Per71] Askol'd Mikhailovich Perelomov.
On the completeness of a system of coherent states.
Theoretical and Mathematical Physics, 6(2):156–164, 1971.

[Per12] Jerome K Percus.
Combinatorial methods, volume 4.
Springer Science & Business Media, 2012.

[PJL+14] Anna Pappa, Paul Jouguet, Thomas Lawson, André Chailloux, Matthieu Legré, Patrick Trinkler, Iordanis Kerenidis, and Eleni Diamanti.
Experimental plug and play quantum coin flipping.
Nat. Commun., 5:3717, 2014.

[Pre98a] John Preskill.
Fault-tolerant quantum computation.
In *Introduction to quantum computation and information*, pages 213–269. World Scientific, 1998.

[Pre98b] John Preskill.
Lecture notes for a course on quantum computation.
Unpublished. Available at http://www. theory. caltech. edu/people/preskill/ph229, 1999, 1998.

[Pre18] John Preskill.
Quantum computing in the NISQ era and beyond.
Quantum, 2:79, 2018.

[PT04] Asher Peres and Daniel R Terno.
Quantum information and relativity theory.
Reviews of Modern Physics, 76(1):93, 2004.

[PWB15] Hakop Pashayan, Joel J Wallman, and Stephen D Bartlett.
Estimating outcome probabilities of quantum circuits using quasiprobabilities.
Physical review letters, 115(7):070501, 2015.

[QA20] Nicolás Quesada and Juan Miguel Arrazola.

Exact simulation of Gaussian Boson Sampling in polynomial space and exponential time.
Physical Review Research, 2(2):023005, 2020.

[Que19] Nicolás Quesada.
Franck–Condon factors by counting perfect matchings of graphs with loops.
The Journal of chemical physics, 150(16):164113, 2019.

[RBCH03] Marián Roško, Vladimír Bužek, Paul Robert Chouha, and Mark Hillery.
Generalized measurements via a programmable quantum processor.
Physical Review A, 68(6):062302, 2003.

[RC08] Renato Renner and J Ignacio Cirac.
A de Finetti representation theorem for infinite dimensional quantum systems and applications to quantum cryptography.
arXiv preprint arXiv:0809.2243, 2008.

[RC09] Renato Renner and J Ignacio Cirac.
de Finetti representation theorem for infinite-dimensional quantum systems and applications to quantum cryptography.
Physical review letters, 102(11):110504, 2009.

[Reg09] Oded Regev.
On lattices, learning with errors, random linear codes, and cryptography.
Journal of the ACM (JACM), 56(6):1–40, 2009.

[Ren08] Renato Renner.
Security of quantum key distribution.
International Journal of Quantum Information, 6(01):1–127, 2008.

[Rie57] Bernhard Riemann.
Theorie der Abel'schen functionen.
Georg Reimer Berlin, 1857.

[RJD+17] Y.-S. Ra, C. Jacquard, A. Dufour, C. Fabre, and C. Treps.
Tomography of a mode-tunable coherent single-photon subtractor.
arXiv:1702.02082, 2017.

[RM11] P Ribeiro and R Mosseri.
Entanglement in the symmetric sector of n qubits.
Physical review letters, 106(18):180502, 2011.

[Roo85] PG Rooney.
Further inequalities for generalized Laguerre polynomials.
CR Math. Rep. Acad. Sci. Canada, 7:273–275, 1985.

[Roy77] Antoine Royer.

Wigner function as the expectation value of a parity operator.
Physical Review A, 15(2):449, 1977.

[RZBB94] Michael Reck, Anton Zeilinger, Herbert J Bernstein, and Philip Bertani.
Experimental realization of any discrete unitary operator.
Physical review letters, 73(1):58, 1994.

[SB07] Nathan Sidoli and J Lennart Berggren.
The Arabic version of Ptolemy's 'planisphere or flattening the surface of the sphere': text, translation, commentary.
Sciamvs, 8:37, 2007.

[SB09] D. Shepherd and M. J. Bremner.
Temporally unstructured quantum computation.
Proc. R. Soc. A, 465:1413, 2009.

[SC83] Francisco Soto and Pierre Claverie.
When is the Wigner function of multidimensional systems nonnegative?
Journal of Mathematical Physics, 24(1):97–100, 1983.

[SEMC13] Francisco Soto-Eguibar and Héctor Manuel Moya-Cessa.
Harmonic oscillator position eigenstates via application of an operator on the vacuum.
Revista mexicana de física E, 59(2):122–127, 2013.

[Ser74] Robert J Serfling.
Probability inequalities for the sum in sampling without replacement.
The Annals of Statistics, pages 39–48, 1974.

[Sho94] Peter W Shor.
Algorithms for quantum computation: discrete logarithms and factoring.
In *Proceedings 35th annual symposium on foundations of computer science*, pages 124–134. Ieee, 1994.

[SJZ+18] Lucas Schweickert, Klaus D Jöns, Katharina D Zeuner, Saimon Filipe Covre da Silva, Huiying Huang, Thomas Lettner, Marcus Reindl, Julien Zichi, Rinaldo Trotta, Armando Rastelli, et al.
On-demand generation of background-free single photons from a solid-state source.
Applied Physics Letters, 112(9):093106, 2018.

[SK19] Maria Schuld and Nathan Killoran.
Quantum machine learning in feature Hilbert spaces.
Physical review letters, 122(4):040504, 2019.

[SLH+18] Ivo Straka, Lukáš Lachman, Josef Hloušek, Martina Miková, Michal Mičuda, Miroslav Ježek, and Radim Filip.

Quantum non-Gaussian multiphoton light.
npj Quantum Information, 4(1):4, 2018.

[SM63] Irving Ezra Segal and George W Mackey.
Mathematical problems of relativistic physics, volume 2.
American Mathematical Soc., 1963.

[SMS19] Daiqin Su, Casey R Myers, and Krishna Kumar Sabapathy.
Conversion of Gaussian states to non-Gaussian states using photon number-resolving detectors.
arXiv preprint arXiv:1902.02323, 2019.

[SPL+14] Luyan Sun, Andrei Petrenko, Zaki Leghtas, Brian Vlastakis, Gerhard Kirchmair, KM Sliwa, Aniruth Narla, Michael Hatridge, Shyam Shankar, Jacob Blumoff, et al.
Tracking photon jumps with repeated quantum non-demolition parity measurements.
Nature, 511(7510):444, 2014.

[SR02] R. W. Spekkens and Terry Rudolph.
A quantum protocol for cheat-sensitive weak coin flipping.
Phys. Rev. Lett., 89:227901, 2002.

[SS83] Rodica Simion and Frank W Schmidt.
On (+ 1,- 1)-matrices with vanishing permanent.
Discrete Mathematics, 46(1):107–108, 1983.

[SS10] Elias M Stein and Rami Shakarchi.
Complex analysis, volume 2.
Princeton University Press, 2010.

[Sto85] Larry Stockmeyer.
On approximation algorithms for #P.
SIAM Journal on Computing, 14(4):849–861, 1985.

[Sud63] ECG Sudarshan.
Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams.
Physical Review Letters, 10(7):277, 1963.

[SW18] Krishna Kumar Sabapathy and Christian Weedbrook.
On states as resource units for universal quantum computation with photonic architectures.
Physical Review A, 97(6):062315, 2018.

[SZP+07] Michal Sedlák, Mário Ziman, Ondřej Přibyla, Vladimír Bužek, and Mark Hillery.

Unambiguous identification of coherent states: searching a quantum database.
Physical Review A, 76(2):022326, 2007.

[TBMS20] Ilan Tzitrin, J Eli Bourassa, Nicolas C Menicucci, and Krishna Kumar Sabapathy.
Progress towards practical qubit computation using approximate gottesman-kitaev-preskill codes.
Physical Review A, 101(3):032315, 2020.

[TD02] Barbara M Terhal and David P DiVincenzo.
Classical simulation of noninteracting-fermion quantum circuits.
Physical Review A, 65(3):032325, 2002.

[TH00] Barbara M Terhal and Paweł Horodecki.
Schmidt number for density matrices.
Physical Review A, 61(4):040301, 2000.

[TM18] Yuki Takeuchi and Tomoyuki Morimae.
Verification of many-qubit states.
Physical Review X, 8(2):021060, 2018.

[TMJ+17] Yong Siah Teo, Christian R Muller, Hyunseok Jeong, Zdenek Hradil, Jaroslav Rehacek, and Luis L Sanchez-Soto.
When heterodyning beats homodyning: an assessment with quadrature moments.
arXiv preprint arXiv:1701.07539, 2017.

[TMM+19] Yuki Takeuchi, Atul Mantri, Tomoyuki Morimae, Akihiro Mizutani, and Joseph F Fitzsimons.
Resource-efficient verification of quantum computing using Serfling's bound.
npj Quantum Information, 5:27, 2019.

[Tod91] Seinosuke Toda.
PP is as hard as the polynomial-time hierarchy.
SIAM Journal on Computing, 20(5):865–877, 1991.

[TV95] J-M Tualle and A Voros.
Normal modes of billiards portrayed in the stellar (or nodal) representation.
Chaos, Solitons & Fractals, 5(7):1085–1102, 1995.

[TZ18] Ryuji Takagi and Quntao Zhuang.
Convex resource theory of non-Gaussianity.
Physical Review A, 97(6):062337, 2018.

[Val79] Leslie G Valiant.
The complexity of computing the permanent.
Theoretical computer science, 8(2):189–201, 1979.

[Val02] Leslie G Valiant.

Quantum circuits that can be simulated classically in polynomial time.
SIAM Journal on Computing, 31(4):1229–1254, 2002.

[VC00] Guifre Vidal and J Ignacio Cirac.
Storage of quantum dynamics on quantum states: a quasi-perfect programmable quantum gate.
arXiv preprint quant-ph/0012067, 2000.

[Vid18] Thomas Vidick.
`http://users.cms.caltech.edu/~vidick/verification_bulletin.pdf`, 2018.

[VKL+13] Brian Vlastakis, Gerhard Kirchmair, Zaki Leghtas, Simon E Nigg, Luigi Frunzio, Steven M Girvin, Mazyar Mirrahimi, Michel H Devoret, and Robert J Schoelkopf.
Deterministically encoding quantum information using 100-photon Schrödinger cat states.
Science, 342(6158):607–610, 2013.

[Vou06] A Vourdas.
Analytic representations in quantum mechanics.
Journal of Physics A: Mathematical and General, 39(7):R65, 2006.

[Wer89] Reinhard F Werner.
Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model.
Physical Review A, 40(8):4277, 1989.

[WHG+03] Jérôme Wenger, Mohammad Hafezi, Frédéric Grosshans, Rosa Tualle-Brouri, and Philippe Grangier.
Maximal violation of Bell inequalities using continuous-variable measurements.
Physical Review A, 67(1):012105, 2003.

[Wig97] Eugene Paul Wigner.
On the quantum correction for thermodynamic equilibrium.
In *Part I: Physical Chemistry. Part II: Solid State Physics*, pages 110–120. Springer, 1997.

[Win14] Andreas Winter.
What does an experimental test of quantum contextuality prove or disprove?
Journal of Physics A: Mathematical and Theoretical, 47(42):424031, 2014.

[WM07] Daniel F Walls and Gerard J Milburn.
Quantum optics.
Springer Science & Business Media, 2007.

[WPGP+12] Christian Weedbrook, Stefano Pirandola, Raúl García-Patrón, Nicolas J Cerf, Timothy C Ralph, Jeffrey H Shapiro, and Seth Lloyd.
Gaussian quantum information.
Reviews of Modern Physics, 84(2):621, 2012.

[WQD+19] Hui Wang, Jian Qin, Xing Ding, Ming-Cheng Chen, Si Chen, Xiang You, Yu-Ming He, Xiao Jiang, L You, Z Wang, et al.
Boson sampling with 20 input photons and a 60-mode interferometer in a 1 0 14-dimensional hilbert space.
Physical review letters, 123(25):250503, 2019.

[WRD+06] SP Walborn, PH Souto Ribeiro, L Davidovich, F Mintert, and A Buchleitner.
Experimental determination of entanglement with a single measurement.
Nature, 440(7087):1022, 2006.

[WSPT18] Mattia Walschaers, Supratik Sarkar, Valentina Parigi, and Nicolas Treps.
Tailoring non-Gaussian continuous-variable graph states.
Physical review letters, 121(22):220501, 2018.

[Wün98] Alfred Wünsche.
Laguerre 2D-functions and their application in quantum optics.
Journal of Physics A: Mathematical and General, 31(40):8267, 1998.

[Wys17] Walter Wyss.
Two non-commutative binomial theorems.
arXiv preprint arXiv:1707.03861, 2017.

[WZ82] William K Wootters and Wojciech H Zurek.
A single quantum cannot be cloned.
Nature, 299(5886):802–803, 1982.

[YBT+18] Benjamin Yadin, Felix C Binder, Jayne Thompson, Varun Narasimhachar, Mile Gu, and MS Kim.
Operational resource theory of continuous-variable nonclassicality.
Physical Review X, 8(4):041038, 2018.

[YUA+13a] Shota Yokoyama, Ryuji Ukai, Seiji C Armstrong, Chanond Sornphiphatphong, Toshiyuki Kaji, Shigenari Suzuki, Jun-ichi Yoshikawa, Hidehiro Yonezawa, Nicolas C Menicucci, and Akira Furusawa.
Optical generation of ultra-large-scale continuous-variable cluster states.
Nature Photonics, 7:982, 2013.

[YUA+13b] Shota Yokoyama, Ryuji Ukai, Seiji C Armstrong, Chanond Sornphiphatphong, Toshiyuki Kaji, Shigenari Suzuki, Jun-ichi Yoshikawa, Hidehiro Yonezawa, Nicolas C Menicucci, and Akira Furusawa.

Ultra-large-scale continuous-variable cluster states multiplexed in the time domain.
Nature Photonics, 7(12):982, 2013.

[YYK+16] J-i. Yoshikawa, S. Yokoyama, T. Kaji, C. Sorphiphatphong, Y. Shiozawa, K. Makino, and A. Furusawa.
Generation of one-million mode continuous-variable cluster state by unlimited time-domain multiplexing.
arXiv:1606.06688, 2016.

[ZB05] Mário Ziman and Vladimír Bužek.
Realization of positive-operator-valued measures using measurement-assisted programmable quantum processors.
Physical Review A, 72(2):022343, 2005.

[ZSS18] Quntao Zhuang, Peter W Shor, and Jeffrey H Shapiro.
Resource theory of non-Gaussian operations.
Physical Review A, 97(5):052317, 2018.

[ZVB04] Alessandro Zavatta, Silvia Viciani, and Marco Bellini.
Quantum-to-classical transition with single-photon-added coherent states of light.
science, 306(5696):660–662, 2004.