

Imperial College London  
Frontiers in Quantum Technology - Doctoral Training Programme  
Department of Physics

**Nonclassicality detection and  
communication bounds  
in quantum networks**

Luca Rigovacca

August 2017

Thesis submitted in partial fulfilment of the requirements  
for the degree of Ph.D. in Physics of Imperial College London



# Abstract

Quantum information investigates the possibility of enhancing our ability to process and transmit information by directly exploiting quantum mechanical laws. When searching for improvement opportunities, one typically starts by assessing the range of outcomes classically attainable, and then investigates to what extent control over the quantum features of the system could be helpful, as well as the best performance that could be achieved. In this thesis we provide examples of these aspects, in linear optics, quantum metrology, and quantum communication.

We start by providing a criterion able to certify whether the outcome of a linear optical evolution cannot be explained by the classical wave-like theory of light. We do so by identifying a tight lower bound on the amount of correlations that could be detected among output intensities, when classical electrodynamics theory is used to describe the fields.

Rather than simply detecting nonclassicality, we then focus on its quantification. In particular, we consider the characterisation of the amount of squeezing encoded on selected quantum probes by an unknown external device, without prior information on the direction of application. We identify the single-mode Gaussian probes leading to the largest average precision in noiseless and noisy conditions, and discuss the advantages arising from the use of correlated two-mode probes.

Finally, we improve current bounds on the ultimate performance attainable in a quantum communication scenario. Specifically, we bound the number of maximally entangled qubits, or private bits, shared by two parties after a communication protocol over a quantum network, without restrictions on their classical communication. As in previous investigations, our approach is based on the evaluation of the maximum amount of entanglement that could be generated by the channels in the network, but it includes the possibility of changing entanglement measure on a channel-by-channel basis. Examples where this is advantageous are discussed.



## Declaration of Originality

I hereby declare that this thesis is the result of my own original research, carried out during the last three years as PhD student at Imperial College London, under the supervision of Prof. M. S. Kim. All work of others has been appropriately referenced and acknowledged.

## Declaration of Copyright

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

## List of Publications

The results discussed in this thesis are based upon the following publications and pre-prints:

- **L. Rigovacca**, C. Di Franco, B.J. Metcalf, I.A. Walmsley, and M.S. Kim, *Nonclassicality criteria in multiport interferometry*, Physical Review Letters 117, 213602 (2016).
- **L. Rigovacca**, A. Farace, L. A. M. Souza, A. De Pasquale, V. Giovannetti, G. Adesso, *Versatile Gaussian probes for squeezing estimation*, Physical Review A 95, 052331 (2017).
- **L. Rigovacca**, G. Kato, S. Bäuml, M. S. Kim, W. J. Munro, K. Azuma, *Versatile relative entropy bounds for quantum networks*, arXiv:1707.05543 [quant-ph].

Other works, non included in this thesis, are:

- **L. Rigovacca** and C. Di Franco, *Two-walker discrete-time quantum walks on the line with percolation*, Scientific Reports 6, 22052 (2016)
- **L. Rigovacca**, A. Farace, A. De Pasquale and V. Giovannetti, *Gaussian Discriminating Strength*, Physical Review A 92, 042331 (2015)



## Acknowledgements

I would first like to thank my supervisor Prof. Myungshik Kim for taking me on as a student, for always being supporting and encouraging, and for leaving me the freedom to follow my interests. I am also indebted to all the researchers I had the pleasure of collaborating with during these years. In particular, I would like to express my gratitude to Vittorio Giovannetti, Antonella De Pasquale, Alessandro Farace, Gerardo Adesso, Carlo Di Franco, Bill Munro and Koji Azuma: discussing with you has been very inspiring, and I am grateful for the opportunity I had to work with all of you.

A sincere thank you to all the people in the CQD theory group, for the friendly and relaxed atmosphere that always characterised our office. In particular, I would like to thank Doug Plato for all his social efforts and for giving me the opportunity to practice my English, and the Italian team (Federico, Ludovico, Carlo, Matteo, Tommaso) for not letting me forget my native language. I am also grateful to the Theoretical Quantum Physics Research Group of NTT BRL, for turning my visit to Japan into such a wonderful experience.

These years would not have been the same without the company of friends, close and far. A special thank you to Luca, for enduring my company and for making home feel like home, to Bea for introducing me to bouldering and for her pizza, and to Riccardo, Giovanna, Daniele and Daniele for our long-standing friendship.

I would also like to thank my parents and my (not so little any more) sister Ilaria, for having always been the best family I could have hoped for. Despite the distance and my absence from home you never stopped believing in me and supporting all my choices, and I am grateful for that.

Last but not least, thank you Giovanna for your love and your patience: among the highs and lows of this PhD you have always been at my side, close despite the distance between us.





# Contents

|                                                                      |           |
|----------------------------------------------------------------------|-----------|
| <b>Abstract</b>                                                      | <b>3</b>  |
| <b>Author's declarations and list of publications</b>                | <b>3</b>  |
| <b>Acknowledgements</b>                                              | <b>6</b>  |
| <b>1 Thesis Overview</b>                                             | <b>15</b> |
| 1.1 Quantum Information: history and perspectives . . . . .          | 15        |
| 1.2 Summary of results . . . . .                                     | 17        |
| 1.3 Required background and comments on notation . . . . .           | 19        |
| <b>2 Nonclassicality detection in multiport interferometry</b>       | <b>21</b> |
| 2.1 Introduction . . . . .                                           | 21        |
| 2.2 Quantum and classical optics . . . . .                           | 23        |
| 2.2.1 Quantum states of light . . . . .                              | 23        |
| 2.2.2 Comparison with classical fields . . . . .                     | 25        |
| 2.2.3 Intensity measurements . . . . .                               | 27        |
| 2.3 Hong-Ou-Mandel effect . . . . .                                  | 29        |
| 2.3.1 Quantum description . . . . .                                  | 29        |
| 2.3.2 Counterpart in classical optics . . . . .                      | 35        |
| 2.4 Multimode generalisation . . . . .                               | 39        |
| 2.4.1 Tight classical lower bound . . . . .                          | 41        |
| 2.4.2 Quantum violation of the classical threshold . . . . .         | 49        |
| 2.4.3 Two practical applications . . . . .                           | 56        |
| 2.4.4 Weakening the assumptions on the classical framework . . . . . | 60        |
| 2.5 Discussion . . . . .                                             | 64        |
| 2.5.1 Can we further weaken the classical assumptions? . . . . .     | 64        |
| 2.5.2 Relation to other forms of nonclassicality . . . . .           | 65        |
| 2.6 Conclusions . . . . .                                            | 67        |
| <b>3 Versatile probes for squeezing estimation</b>                   | <b>69</b> |
| 3.1 Introduction . . . . .                                           | 69        |

|          |                                                                       |            |
|----------|-----------------------------------------------------------------------|------------|
| 3.1.1    | From black-box metrology to squeezing estimation . . . . .            | 70         |
| 3.2      | Preliminary notions . . . . .                                         | 72         |
| 3.2.1    | Quantum estimation theory . . . . .                                   | 73         |
| 3.2.2    | Gaussian states and operations . . . . .                              | 75         |
| 3.2.3    | Quantum Fisher information for Gaussian states . . . . .              | 81         |
| 3.2.4    | Uniform sampling of single-mode Gaussian states . . . . .             | 83         |
| 3.3      | Quantification of probe versatility . . . . .                         | 85         |
| 3.3.1    | The model . . . . .                                                   | 85         |
| 3.3.2    | Average quantum Fisher information . . . . .                          | 87         |
| 3.3.3    | AvQFI for squeezing estimation . . . . .                              | 89         |
| 3.4      | Results on squeezing estimation . . . . .                             | 91         |
| 3.4.1    | Noiseless evolution . . . . .                                         | 92         |
| 3.4.2    | Lossy evolution . . . . .                                             | 96         |
| 3.5      | Conclusions . . . . .                                                 | 99         |
| <b>4</b> | <b>Versatile communication bounds for quantum networks</b>            | <b>103</b> |
| 4.1      | Introduction . . . . .                                                | 103        |
| 4.1.1    | Quantum communication . . . . .                                       | 103        |
| 4.1.2    | Our contribution in light of recent results . . . . .                 | 105        |
| 4.2      | Preliminary notions . . . . .                                         | 108        |
| 4.2.1    | Entanglement measures . . . . .                                       | 108        |
| 4.2.2    | Target states: maximally entangled or private states . . . . .        | 110        |
| 4.2.3    | Choi-stretchable channels . . . . .                                   | 112        |
| 4.2.4    | Adaptive communication protocols . . . . .                            | 114        |
| 4.2.5    | Quantifying the performance of a communication protocol . . . . .     | 117        |
| 4.3      | Entanglement-based upper bounds . . . . .                             | 118        |
| 4.3.1    | Key properties . . . . .                                              | 119        |
| 4.3.2    | Upper bounds based on a single entanglement measure . . . . .         | 122        |
| 4.3.3    | Proofs of Theorems 4.1 and 4.2 . . . . .                              | 124        |
| 4.4      | Versatile upper bound on quantum networks . . . . .                   | 125        |
| 4.4.1    | Versatile Property P2 . . . . .                                       | 126        |
| 4.4.2    | Upper bound combining multiple entanglement measures . . . . .        | 127        |
| 4.5      | Max-relative entropy of entanglement of some qubit channels . . . . . | 129        |
| 4.5.1    | Tools for qubit channels invariant under phase rotations . . . . .    | 130        |
| 4.5.2    | Evaluation on qubit channels . . . . .                                | 132        |
| 4.6      | Examples . . . . .                                                    | 134        |
| 4.6.1    | Bipartition composed by flower and distillable channels . . . . .     | 136        |
| 4.6.2    | Bipartition composed by amplitude damping and dephasing channels      | 138        |
| 4.7      | Conclusions . . . . .                                                 | 140        |

|          |                                                                             |            |
|----------|-----------------------------------------------------------------------------|------------|
| <b>5</b> | <b>A global perspective</b>                                                 | <b>141</b> |
|          | <b>Bibliography</b>                                                         | <b>142</b> |
| <b>A</b> | <b>Technicalities on squeezing estimation</b>                               | <b>157</b> |
| A.1      | Two-mode AvQFI in noiseless conditions . . . . .                            | 157        |
| A.2      | Single-mode probes with extreme AvQFI values . . . . .                      | 158        |
| A.3      | Single-mode QFI under noisy conditions . . . . .                            | 159        |
| A.4      | Optimal displacement direction for single-mode probes . . . . .             | 160        |
| <b>B</b> | <b>Technicalities on communication bounds</b>                               | <b>163</b> |
| B.1      | Proof of property P1 for the max-relative entropy of entanglement . . . . . | 163        |
| B.2      | Proof of the upper bound in Proposition 4.2 . . . . .                       | 165        |
| B.3      | Proof of the lower bound in Proposition 4.2 . . . . .                       | 167        |



# List of Figures

|     |                                                                                                                                             |     |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 2.1 | Schematic representation of the Hong-Ou-Mandel effect. . . . .                                                                              | 31  |
| 2.2 | Hong-Ou-Mandel dip in the probability of observing a coincident detection. . . . .                                                          | 33  |
| 2.3 | Sketch of the studied classical interferometric setup, with $N_s$ sources and $N_d$ detectors. . . . .                                      | 41  |
| 2.4 | Classical lower bound on correlations among pairs of output intensities. . . . .                                                            | 42  |
| 2.5 | Optimal interferometric structure. . . . .                                                                                                  | 44  |
| 2.6 | Sketch of the studied quantum interferometric setup, with $N_s$ sources and $N_d$ detectors. . . . .                                        | 49  |
| 2.7 | Certification of the impossibility of dividing an interferometer into smaller independent blocks. . . . .                                   | 58  |
| 3.1 | Sketch of a physical scenario in which a squeezing operation along an initially unknown direction is applied on the probing system. . . . . | 87  |
| 3.2 | Performance of single-mode Gaussian states for noiseless squeezing estimation. . . . .                                                      | 94  |
| 3.3 | Numerical investigations leading to the identifications of the optimal single-mode probe in noisy condition. . . . .                        | 98  |
| 3.4 | Dependence of the optimal single-mode probe on the squeezing strength. . . . .                                                              | 99  |
| 3.5 | Comparison of the average precisions attained by two-mode squeezed states and the best single-mode probe. . . . .                           | 100 |
| 4.1 | Pictorial representation of a Choi-stretchable channel. . . . .                                                                             | 113 |
| 4.2 | Example of quantum network. . . . .                                                                                                         | 115 |
| 4.3 | Bounds on the capacity of the qubit amplitude damping channel. . . . .                                                                      | 134 |
| 4.4 | Bounds on the capacity of three Choi-stretchable qubit channels. . . . .                                                                    | 135 |
| 4.5 | Example of network bipartition where the versatile bound can be advantageous. . . . .                                                       | 138 |
| 4.6 | Advantage of the versatile bound applied on a simple quantum network. . . . .                                                               | 139 |
| B.1 | Typical plot of the function $f_0(\mu \lambda, \nu, \xi)$ defined in Equation (B.26). . . . .                                               | 168 |



# Chapter 1

## Thesis Overview

### 1.1 Quantum Information: history and perspectives

The history of quantum mechanics can be traced back to Max Planck and his seminal work on black-body radiation, followed by Einstein's explanation of the photoelectric effect. From that point, physicists were forced to reconsider not only the structure of light, but also to develop a whole new theory to describe the behaviour of nature at small scales and low energies. This process required several decades to be completed, during which the tentative semi-classical approaches initially proposed turned into formal mathematical frameworks. The conceptual revolution brought by quantum mechanics allowed for a much deeper understanding of nature, which led to technological applications before unimaginable. Among the most remarkable examples, that greatly impacted our society, we can mention innovative imaging methods based on the wave-like character of electrons, the possibility of generating powerful coherent light beams by means of lasers, and the development of small processors based on the properties of transistors.

Another shift in perspective came around nearly thirty years ago, with the realisation that information is physical, and as such its manipulation has to obey the laws of physics. The fact that any device ultimately has a quantum description, along with the progressive reduction in size of technological components, naturally led to the development of a new field: quantum information. Combining physics with information theory, which was developed during the second half of the twentieth century after Shannon's seminal work [1], this is now a well-established research topic. On the one hand, it aims at exploiting the laws of quantum mechanics to improve our ability of transmitting and processing information. On the other hand, it wants to explore to what extent it is possible to do so, and which are the features that make this possible. In addition to the intellectual interest of the problem, the possibility of harnessing the full power allowed by quantum mechanics in order to improve humankind's computing and communication capabilities is very appealing. Indeed, in today's world the amount of information produced on a daily basis is increasing exponentially, while our ability to elaborate it is lagging behind. For this

reason, many political and industrial organisations are nowadays funding the development of quantum technologies, with the aim of bridging the gap between academic knowledge and technological applications. It is very likely that this “second quantum revolution” will bring a considerable impact to the world, similarly to what already happened a century ago.

One of the long-term objectives in quantum information is the creation of the first programmable universal quantum computer, whose realisation is still quite far in the future. Nonetheless, it is interesting to search for those situations and specific tasks, perhaps easier to realise, where quantum mechanical laws could yield an advantage over their classical counterpart. The concept of “classicality”, however, can change its meaning depending on the context and on the task at hand. In order to identify a quantum advantage, one should first provide a precise definition of classicality, and then understand which experimental outcomes are obtainable within those constraints. Only on a second moment, we can ask ourselves whether some figure of merit for the considered task could be improved if we abandon the chosen classical regime, by embracing the full power of quantum mechanics. From this perspective, having a way to detect and quantify non-classical effects is the first step towards their exploitation. The first part of this thesis deals with these issues of detecting (Chapter 2) and measuring (Chapter 3) nonclassicality in two situations of interest for the quantum optical community. Further details on the considered scenarios will be provided in the following.

Another long-term objective in the field of quantum information is the construction of a large-scale network, which can be used to faithfully transmit quantum states among its nodes. In this scenario, each node will possess a fully functional quantum computer, so that the aforementioned network would be the analogue of today’s internet, also known as the “quantum internet” [2]. Although the manipulation of local quantum states is still severely limited by current technology, the realisation of the first metropolitan quantum networks is already under way in several areas of the world [3–6]. Typically, the main quantum information carriers in these architectures are photons, which are exchanged through optical fibres or free-space links. Each site, then, can locally process the received information by exploiting light-matter interactions (see e.g. Ref. [7]). Applications within experimental reach include the distribution of secure private keys for cryptographic purposes, and the transmission of entangled quantum states, which could then be used as building blocks for more complex protocols. As in any classical communication system, also the performance of this sort of network has to be limited by the noise affecting the transmission lines, which reduces the amount of information faithfully transmitted per channel use. Finding the ultimate communication rate obtainable with a given quantum channel, or network, is one of the main theoretical goals within this branch of quantum information [8]. In Chapter 4 we will contribute to these investigations, by improving known upper bounds on the number of private bits, or maximally entangled qubits, that could



be shared by two local parties at the end of a communication protocol over a quantum network, when free and unlimited classical communication is allowed.

In addition to the aforementioned challenges, there are many other branches of research within the realm of quantum information. As non-exhaustive examples, we can mention the research on quantum thermodynamics, open systems dynamics, quantum control or quantum biology. Overall, quantum information is a recent and exciting field, which not only can be used to reinterpret and further develop several branches of existing knowledge, but that also holds promise for important technological advancements. However, at the moment several challenges still remain to be addressed before quantum information could achieve its full potential, and only time will tell to what extent it will be able to impact our society.

## 1.2 Summary of results

This thesis can be divided into three, almost independent, main blocks, composed by Chapters 2, 3, and 4. In the first two we respectively address the issues of detecting and measuring nonclassical effects in optical systems, while in the last one we consider a problem of quantum communication. Each of them begins with a detailed introduction, in which we discuss the motivations behind our investigation and some connections with existing literature, followed by an overview of the preliminary notions needed to present our findings. The core of each of these chapters consists in the presentation of our original results, followed by a discussion and by some conclusive remarks. A final comment on the novel contributions discussed in this thesis can be found in Chapter 5.

Here we provide a short and (almost) free from technicalities summary of each main chapter, from which the reader can gain a broad understanding of the content of this thesis.

- In Chapter 2 we consider a generic linear optical setup, which can be thought of as being made by lenses, mirrors, waveplates, and other components that act linearly on the input light. In addition to respecting the superposition principle, these must also preserve the frequency of any monochromatic input light. These kind of apparatus have been widely used to study interference effects for hundreds of years, and are still at the core of any optical experiment. In particular, we study a multimode setting, that accepts several light beams in input and split them among multiple outputs, on which intensity measurements are performed. Upon repeating the experiment many times, interference effects can be observed in the correlation among the recorded intensity values, which carry information on the input light as well as on the interferometer.

The goal of this chapter is to set a tight bound on the amount of intensity correlations that could be possibly obtained by using only classical sources, fully described

by Maxwell’s theory of light. We do so by fixing a measure of correlation, and by taking two assumptions: the independence of the sources, and the randomness of the phase characterising each emitted light pulse. By optimising over all input states of this kind, and over all linear optical networks, we find the minimum value of correlations that could be obtained classically. Hence, the violation of this threshold in an experiment certifies the nonclassicality of the result, which cannot be explained in any way within the aforementioned classical framework. With a fully quantum treatment of the phenomenon, we then show that only quantum states with sub-Poissonian photon-number statistics can violate the classical bound. Our findings generalise to a multimode setting the Hong-Ou-Mandel effect, one of the cornerstones in the field of quantum optics.

This chapter is based upon the paper in Ref. [9].

- In Chapter 3 we consider a particular notion of optical nonclassicality, known as “squeezing”, which arises when the intrinsic noise of the light is compressed below the vacuum level. Needless to say, this can be done only for one degree of freedom, while the complementary one undergoes the opposite process, in such a way that Heisenberg’s indetermination principle is not violated. The possibility of “hiding” the noise where it is less detrimental can be a very useful tool in many situations of interest, and squeezed light is nowadays accepted as one of the typical quantum features which can arise in optical systems.

In particular, in this chapter we assume to have a squeezing device at our disposal, which takes some input light and squeezes it along an unknown direction, possibly fluctuating in time. Our goal is to select the optimal “versatile” probing system, capable of capturing, on average, as much information as possible on the amount of squeezing applied by the device. We also assume that the squeezing direction can be exactly deduced by the experimenter upon recollection of the probe, so that this information can be used to select and perform the optimal readout measurement. In this framework, we introduce a figure of merit which characterises the average estimation precision associated with each probe, and we use it in order to identify good probing systems among a set of isoenergetic and experimentally feasible states. The effects of input correlations and photon losses are also discussed.

This chapter is based upon the paper in Ref. [10].

- In Chapter 4 we consider a network of quantum channels, able to transmit quantum information between distant locations, with the aim of quantifying the ultimate communication performance achievable by two parties having access to the network. In doing so, we assume that each node of the network has full control over its local quantum systems, is cooperative, and can freely exchange classical information over a public transmission line. As paradigmatic communication task, we consider the

generation of a string of shared private bits, which could then be used as secret key within a cryptographic protocol. Alternatively, we can also deal with the problem of generating pairs of maximally entangled states, whose correlations cannot be explained classically. Indeed, this second problem is a particular instance of the former, because each maximally entangled qubit straightforwardly leads to a pair of secret shared bits whenever the exchange of additional public information is allowed.

More precisely, under the aforementioned assumptions, in this chapter we develop a versatile upper bound on the number of private bits shared by two parties after a generic communication protocol over the network. As in previous studies, we split the network into two components and focus our study to the channels connecting them: for each channel we compute the maximum amount of entanglement that it could possibly generate, and then we average these values according to the number of times each channel has been used. For some particular entanglement measure, this approach is known to lead to an upper bound on the number of private bits shared at the end of the protocol. However, typically one has to sacrifice either the tightness of the bound or the possibility of applying it to a broad class of channels, depending on the chosen entanglement measure. Instead, in our approach we are not forced to select a single measure, but we can tailor this choice on a channel-by-channel basis, in order to make our bound as tight as possible with limited requirements on the networks where it can be applied.

This chapter is based upon the pre-print article in Ref. [11].

## 1.3 Required background and comments on notation

Throughout the thesis we will introduce most of the tools and concepts we will be using. However, some preliminary notions of quantum mechanics and quantum information theory are given for granted here and there. In particular, we will assume familiarity with:

- Dirac's ket-bra notation, and the notion of density matrix;
- the Fock space and the quantum harmonic oscillator;
- the Kraus and Stinespring representations of any physical evolution allowed by quantum mechanics, and the equivalent characterisation of the latter as a completely positive trace-preserving linear map;
- the positive operators formalism for generalised quantum measurements (POVM);
- the notions of entanglement, qubit and Bloch's sphere.

For an introduction on these important topics, we refer the reader to John Watrous' lecture notes on quantum information theory, which can be found at:  
<https://cs.uwaterloo.ca/~watrous/LectureNotes.html>.

Despite the variety of topics addressed in this thesis, a lot of effort has been put to keep the notation consistent. Nonetheless, although we tried to limit as much as possible the number of instances in which the same symbol changes meaning across different chapters, in some occasions this could not be avoided without adding additional indices or changing standard notations. In these cases the repetition of a symbol seemed the least confusing choice, and we are confident in the fact that context will remove any possible doubt. This said, there are a few conventions that are consistently followed throughout the thesis. A first and simple one is that vectors are written in bold, e.g., as  $\mathbf{A}$  or  $\boldsymbol{\xi}$ . A second one concerns the use of “hats”, e.g. as in  $\hat{U}$  or  $\hat{a}$ , which are added to every quantum operator acting on vectors of a Hilbert space. However, although density matrices could be considered operators of this kind, in order to keep the notation simple we decided not to use hats for them. The superoperators characterising the evolution of quantum states, instead, are typically written with a calligraphic style, as  $\mathcal{U}$ ,  $\mathcal{R}$  or  $\mathcal{S}$ , or with a capital Greek letter, as  $\Phi$  or  $\Lambda$ . A final important remark concerns the particular case of unitary evolutions. If  $\hat{U}$  is the quantum operator characterising the evolution, the superoperator acting on density matrices as  $\hat{U}[\cdot]\hat{U}^\dagger$  is always represented with the calligraphic version of the same letter, in this case  $\mathcal{U}$ . Finally, finite-dimensional matrices are typically represented by capital Roman letters, as  $U$  or  $T$ . Of course, for finite-dimensional quantum systems one can always move from  $\hat{U}$  to  $U$  simply by fixing a reference basis.

# Chapter 2

## Nonclassicality detection in multiport interferometry

### 2.1 Introduction

Light interference has been studied for a very long time, and has found a wide range of applications, going from fundamental science (e.g., Young [12] or Michelson-Morley [13] experiments) to the technological development of imaging systems, as microscopes or radars. Towards the end of last century, with a fully developed theory of quantum mechanics and the first experimental realisations of exotic states of light, physicists began to predict and observe unexpected interference behaviours. One of the most striking examples goes under the name of Hong-Ou-Mandel (HOM) effect [14], from the names of the scientists who first observed it. In short, when two pure, indistinguishable, and independent photons evolve through a balanced beamsplitter, they always exit together from one of the two output ports. This behaviour is rather surprising because a beamsplitter, that can be simply modelled as a half-reflecting and half-transmitting mirror, seems to be able to correlate otherwise independent input light beams. This phenomenon has no classical counterpart, as we will extensively discuss in the following, and over the years it has found many applications, so that it is now considered one of the cornerstones of quantum optics. For example, it can be used as a tool to measure the degree of distinguishability between the two injected photons [15], and it underlies the most basic entangling mechanism used in linear optical quantum computing [16]. Moreover, it can be used to produce the simplest non-trivial N00N state [17, 18] useful for quantum metrology and sensing [19, 20], and it found applications in entanglement detection [21] or swapping [22].

Recently, several authors have investigated interference effects of non-interacting particles, aiming at reproducing or generalising HOM's result to different situations. For example, by considering highly symmetric interferometers, the authors of Refs. [23, 24] found a strict suppression law for most possible output events, while the possibility of having coincident detections in all output ports was investigated in Ref. [25]. Although

these sort of studies are not necessarily constrained to a linear optical setting [26–30], photonics seems to remain the physical platform of choice. Indeed, with current technology multiple photons can be prepared and manipulated at the same time, and can then be injected into multimode interferometers [31–39]. Recent studies on many-particle interference effects revealed a need for a deeper understanding of the phenomenon, and current research on the topic is mostly developing along two directions. On the one hand, from a computational perspective the problem of boson sampling is currently the best candidate to show the possibility of outperforming classical computers by exploiting quantum mechanical laws [35, 36, 40–44]. This is because it is remarkably hard to predict the evolution of photons through an array of passive linear optical devices [45]. On the other hand, from a foundational point of view, it is interesting to study the interplay between the wave-like behaviour of photons and the many-particle interference effects arising because of their bosonic nature [25, 46–51]. Both these features heavily influence the probabilities of detection events, and their joint action is often hard to predict and interpret [46, 49, 50].

In order to identify events and properties that can be considered nonclassical, it is necessary to compare the predictions of quantum mechanics with those of a classical framework. The typical choice for the latter involves considering completely distinguishable photons, which cannot interfere with each other. In this way, their wave-like character is completely lost, and their evolution through a linear network can be easily studied, because it can be mapped into the scattering of classical particles according to some given probabilities. An alternative choice consists in choosing a classical framework where the wave-like character of light is retained, while its particle-like bosonic features are completely lost. This regime is naturally described by the classical theory of electrodynamics, which is elegantly summarised by Maxwell’s equations. In the past, this theory of light was used to predict the intensity correlations expected to be observed in a HOM setup receiving classical fields in input. The inconsistency of these predictions with the actual experiment provided a clear indication for the fact that HOM effect had to be considered an intrinsically quantum phenomenon [52–54].

In this chapter we generalise the same reasoning to a multimode linear optical setup, with several independent sources and many intensity-sensitive detectors. In particular, we find a tight lower bound on the amount of correlations that could be found among the output intensities of a generic multimode interferometer, when classical electric fields are injected in input. By using the standard quantum mechanical formalism, we then show that quantum input states of light can lead to a violation of this threshold. The remainder of this chapter is structured as follows. After a comparison between the theories of classical and quantum optics, in Section 2.3 we review the original Hong-Ou-Mandel effect. In doing so, we will emphasise the reasons why it represents a signature of nonclassicality, and we set the basis for its multimode generalisation, which will be formally developed in Section 2.4. A discussion about our results is presented in Section 2.5, where we also

compare them with other notions of nonclassicality. Our conclusive remarks on the topic of this chapter are left for Section 2.6.

## 2.2 Quantum and classical optics

A complete description of the canonical procedure for the quantisation of the electromagnetic field is beyond the purpose of this thesis, and we shall give it for granted. For an in depth discussion, we refer the reader to one of the many available books on the topic (see, e.g., Ref. [55]). Nonetheless, it is beneficial to compare the formalism used to describe quantum states of the fields with its classical counterpart, if only to introduce the reader to the notation that will be used throughout the chapter. At the end of this section, particular attention will be devoted to intensity measurements.

The typical approach to quantum electrodynamics (QED) assumes that the fields are confined within an optical cavity, whose dimensions can then be sent to infinity in order to recover an expression for the electric and magnetic fields over the whole three-dimensional space. However, the arrangement of most optical experiments involves light beams travelling in straight lines with a given cross-sectional area  $\mathcal{A}$ , for example in optical fibres or on-chip waveguides. For this reason, it is convenient to consider a one-dimensional version of QED, that naturally describes field excitations in this framework. Here we take this approach, closely following the discussion detailed in Ref. [56].

### 2.2.1 Quantum states of light

Let us consider the propagation of light along a one-dimensional path parametrised by the coordinate  $z$ . After the quantisation procedure, the vector potential operator  $\hat{\mathbf{A}}(z, t)$  in Coulomb gauge for a field propagating in free space towards the positive  $z$  direction can be written as

$$\hat{\mathbf{A}}(z, t) = \hat{\mathbf{A}}^+(z, t) + \hat{\mathbf{A}}^-(z, t), \quad (2.1)$$

where

$$\hat{\mathbf{A}}^+(z, t) = \sum_{\lambda=1,2} \int_0^{+\infty} d\omega \left( \frac{\hbar}{4\pi\epsilon_0\omega\mathcal{A}} \right)^{\frac{1}{2}} \hat{a}_{\omega,\lambda} \boldsymbol{\epsilon}_{\omega,\lambda} e^{-i\omega(t-z/c)}, \quad (2.2)$$

while  $\hat{\mathbf{A}}^-(z, t)$  is its adjoint operator. Here,  $\lambda$  denotes the polarisation of the field through the vectors  $\boldsymbol{\epsilon}_{\omega,\lambda}$ , which satisfy

$$\boldsymbol{\epsilon}_{\omega,\lambda} \cdot \boldsymbol{\epsilon}_{\omega,\lambda'}^* = \delta_{\lambda,\lambda'}, \quad \vec{z} \cdot \boldsymbol{\epsilon}_{\omega,\lambda} = 0, \quad (2.3)$$

where  $\vec{z}$  is the unit vector pointing towards the propagation direction, and the operators  $\{\hat{a}_{\omega,\lambda}\}_{\omega,\lambda}$  satisfy the commutation relations

$$[\hat{a}_{\omega,\lambda}, \hat{a}_{\omega',\lambda'}] = 0, \quad [\hat{a}_{\omega,\lambda}, \hat{a}_{\omega',\lambda'}^\dagger] = \delta(\omega - \omega')\delta_{\lambda,\lambda'}. \quad (2.4)$$

The operators for the electric and magnetic fields can be obtained as

$$\hat{\mathbf{E}} = -\frac{\partial}{\partial t}\hat{\mathbf{A}}, \quad \hat{\mathbf{B}} = \nabla \wedge \hat{\mathbf{A}}, \quad (2.5)$$

in analogy with the classical case, so that their contributions involving the annihilation operators are given by

$$\hat{\mathbf{E}}^+(z, t) = i \left( \frac{1}{2\epsilon_0 c \mathcal{A}} \right)^{\frac{1}{2}} \sum_{\lambda=1,2} \int_0^{+\infty} d\omega \left( \frac{\hbar\omega}{2\pi} \right)^{\frac{1}{2}} \hat{a}_{\omega,\lambda} \boldsymbol{\epsilon}_{\omega,\lambda} e^{-i\omega(t-z/c)}, \quad (2.6)$$

$$\hat{\mathbf{B}}^+(z, t) = i \left( \frac{1}{2\epsilon_0 c^3 \mathcal{A}} \right)^{\frac{1}{2}} \sum_{\lambda=1,2} \int_0^{+\infty} d\omega \left( \frac{\hbar\omega}{2\pi} \right)^{\frac{1}{2}} \hat{a}_{\omega,\lambda} \vec{z} \wedge \boldsymbol{\epsilon}_{\omega,\lambda} e^{-i\omega(t-z/c)}. \quad (2.7)$$

By using these expressions, the Hamiltonian of the field can be decomposed as a collection of harmonic oscillators

$$\begin{aligned} \hat{H} &= \mathcal{A} \int_{-\infty}^{+\infty} dz \frac{\epsilon_0}{2} \left[ c^2 \hat{\mathbf{B}}(z, t) \cdot \hat{\mathbf{B}}(z, t) + \hat{\mathbf{E}}(z, t) \cdot \hat{\mathbf{E}}(z, t) \right] \\ &= \sum_{\lambda=1,2} \int_0^{+\infty} d\omega \hbar\omega \left( \hat{a}_{\omega,\lambda}^\dagger \hat{a}_{\omega,\lambda} + \frac{1}{2} \right), \end{aligned} \quad (2.8)$$

as can be easily verified thanks to Equations (2.3) and (2.4) by making use of the relation

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dx e^{-ix(\omega-\omega')} = \delta(\omega - \omega'). \quad (2.9)$$

As it is customary, we call “photons” the excitations of the field, and we call “vacuum” the ground state  $|0\rangle$  of the system, which satisfies the set of conditions  $\hat{a}_{\omega,\lambda}|0\rangle = 0$ , for all  $\omega > 0$  and  $\lambda = 1, 2$ . The whole Hilbert space of the system can then be generated by taking linear superpositions of states obtained by applying any number of creation operators on the vacuum:

$$|\psi\rangle = \sum_{\{n_{\omega,\lambda}\}_{\omega,\lambda}} \psi(\{n_{\omega,\lambda}\}_{\omega,\lambda}) \left[ \prod_{\omega,\lambda} \left( \hat{a}_{\omega,\lambda}^\dagger \right)^{n_{\omega,\lambda}} \right] |0\rangle. \quad (2.10)$$

Experimentally obtainable photons are not completely delocalised in space. Therefore, they cannot be written as  $\hat{a}_{\omega,\lambda}^\dagger|0\rangle$ , but only as linear superposition of different states of that form. Given a normalised function  $g : (0, \infty) \times \{1, 2\} \rightarrow \mathbb{C}$ , which associates to each



pair  $(\omega, \lambda)$  a weight  $g(\omega, \lambda)$  such that

$$\sum_{\lambda=1,2} \int_0^{+\infty} d\omega |g(\omega, \lambda)|^2 = 1, \quad (2.11)$$

we can define a new annihilation operator  $\hat{a}_g$  as

$$\hat{a}_g = \sum_{\lambda=1,2} \int_0^{+\infty} d\omega g^*(\omega, \lambda) \hat{a}_{\omega,\lambda}. \quad (2.12)$$

The normalisation condition in Equation (2.11) guarantees the canonical form of the commutator  $[\hat{a}_g, \hat{a}_g^\dagger] = 1$ , so that the state  $\hat{a}_g^\dagger |0\rangle$  represents a well-defined single excitation of the field. In the following, we will say that  $\hat{a}_g^\dagger$  creates photons in mode  $g$ . Any pure quantum state of the field in this mode can then be easily expanded as

$$|\psi\rangle_g = \sum_{n=0}^{+\infty} \psi(n) (\hat{a}_g^\dagger)^n |0\rangle. \quad (2.13)$$

### 2.2.2 Comparison with classical fields

At this point, it is worthwhile to take a step back, write the classical expressions of the electromagnetic fields, and then compare them with the quantum description of light previously discussed. The complex classical vector potential for a light beam propagating along the positive  $z$  direction, analogous to Equation (2.2), can be written as

$$\mathbf{A}(z, t) = A \sum_{\lambda=1,2} \int_0^{+\infty} d\omega \left( \frac{\hbar}{4\pi\epsilon_0\omega\mathcal{A}} \right)^{\frac{1}{2}} \tilde{g}_{\omega,\lambda} \boldsymbol{\epsilon}_{\omega,\lambda} e^{-i\omega(t-z/c)}, \quad (2.14)$$

where  $A \in \mathbb{C}$  represents the amplitude of the field. The complex coefficients  $\{\tilde{g}_{\omega,\lambda}\}_{\omega,\lambda}$  play the same role of the weights  $\{g(\omega, \lambda)\}_{\omega,\lambda}$  appearing in Equation (2.12), and are similarly normalised:

$$\sum_{\lambda=1,2} \int_0^{+\infty} d\omega |\tilde{g}_{\omega,\lambda}|^2 = 1. \quad (2.15)$$

Note that the factor appearing within the square root in the definition of  $\mathbf{A}(z, t)$  is inserted with the only purpose of easing the comparison with the quantum description [see Equation (2.2)], and it could be removed by suitably changing the coefficients  $\tilde{g}_{\omega,\lambda}$  and the amplitude  $A$ . By introducing a function  $\zeta_{\tilde{g}}(z, t)$  describing the mode of the fields

$$\zeta_{\tilde{g}}(z, t) = \sum_{\lambda=1,2} \int_0^{+\infty} d\omega \left( \frac{\hbar\omega}{2\pi} \right)^{\frac{1}{2}} \tilde{g}_{\omega,\lambda} \boldsymbol{\epsilon}_{\omega,\lambda} e^{-i\omega(t-z/c)}, \quad (2.16)$$

the classical complex electric and magnetic fields can be written as:

$$\mathbf{E}(z, t) = i \left( \frac{1}{2\epsilon_0 c \mathcal{A}} \right)^{\frac{1}{2}} A \boldsymbol{\zeta}_{\tilde{g}}(z, t), \quad (2.17)$$

$$\mathbf{B}(z, t) = i \left( \frac{1}{2\epsilon_0 c^3 \mathcal{A}} \right)^{\frac{1}{2}} A [\vec{z} \wedge \boldsymbol{\zeta}_{\tilde{g}}(z, t)]. \quad (2.18)$$

The quantum and classical descriptions of light differ for two main reasons.

- *The selection of a particular mode of light.* In the classical case, any field is characterised by a certain mode  $\tilde{g}$ , determined by the expansion in the frequency and polarisation degrees of freedom given in Equation (2.14), whereas a quantum state of light can contain multiple excitations of different modes. For example, quantum mechanics allows for superpositions of multiple states, composed by different numbers of photons and by different modes of light.
- *The amplitude of the field.* Once a particular mode of light has been selected, the classical field is completely determined by a single complex number  $A$ . On the contrary, a quantum state of light is characterised by a possibly infinite number of amplitudes, e.g., the coefficients  $\{\psi(n)\}_n$  in Equation (2.13).

The classical picture can be recovered from the quantum one by considering a *coherent state* with amplitude  $A \in \mathbb{C}$  and mode  $g$  [57, 58], defined as

$$|A\rangle_g = \sum_{n=0}^{\infty} e^{-\frac{|A|^2}{2}} \frac{A^n}{n!} (\hat{a}_g^\dagger)^n |0\rangle. \quad (2.19)$$

By exploiting the commutation relation

$$[\hat{a}_{\omega, \lambda}, (\hat{a}_g^\dagger)^n] = n g(\omega, \lambda) (\hat{a}_g^\dagger)^{n-1}, \quad (2.20)$$

together with the condition  $\hat{a}_{\omega, \lambda} |0\rangle = 0$ , we can easily show that for a generic quantum state  $|\psi\rangle_g$  characterised by a single mode of light  $g$  [see Equation (2.13)], one has

$$\hat{a}_{\omega, \lambda} |\psi\rangle_g = g(\omega, \lambda) \hat{a}_g |\psi\rangle_g. \quad (2.21)$$

Therefore, on this subset of quantum states, the annihilation operators  $\hat{a}_{\omega, \lambda}$  and  $\hat{a}_g$  have the same effect, up to a proportionality factor. We can use this property, together with the relation  $\hat{a}_g |A\rangle_g = A |A\rangle_g$ , in order to show that  $|A\rangle_g$  is an eigenvector of the electric field operator  $\hat{\mathbf{E}}^+$ :

$$\hat{\mathbf{E}}^+(z, t) |A\rangle_g = i \left( \frac{1}{2\epsilon_0 c \mathcal{A}} \right)^{\frac{1}{2}} A \boldsymbol{\zeta}_g(z, t) |A\rangle_g = \mathbf{E}(z, t) |A\rangle_g, \quad (2.22)$$

whose eigenvalue corresponds to the complex classical electric field given in Equation (2.17), with mode of light  $\tilde{g}_{\omega,\lambda} = g(\omega, \lambda)$ . Because of this relation, in the following we will not distinguish between  $\tilde{g}$  and  $g$  any more, and for both scenarios we will simply refer to a state of light in mode  $g$ .

*Remark:* In writing expressions for the classical fields, we implicitly assumed that the weights  $\tilde{g}_{\omega,\lambda}$  could be normalised as in Equation (2.15). However, the integration on the left-hand side is not necessarily finite, as the requirement coming from the fact that the field must have a finite energy only yields:

$$\sum_{\lambda=1,2} \int_0^{+\infty} d\omega \hbar\omega |\tilde{g}_{\omega,\lambda}|^2 < \infty. \quad (2.23)$$

Equation (2.15) is an independent condition, that guarantees the possibility of interpreting the light as composed by photons with a frequency-dependent energy given by  $\hbar\omega$ . This requirement cannot arise from a classical theory of light, but it has to be adopted if we want to compare the predictions of the latter with those arising from a quantum framework.

### 2.2.3 Intensity measurements

Light-matter interactions and the theory of photo-detection of light have been extensively studied, and we refer the reader to any quantum optics book for a detailed discussion on the subject. Here we prefer to take an abstract approach, by assuming to deal with detectors that can measure the whole energy carried to their location by the electromagnetic field, within the time window in which they operate. In the following we connect this measurable quantity to other features of the field, as the energy associated with its mode of light or its amplitude.

Classically, the energy flux density is characterised by the real Poynting vector

$$\mathbf{S}_R(z, t) \equiv \frac{1}{\mu_0} [\mathbf{E}(z, t) \wedge \mathbf{B}^*(z, t) + \mathbf{E}^*(z, t) \wedge \mathbf{B}(z, t)], \quad (2.24)$$

whose magnitude for a field in mode  $g$  is given by:

$$|\mathbf{S}_R(z, t)| = \frac{2}{\mu_0 c} |\mathbf{E}(z, t)|^2 = \frac{|A|^2}{\mathcal{A}} |\zeta_g(z, t)|^2, \quad (2.25)$$

where we used Equations (2.17) and (2.18). The total intensity that a detectors receives can be found by integrating this quantity over the detection time  $\tau_M$ . For a pulse of light with duration much smaller than  $\tau_M$ , reaching a detector placed at  $z = z_0$  at time  $t \sim 0$ ,

the integrated intensity of the pulse is given by

$$I = \int_{-\infty}^{+\infty} dt |\mathbf{S}_R(z_0, t)| = \frac{\mathcal{E}_g}{\mathcal{A}} |A|^2, \quad (2.26)$$

where  $\mathcal{E}_g$  is the energy associated with mode  $g$ :

$$\mathcal{E}_g = \sum_{\lambda=1,2} \int_0^{+\infty} d\omega \hbar \omega |g(\omega, \lambda)|^2. \quad (2.27)$$

In the same regime, we can similarly define a normal-ordered Poynting vector operator

$$\hat{\mathbf{S}}(z, t) = \frac{1}{\mu_0} \left[ \hat{\mathbf{E}}^-(z, t) \wedge \hat{\mathbf{B}}^+(z, t) + \hat{\mathbf{B}}^-(z, t) \wedge \hat{\mathbf{E}}^+(z, t) \right], \quad (2.28)$$

with modulus

$$|\hat{\mathbf{S}}(z, t)| = \frac{2}{\mu_0 c} \hat{\mathbf{E}}^-(z, t) \cdot \hat{\mathbf{E}}^+(z, t), \quad (2.29)$$

which can be used to obtain the intensity operator of the field at a fixed position  $z = z_0$

$$\hat{I} = \int_{-\infty}^{+\infty} dt |\hat{\mathbf{S}}(z_0, t)| = \frac{1}{\mathcal{A}} \sum_{\lambda=1,2} \int_0^{+\infty} d\omega \hbar \omega \hat{a}_{\omega,\lambda}^\dagger \hat{a}_{\omega,\lambda}. \quad (2.30)$$

For a pure quantum state  $|\psi\rangle_g$ , involving a single mode of light  $g$  as in Equation (2.13), the quantum counterpart of the classical intensity  $I$  can then be obtained by evaluating the expectation value:

$$\text{Tr} \left[ \hat{I} |\psi\rangle_g \langle\psi| \right] = \frac{\mathcal{E}_g}{\mathcal{A}} \text{Tr} \left[ \hat{a}_g^\dagger \hat{a}_g |\psi\rangle_g \langle\psi| \right] \quad (2.31)$$

where we exploited Equation (2.21). Therefore, whenever we are dealing with photons in the same mode of light, the expectation value of the intensity operator  $\hat{I}$  corresponds, up to a proportionality factor, to the expectation value of the total photon-number operator

$$\hat{n} = \sum_{\lambda=1,2} \int_0^{+\infty} d\omega \hat{a}_{\omega,\lambda}^\dagger \hat{a}_{\omega,\lambda}, \quad (2.32)$$

because

$$\text{Tr} \left[ \hat{a}_g^\dagger \hat{a}_g |\psi\rangle_g \langle\psi| \right] = \text{Tr} \left[ \hat{n} |\psi\rangle_g \langle\psi| \right]. \quad (2.33)$$

If we compare the classical intensity of Equation (2.26) with the quantum one in Equation (2.31), we see that they have a common term given by the energy associated with the considered mode of light, divided by the beam cross section. The remaining contribution is given respectively by the squared amplitude  $|A|^2$  and by the average number of photons in the quantum state.

## 2.3 Hong-Ou-Mandel effect

In this section, we give an overview of the original HOM effect, and we discuss its nonclassicality by comparing it to an analogous scenario in the framework of classical optics. This gives us an opportunity to introduce the reader to the key ideas behind the phenomenon, to the mathematical tools needed to describe it, and to the notation that will be used throughout the chapter. Later on, in Section 2.4, we will build on these concepts in order to extend this result to a multimode scenario.

More specifically, in Section 2.3.1 we begin by introducing a convenient formalism to describe the degrees of freedom of photons, and we use it to compute the probability of detecting one photon per output mode of a balanced beamsplitter, when two of them are injected in input. The HOM effect is very easily formulated in terms of this probability, but this approach has the drawback of not having a natural classical counterpart. In order to address this issue, we first rewrite the condition characterising the HOM effect in terms of intensity measurements, and then in Section 2.3.2 we show how the same result cannot be explained within the framework of classical electrodynamics.

It is worth stressing that all the results presented in this section are not original, and are well known by the community. Along the way, we will provide references to the original papers and to useful reviews on the topics that are being discussed.

### 2.3.1 Quantum description

It is important to emphasise that, from now on, we will deal with multiple light beams that are separately quantised with the procedure previously mentioned in Section 2.2. For this reason, the annihilation and creation operators characterising the system will be labelled by two indices separated by a semicolon, e.g., as in  $\hat{a}_{\alpha;g}$ . With this notation, the first index represents the path degree of freedom on which the operator annihilates or creates a photon. The second index, instead, describes the polarisation and spectral mode of the photon, or its “mode of light”, with the terminology of Section 2.2. In order to avoid confusion, from now on we will refer to these degrees of freedom as to the “spatial” and “spectral-polarisation” modes of the system. As different paths corresponds to independent quantisations, the following commutation relations hold:

$$[\hat{a}_{\alpha;g}, \hat{a}_{\beta;g'}] = 0, \quad [\hat{a}_{\alpha;g}, \hat{a}_{\beta;g'}^\dagger] = \delta_{\alpha,\beta}, \quad [\hat{a}_{\alpha;g}, \hat{a}_{\alpha;g'}^\dagger] = \delta_{\alpha,\beta} \sum_{\lambda=1,2} \int_0^{+\infty} d\omega g(\omega, \lambda)^* g(\omega, \lambda)', \quad (2.34)$$

where in order to obtain the rightmost equality we had to resort to the expansion of  $\hat{a}_{\alpha;g}$  in terms of its simpler components  $\{\hat{a}_{\alpha;\omega,\lambda}\}_{\omega,\lambda}$ , as in Equation (2.12).

It is now worth to introduce a shorthand notation to describe the spectral-polarisation degrees of freedom, that will be extremely useful throughout this chapter. Any function whose squared modulus is integrable can be thought as a vector in a  $L^2$  Hilbert space.

Hence, any function  $g(\omega, \lambda)$  that characterises a mode of light corresponds to a normalised vector  $|g\rangle_{\mathcal{M}} \in \mathcal{M} \equiv L^2 \otimes \mathbb{C}^2$  defined by

$$|g\rangle_{\mathcal{M}} = \sum_{\lambda=1,2} \int_0^{+\infty} d\omega |\omega, \lambda\rangle_{\mathcal{M}} g(\omega, \lambda). \quad (2.35)$$

Note that the component  $\mathbb{C}^2$  appearing in  $\mathcal{M}$  needs to be inserted because there are two independent polarisations for any given angular frequency. The integral in Equation (2.34) can now be rewritten by noticing that it corresponds to the inner product between  $|g\rangle_{\mathcal{M}}$  and  $|g'\rangle_{\mathcal{M}}$  in the Hilbert space  $\mathcal{M}$ :

$$[\hat{a}_{\alpha;g}, \hat{a}_{\beta;g'}^\dagger] = \delta_{\alpha,\beta} \langle g|g'\rangle_{\mathcal{M}}. \quad (2.36)$$

Photons characterised by orthogonal spectral-polarisation mode vectors are said to be “distinguishable”, and behave similarly to classical particles because the commutators in Equation (2.36) become zero. The opposite regime arises when the photons share the same spectral-polarisation mode of light: in this case they are said to be “indistinguishable”, and the most prominent quantum effects can be expected. The study of the intermediate regime, in which photons are only “partially distinguishable”, is more involved and we refer the reader to Refs. [59,60] for a detailed discussion on the subject. A simple approach that is used to study this scenario consists in expanding the modes of light of the considered photons in an orthonormal basis. For example, the space spanned by two vectors  $|g\rangle_{\mathcal{M}}$  and  $|g'\rangle_{\mathcal{M}}$  can be equivalently spanned by the orthonormal vectors  $|g\rangle_{\mathcal{M}}$  and  $|g^\perp\rangle_{\mathcal{M}}$ , where the latter is obtained through the standard Gram-Schmidt procedure:

$$|g^\perp\rangle_{\mathcal{M}} = \frac{|g'\rangle_{\mathcal{M}} - |g\rangle_{\mathcal{M}} \langle g|g'\rangle_{\mathcal{M}}}{\sqrt{1 - |\langle g|g'\rangle_{\mathcal{M}}|^2}}. \quad (2.37)$$

Therefore, the function  $g'(\omega, \lambda)$  can be expanded as

$$g'(\omega, \lambda) = \langle g|g'\rangle_{\mathcal{M}} g(\omega, \lambda) + \sqrt{1 - |\langle g|g'\rangle_{\mathcal{M}}|^2} g^\perp(\omega, \lambda), \quad (2.38)$$

which, together with the definition of  $\hat{a}_{\alpha,g'}^\dagger$  obtained from Equation (2.12), yields:

$$\hat{a}_{\alpha,g'}^\dagger = \langle g|g'\rangle_{\mathcal{M}} \hat{a}_{\alpha,g}^\dagger + \sqrt{1 - |\langle g|g'\rangle_{\mathcal{M}}|^2} \hat{a}_{\alpha,g^\perp}^\dagger. \quad (2.39)$$

The advantage of this formulation is that now  $\hat{a}_{\alpha,g}$  and  $\hat{a}_{\beta,g^\perp}$  satisfy canonical commutation relations, and the effect of partial distinguishability has been moved into some complex coefficients. Having introduced this formalism, we now turn to discuss the details of the HOM effect.

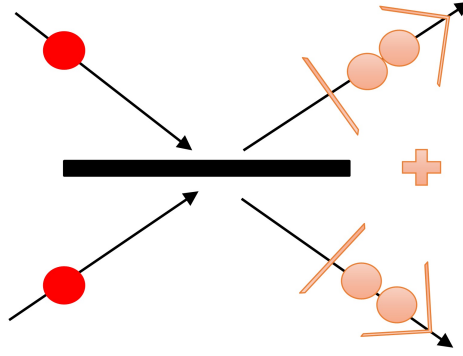


Figure 2.1: Schematic representation of the HOM effect, in which two pure indistinguishable photons impinge on the input ports of a balanced beamsplitter. The output state is a superposition of photons leaving from the same port.

### The HOM effect

The HOM effect concerns the evolution of two single photons impinging on the two input ports of a balanced beamsplitter. As we will shortly show, if the two photons are pure and indistinguishable, quantum mechanics forces them to always leave the beamsplitter with the same spatial mode [see Figure 2.1 for a pictorial representation]. This phenomenon of optical interference was discussed by Mandel in 1983 [52], and experimentally verified in 1987 [14]. For an historical overview on the subject, we refer the reader to the review in Ref. [53]. Here, instead, we follow an approach similar to Ref. [59].

Mathematically, a beamsplitter is simply a unitary operator  $\hat{U}$  that mixes the path indices of the creation operators on which it is applied, while leaving the spectral polarisation modes unaltered. We can formally write this as [61]

$$\hat{U} \hat{a}_{i,g}^\dagger \hat{U}^\dagger = \sum_{\alpha} U_{\alpha,i} \hat{a}_{\alpha,g}^\dagger, \quad (2.40)$$

where  $U$  is the unitary matrix associated with  $\hat{U}$ , and  $\hat{U} |0\rangle = |0\rangle$ . An example of physical implementation is given by a half-silvered mirror, which reflects part of the light while letting the rest go through. A beamsplitter is said to be “balanced” if it equally splits the intensity of the incident light. This condition can be achieved, for example, by employing in Equation (2.40) the unitary matrix

$$U_{\text{BS}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad (2.41)$$

and we label the corresponding operator by  $\hat{U}_{\text{BS}}$ .

Let  $|f\rangle_{\mathcal{M}}$  and  $|g\rangle_{\mathcal{M}}$  be the spectral-polarisation modes of two photons in different spatial modes. If they evolve through a balanced beamsplitter, characterised by the unitary operation  $\hat{U}_{\text{BS}}$ , the final state of the system can be deduced from Equation (2.40).

It can be written as:

$$\begin{aligned} |\psi_{\text{out}}\rangle &= \hat{U}_{\text{BS}} \hat{a}_{1,f}^\dagger \hat{a}_{2,g}^\dagger |0\rangle = \frac{1}{2} \left( \hat{a}_{1,f}^\dagger - \hat{a}_{2,f}^\dagger \right) \left( \hat{a}_{1,g}^\dagger + \hat{a}_{2,g}^\dagger \right) |0\rangle \\ &= \frac{1}{2} \left( \hat{a}_{1,f}^\dagger \hat{a}_{1,g}^\dagger - \hat{a}_{2,f}^\dagger \hat{a}_{2,g}^\dagger \right) |0\rangle + \frac{1}{2} \left( \hat{a}_{1,f}^\dagger \hat{a}_{2,g}^\dagger - \hat{a}_{2,f}^\dagger \hat{a}_{1,g}^\dagger \right) |0\rangle, \end{aligned} \quad (2.42)$$

where we grouped together the terms leading to a “bunching” or to a “coincidence” event, where the photons exit respectively from the same side or from different ports. In this way, the probability of detecting a coincidence event is given by the norm of the second vector appearing in Equation (2.42), but its calculation is complicated by the non orthogonality between  $|f\rangle_{\mathcal{M}}$  and  $|g\rangle_{\mathcal{M}}$ . By taking advantage of the expansion in Equation (2.39), we can expand  $|f\rangle_{\mathcal{M}}$  in terms of  $|g\rangle_{\mathcal{M}}$  and  $|g^\perp\rangle_{\mathcal{M}}$ , thus obtaining

$$|\psi_{\text{out}}\rangle = \frac{1}{2} \left( \hat{a}_{1,f}^\dagger \hat{a}_{1,g}^\dagger - \hat{a}_{2,f}^\dagger \hat{a}_{2,g}^\dagger \right) |0\rangle + \frac{1}{2} \sqrt{1 - |\langle g|f\rangle_{\mathcal{M}}|^2} \left( \hat{a}_{1,g^\perp}^\dagger \hat{a}_{2,g}^\dagger - \hat{a}_{2,g^\perp}^\dagger \hat{a}_{1,g}^\dagger \right) |0\rangle, \quad (2.43)$$

where the two contributions proportional to  $\langle g|f\rangle_{\mathcal{M}}$  cancelled because  $[\hat{a}_{1,g}^\dagger, \hat{a}_{2,g}^\dagger] = 0$ . As the vector  $\hat{a}_{1,g^\perp}^\dagger \hat{a}_{2,g}^\dagger |0\rangle$  and its counterpart with exchanged indices  $1 \leftrightarrow 2$  are orthogonal, the probability of a coincidence event can now be easily calculated as

$$p_{11}(f, g) = \frac{1 - |\langle g|f\rangle_{\mathcal{M}}|^2}{2}. \quad (2.44)$$

From this, we see that when the two photons are perfectly indistinguishable, i.e., when  $\langle g|f\rangle_{\mathcal{M}} = 1$ , the probability of detecting a coincidence event goes to zero and the photons always “bunch” together and exit the beamsplitter from the same side.

In order to observe this behaviour in experiments, typically one of the two photons is delayed with the respect to the other by a certain time  $\tau$ , and the probability  $p_{11}$  of observing a coincident detection is measured. By sweeping the parameter  $\tau$  over a broad set of values, a dip in  $p_{11}$  can be observed for indistinguishable photons when  $\tau$  approaches zero. For example, consider two photons with a Gaussian frequency distribution

$$g(\omega, \lambda) = \delta_{\lambda, \lambda_0} \frac{1}{(\sqrt{2\pi}\sigma_\omega)^{1/2}} e^{-\frac{(\omega - \omega_0)^2}{4\sigma_\omega^2}}, \quad (2.45)$$

with central frequency  $\omega_0$ , variance  $\sigma_\omega^2$ , and polarisation  $\lambda_0$ . When one of the two is delayed by  $\tau$ , its spectral-polarisation mode is described by  $g_\tau(\omega, \lambda) = e^{-i\omega\tau} g(\omega, \lambda)$ , so that the overlap between the two modes becomes

$$\langle g|g_\tau\rangle_{\mathcal{M}} = e^{-\frac{1}{2}(\sigma_\omega\tau)^2}, \quad (2.46)$$

and the dip in coincidence events assumes the typical form shown in Figure 2.2. The minimum value of  $p_{11}$  experimentally observed can thus be used as an indication of the



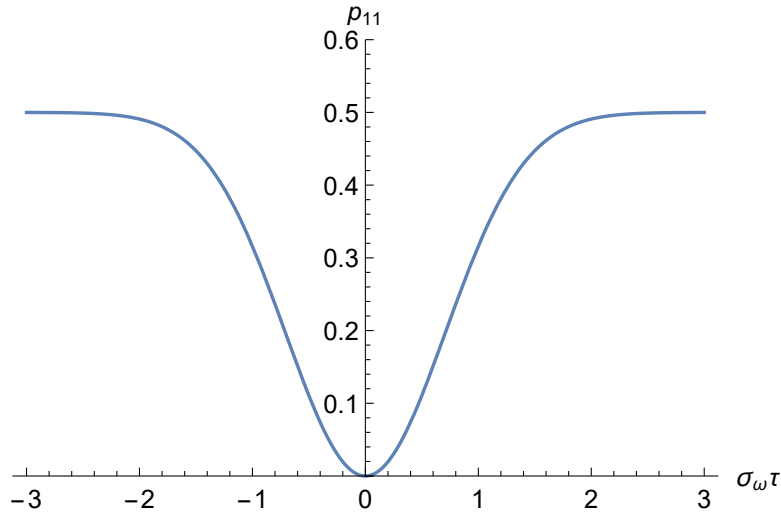


Figure 2.2: Typical HOM dip in the probability of observing a coincident detection, for photons with the same Gaussian wave-packet delayed by a time  $\tau$  with respect to each other.

distinguishability of the two photons.

### A different perspective

In order to appreciate the quantumness of this result, we will shortly make a comparison with a classical realisation of the same experiment. However, in order to do this, we firstly need to rewrite the condition  $p_{11} = 0$  in a more suitable form. The coincidence probability in Equation (2.44) could be formally written as the expectation value

$$p_{11} = \text{Tr} \left[ \hat{\Pi}_{11} |\psi_{\text{out}}\rangle \langle \psi_{\text{out}}| \right], \quad (2.47)$$

where  $\hat{\Pi}_{11} = |1\rangle_1 \langle 1| \otimes |1\rangle_2 \langle 1|$  represents the projector on the subspace with one photon per spatial mode. As the photons could be in any spectral-polarisation mode, each projector  $|1\rangle_i \langle 1|$  can be expanded as

$$|1\rangle_i \langle 1| = \sum_{x \in \mathcal{O}} \hat{a}_{i;x}^\dagger |0\rangle \langle 0| \hat{a}_{i;x}, \quad (2.48)$$

where the index  $x$  runs over any complete set  $\mathcal{O}$  of orthonormal modes. For example, one could choose  $\mathcal{O}$  to be the set of pairs  $(\omega, \lambda)$ , but this is not the only option. It is easily proven that the definition of  $|1\rangle_i \langle 1|$  is independent from the particular choice of basis, and the simplest way to recover Equation (2.44) is by choosing a set  $\mathcal{O}$  that includes the modes  $g$  and  $g^\perp$ .

At this stage, the measurement operator  $\hat{\Pi}_{11}$  is intrinsically quantum mechanical, because it is explicitly defined in terms of projectors onto single photon states. However, the expectation value defining  $p_{11}$  in Equation (2.47) does not change if we replace  $\hat{\Pi}_{11}$

with  $\hat{n}_1 \otimes \hat{n}_2$ , where  $\hat{n}_i$  counts the total number of photons in spatial mode  $i$ , for  $i = 1, 2$ :

$$\hat{n}_i = \sum_{x \in O} \hat{a}_{i;x}^\dagger \hat{a}_{i;x}. \quad (2.49)$$

This is because the output state is composed by two photons only, and  $\hat{n}_i$  acts exactly like  $|1\rangle_i\langle 1|$  on all single-photon states, and it yields zero when applied on a state with no photons on the  $i$ th spatial mode. Therefore, the HOM result can be alternatively stated by saying that the expectation value  $\langle \hat{n}_1 \hat{n}_2 \rangle \equiv \text{Tr} [\hat{n}_1 \otimes \hat{n}_2 \rho_{\text{out}}]$  can go to zero when two indistinguishable pure photons are injected on a balanced beamsplitter, where the output state of the system has been generically represented by the density matrix  $\rho_{\text{out}}$ .

As we discussed at the end of Section 2.2, for indistinguishable photons (i.e., characterised by the same spectral-polarisation mode of light) photon counting and intensity measurements are equivalent, up to a proportionality factor. Therefore, we can also express the HOM effect by stating that the expectation value  $\langle \hat{I}_1 \hat{I}_2 \rangle$  can reach zero, when evaluated on  $\rho_{\text{out}}$ . This is the formulation typically mentioned in the literature [52–54], because it is the well-suited to describe the detectors response according to photodetection theory [57, 58]. We decided to introduce this approach only on a second moment because we find more intuitive the formulation of Ref. [59], based on the evolution of creation operators. Moreover, in order to properly compare the light behaviour in the HOM effect with a classical scenario, it is convenient to introduce the correlation function

$$G_{12} = \frac{\langle \hat{I}_1 \hat{I}_2 \rangle}{\langle \hat{I}_1 \rangle \langle \hat{I}_2 \rangle}, \quad (2.50)$$

obtained by normalising  $\langle \hat{I}_1 \hat{I}_2 \rangle$  with the average light intensity detected at the output ports of the beamsplitter, i.e.,  $\langle \hat{I}_i \rangle = \text{Tr} [\hat{I}_i \rho_{\text{out}}]$  for  $i = 1, 2$ . As both  $\langle \hat{I}_1 \rangle$  and  $\langle \hat{I}_2 \rangle$  are non zero in the HOM setup, the condition  $\langle \hat{I}_1 \hat{I}_2 \rangle = 0$  can be equivalently expressed as  $G_{12} = 0$ .

Overall, the HOM result can be equivalently stated by saying that there exists an input state  $\rho_{\text{in}}$  and a unitary evolution  $\hat{U}_{BS}$ , acting linearly on the modes of the field [see Equation (2.40)], such that the correlation function  $G_{12}$  reaches zero when evaluated on the output state  $\rho_{\text{out}} = \hat{U}_{BS} \rho_{\text{in}} \hat{U}_{BS}^\dagger$ . In particular, the aforementioned input state can be written in the factorised form

$$\rho_{\text{in}} = \rho_1 \otimes \rho_2, \quad (2.51)$$

where  $\rho_\alpha = \hat{a}_{\alpha;g}^\dagger |0\rangle \langle 0| \hat{a}_{\alpha;g}$ , for  $\alpha = 1, 2$ , and the spectral-polarisation mode  $g$  is the same for both photons. Furthermore, we note that each  $\rho_\alpha$  is invariant under any phase transformation of the light mode  $g(\omega, \lambda) \rightarrow e^{-i\phi} g(\omega, \lambda)$ , for  $\phi \in [0, 2\pi]$ , because  $\hat{a}_{\alpha,ge^{-i\phi}}^\dagger = e^{-i\phi} \hat{a}_{\alpha,g}^\dagger$ . Should the reader be familiar with the phase space formulation of quantum optics, this is the well-known fact that the Wigner functions of Fock states are invariant

under phase rotations.

### 2.3.2 Counterpart in classical optics

An advantage of formulating the HOM effect in terms of  $G_{12}$  is that this expression has a natural classical counterpart, obtained by shining classical light fields on a beamsplitter and by looking at correlations between the output intensities. If this classical quantity was unable to reach zero under conditions compatible with the quantum HOM effect, the latter could be considered a signature of nonclassicality.

This is indeed what happens [52–54, 62], but in order to show it explicitly we first need to discuss which input classical states of light should be considered when looking for the minimum value of the classical counterpart of  $G_{12}$ . The usual assumptions are the following [52–54]:

1. The sources are independent;
2. The phase of the field emitted by each source fluctuates randomly.

It is important to stress that both these requirements are compatible with the structure of the single-photon sources used in the quantum case, which are independent and invariant under phase transformations [see Equation (2.51) and comments below]. The first condition is mostly introduced for the sake of simplicity, while the second one naturally arises in optical experiments. On top of these fundamental assumptions, most authors add some extra simplifying hypotheses. For example, Ref. [54] deals with identical pulses of light, although potentially delayed with respect to each other, while in Ref. [52] quasi monochromatic light beams are considered. For the time being, we take the following additional hypothesis:

3. The fields emitted by the sources are characterised by the same spectral-polarisation mode of light, and reach the detectors after travelling for the same optical length.

Notice that this condition is satisfied by the indistinguishable photons of the HOM experiment, and that it intuitively allows the light beams to maximally interfere, because their wave-packets perfectly overlap at the detectors.

In the remainder of this section we will show how these hypotheses enforce a non zero lower bound on the classical counterpart of Equation (2.50). Before doing this, however, we introduce the mathematical formalism needed to rigorously describe the fields emitted by the classical sources.

#### Details on the classical framework

As the considered classical sources are stochastic, we introduce two random variables  $\xi_1, \xi_2$ , whose values are in a one-to-one correspondence to all possible realisations of the

fields emitted by the sources, up to a random phase. More precisely, with probability  $p_\alpha(\xi_\alpha)$  the  $\alpha$ th source ( $\alpha = 1, 2$ ) emits a pulse of light characterised by the electric field

$$\mathbf{E}_\alpha^{(\xi_\alpha)}(z, t) = i \left( \frac{1}{2\epsilon_0 c \mathcal{A}} \right)^{\frac{1}{2}} A_\alpha(\xi_\alpha) \zeta_{g(\xi_\alpha)}(z - z_\alpha, t), \quad (2.52)$$

where  $z_\alpha$  is the source position, the phase of  $A_\alpha(\xi_\alpha)$  is uniformly distributed in  $[0, 2\pi]$  for any  $\xi_\alpha$ , and the function  $\zeta$  has been formally defined in Equation (2.16). Notice that at this stage we leave the emitted spectral-polarisation mode of light  $g(\xi_\alpha)$  free to change from one realisation to another, thus momentarily ignoring the constraint given by condition 3.

When the electric fields evolve through a beamsplitter, they are linearly mixed similarly to the annihilation operators in the quantum case. Therefore, from a comparison with Equation (2.40), it follows that the output field in path  $i = 1, 2$  can be written as

$$\mathbf{E}_i^{(\xi)}(z, t) = \sum_{\alpha=1}^2 U_{i\alpha} \mathbf{E}_\alpha^{(\xi_\alpha)}(z, t), \quad (2.53)$$

where  $\xi = \{\xi_1, \xi_2\}$ . More generally, with the same effort we can also take into account possible losses by substituting  $U$  with a generic complex matrix  $T$ . Therefore, if the detectors are in positions  $\bar{z}_i$ , for  $i = 1, 2$ , the fields at their locations at time  $t$  can be generally written as

$$\mathbf{E}_i^{(\xi)}(\bar{z}_i, t) = i \left( \frac{1}{2\epsilon_0 c \mathcal{A}} \right)^{\frac{1}{2}} \sum_{\alpha=1}^2 T_{i\alpha} A_\alpha(\xi_\alpha) \zeta_{g(\xi_\alpha)}(\bar{z}_i - z_\alpha, t). \quad (2.54)$$

Because of the aforementioned condition 3, in the situation of interest the modes  $|g(\xi_\alpha)\rangle_{\mathcal{M}}$  are all equal to the same  $|g\rangle_{\mathcal{M}}$ , and the distances  $\bar{z}_i - z_\alpha$  are independent from  $i$  and  $\alpha$ . Therefore, all functions  $\zeta_{g(\xi_\alpha)}(\bar{z}_i - z_\alpha, t)$  are the same and equal to  $\zeta_h(t)$ , where

$$h(\omega, \lambda) = g(\omega, \lambda) e^{+i\omega(\bar{z}_i - z_\alpha)}. \quad (2.55)$$

Thanks to Equations (2.25) and (2.26), the intensity measured by the  $i$ th detector is thus given by

$$I_i^{(\xi)} = \frac{\mathcal{E}_h}{\mathcal{A}} \sum_{\alpha, \beta=1}^2 T_{i\alpha} T_{i\beta}^* A_\alpha(\xi_\alpha) A_\beta^*(\xi_\beta), \quad (2.56)$$

where  $\mathcal{E}_h$  has been defined in Equation (2.27).

The classical counterpart of Equation (2.50) is obtained by looking at the correlations between the intensities measured by the detectors for different realisations of the sources.

Formally, we can define  $G_{12}^{(\text{cl})}$  as

$$G_{12}^{(\text{cl})} \equiv \frac{\langle I_1 I_2 \rangle}{\langle I_1 \rangle \langle I_2 \rangle}, \quad (2.57)$$

where now the brackets  $\langle O \rangle$  stand for the average over all possible vectors  $\{\xi\}$ , i.e.

$$\langle O \rangle \equiv \sum_{\xi_1} \sum_{\xi_2} p_1(\xi_1) p_2(\xi_2) O(\xi). \quad (2.58)$$

Note that in this expansion we used the independence of the sources guaranteed by condition 1. The evaluation of these averages for the intensities in Equation (2.56), instead, is greatly simplified by condition 2, which implies

$$\langle A_\alpha \rangle = \langle A_\alpha^2 \rangle = 0, \quad (2.59)$$

for any  $\alpha = 1, 2$ , because of the phase randomness of each  $A_\alpha$ . Therefore, one has

$$\langle I_1 \rangle = \frac{\mathcal{E}_h}{\mathcal{A}} \sum_{\alpha=1}^2 |T_{1\alpha}|^2 \langle |A_\alpha|^2 \rangle, \quad \langle I_2 \rangle = \frac{\mathcal{E}_h}{\mathcal{A}} \sum_{\alpha=1}^2 |T_{2\alpha}|^2 \langle |A_\alpha|^2 \rangle, \quad (2.60)$$

and a slightly more involved but straightforward calculation yields:

$$\begin{aligned} \langle I_1 I_2 \rangle &= \langle I_1 \rangle \langle I_2 \rangle + \frac{\mathcal{E}_h^2}{\mathcal{A}^2} \sum_{\alpha=1}^2 |T_{1\alpha}|^2 |T_{2\alpha}|^2 \left( \langle |A_\alpha|^4 \rangle - \langle |A_\alpha|^2 \rangle^2 \right) \\ &\quad + \frac{\mathcal{E}_h^2}{\mathcal{A}^2} \sum_{\alpha \neq \beta}^2 T_{1\alpha} T_{1\beta}^* T_{2\beta} T_{2\alpha}^* \langle |A_\alpha|^2 \rangle \langle |A_\beta|^2 \rangle. \end{aligned} \quad (2.61)$$

The second term appearing in the expression of  $\langle I_1 I_2 \rangle$  is always positive, because it represents the variance of the random variable  $|A_\alpha(\xi_\alpha)|^2$  characterising the intensity of the emitted field. Hence, the minimum  $G_{12}^{(\text{cl})}$  can be found by dropping it and minimising

$$G_{12}^{(\text{cl,nf})} = 1 + \frac{\sum_{\alpha \neq \beta}^2 T_{1\alpha} T_{1\beta}^* T_{2\beta} T_{2\alpha}^* \langle |A_\alpha|^2 \rangle \langle |A_\beta|^2 \rangle}{\left( \sum_{\gamma=1}^2 |T_{1\gamma}|^2 \langle |A_\gamma|^2 \rangle \right) \left( \sum_{\delta=1}^2 |T_{2\delta}|^2 \langle |A_\delta|^2 \rangle \right)}, \quad (2.62)$$

where the subscript “nf” reminds us that we are considering a situation where the sources emit fields with non fluctuating intensities. In what follows, we show that this value can never drop below 1/2, for any choice of average intensities  $\{\langle |A_\alpha|^2 \rangle\}_{\alpha=1}^2$  and any evolution matrix  $T$ .

### Minimisation of the classical correlation function

Because of the low number of parameters in the problem, we can explicitly perform this minimisation by parametrising the entries of  $T$  as:

$$T = \begin{pmatrix} t_{11}e^{i\theta_{11}} & t_{12}e^{i\theta_{12}} \\ t_{21}e^{i\theta_{21}} & t_{22}e^{i\theta_{22}} \end{pmatrix}, \quad (2.63)$$

with  $t_{i\alpha} \geq 0$  for  $i, \alpha = 1, 2$ . Note that this can always be done, and the result will hold with great generality. For example, the unitarity of the evolution matrix is not required, and the conclusion will not be affected by the loss of part of the emitted light. With this parametrisation,  $G_{12}^{(\text{cl})}$  becomes

$$G_{12}^{(\text{cl,nf})} = 1 + t_{\text{TOT}} \cos(\theta_{\text{TOT}}) \frac{2a_1a_2}{(t_{11}^2a_1 + t_{12}^2a_2)(t_{21}^2a_1 + t_{22}^2a_2)}, \quad (2.64)$$

where we defined

$$t_{\text{TOT}} \equiv t_{11}t_{12}t_{21}t_{22}, \quad \theta_{\text{TOT}} \equiv \theta_{11} - \theta_{12} - \theta_{21} + \theta_{22}, \quad (2.65)$$

and  $a_\alpha \equiv \langle |A_\alpha|^2 \rangle$  for simplicity.

In some papers on the subject (e.g., see Ref. [52]) only the simpler scenario of interference in the “far field” is discussed, where the weights  $t_{i\alpha}$  are all the same and equal to  $t$ . For example, this condition naturally arises when the sources emit spherical waves and are very far, and thus approximately equidistant, from the detectors. In this case, the desired bound is easily obtained from the inequality  $\cos(\theta_{\text{TOT}}) \geq -1$ :

$$G_{12}^{(\text{cl,nf})}|_{t_{i\alpha} \equiv t} = 1 + \cos(\theta_{\text{TOT}}) \frac{2a_1a_2}{(a_1 + a_2)^2} \geq \frac{1}{2} + \frac{1}{2} \frac{(a_1 - a_2)^2}{(a_1 + a_2)^2} \geq \frac{1}{2}. \quad (2.66)$$

Interestingly, this reasoning also shows that the value  $1/2$  can be reached by sources emitting light with the same, non fluctuating, intensity whenever a beamsplitter characterised by  $t_{i\alpha} \equiv t$  and  $\cos(\theta_{\text{TOT}}) = -1$  is used. Two examples satisfying these conditions are worth to keep in mind: a balanced beamsplitter  $\hat{U}_{\text{BS}}$  [see Equation (2.41)], and a balanced beamsplitter which follows a  $\pi$  phase rotation. The overall operator  $\hat{U}_{\text{FT}(2)}$  characterising this second evolution acts on creation and annihilation operators, as in Equation (2.40), via the unitary matrix

$$U_{\text{FT}(2)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (2.67)$$

The subscript FT(2) indicates that this corresponds to a 2-dimensional Fourier transform matrix. In the following we will see that the class of higher dimensional Fourier transform matrices play a similar role in a multiport interferometer.

The same lower bound  $G_{12}^{(\text{cl,nf})} \geq 1/2$  holds also without the “far field” hypothesis, as it is mentioned (but not explicitly proven) in Ref. [54]. However, a formal proof can be easily found. As in the previous case, for fixed  $a_1, a_2$  and  $\{t_{i\alpha}\}_{i,\alpha}$ , the condition  $\cos(\theta_{\text{TOT}}) = -1$  minimises the correlation function. Therefore, without loss of generality we can choose  $\theta_{\text{TOT}} = \pi$ . The idea now is to introduce four new non negative variables:

$$z_1 = t_{11}\sqrt{a_1}, \quad z_2 = t_{12}\sqrt{a_2}, \quad z_3 = t_{21}\sqrt{a_1}, \quad z_4 = t_{22}\sqrt{a_2}, \quad (2.68)$$

which can be used to rewrite the inequality  $G_{12}^{(\text{cl,nf})}|_{\theta_{\text{TOT}}=\pi} \geq 1/2$  as

$$4z_1z_2z_3z_4 \leq (z_1^2 + z_2^2)(z_3^2 + z_4^2). \quad (2.69)$$

This is equivalent to  $(z_1z_3 - z_2z_4)^2 + (z_1z_4 - z_2z_3)^2 \geq 0$ , and the proof is thus concluded. A parametrisation similar to Equation (2.68) will also be useful in Section 2.4, when we will consider a multimode scenario.

## 2.4 Multimode generalisation

We have now all the tools that we need to describe the main results of this chapter. Our aim is to set a tight lower bound on the amount of intensity correlations that could possibly be detected in output of a linear optical interferometer, whenever classical light fields are used as inputs. This threshold can then be used as a benchmark, whose violation in a quantum experiment certifies the impossibility of reproducing the same results by classical means. More precisely, a violation rules out any classical explanation satisfying the hypotheses that have been assumed in deriving the threshold itself. For this reason, and not only for the sake of precision, it is highly desirable to keep these hypotheses to a minimum. If we look at the conditions that were assumed in Section 2.3.2 to prove the classical bound for the HOM setup, we recognise that assumptions 1 and 2 characterise the structure of the sources (independence and phase randomness), and cannot be easily removed. However, the situation is different for hypothesis 3, which only assures that the wave packets characterising the photons emitted by different sources can perfectly overlap at the detectors. This assumption greatly simplified the analysis, but it is reasonable to expect that a lower “amount” of interference cannot help in reducing the minimum achievable value of intensity correlations. It turns out that this is indeed the case, and we can prove the aforementioned classical threshold by using only assumptions 1 and 2. As the proof is rather technical, it has been postponed to the end of this section. For the time being, we will find and discuss the classical lower bound on intensity correlations by also assuming the aforementioned condition 3. The reader should keep in mind, though, that this has been done just for the sake of simplicity, in order to emphasise the main physical ideas over the technical subtleties.

Of course, in order to talk about an amount of intensity correlations in a multimode setting, we need a suitable correlation measure. This is an important choice, because all our results, and even the possibility of analytically obtaining a classical threshold, depend on it. For consistency with the two-mode scenario previously discussed, we want to generalise the quantity  $G_{12}$  that has been introduced in Equation (2.50) and Equation (2.57) respectively in the quantum and classical case. One option is to look at higher-order correlation functions, in which the products of more than two intensities appear in the expectation values. This approach has been recently adopted by several authors to study multiparticle interference [47, 63] or to obtain advantages in imaging resolution [64–67]. Unfortunately, as it usually happens when studying the evolution of photons through multimode interferometers, these correlation functions depend on permanents of matrices with dimension larger than or equal to 3. Although they could be easily evaluated numerically, their presence greatly increases the difficulty of performing analytical manipulations, because of the complexity of their expressions. On the other hand, a different approach has been taken by Walschaers and coworkers, who considered the average over many output pairs of modes  $i \neq j$  of the simpler quantities  $\langle \hat{n}_i \hat{n}_j \rangle - \langle \hat{n}_i \rangle \langle \hat{n}_j \rangle$ . In particular, they showed that this average can yield information on the statistics of the interfering quantum particles [68] and on their distinguishability [69]. However, the classical counterpart of those differences can be sent to zero simply by considering sources of light with vanishing intensity. Therefore, in that form the correlation function studied by Walschaers *et al.* is not suitable for our purposes.

In order to address these issues, in the following we study a similar, but different, quantifier. If there are  $N_d$  detectors, we can define the normalised average

$$\overline{G} = \frac{1}{\binom{N_d}{2}} \sum_{i < j} \frac{\langle \hat{I}_i \hat{I}_j \rangle}{\langle \hat{I}_i \rangle \langle \hat{I}_j \rangle}, \quad (2.70)$$

where the constraint  $i < j$  enforces the sum to be over all pairs of detectors. Similarly to  $G_{12}$ , the quantity  $\overline{G}$  only depends on the correlation between intensities, and does not change if the electric fields emitted by the sources are multiplied by a common factor. Moreover, with respect to considering higher-order correlators, it has the advantage of being composed by many simple contributions, while still taking into account the information provided by each detector. For these reasons,  $\overline{G}$  is a promising measure of intensity correlations, whose classical counterpart can be lower bounded through an optimisation over all linear optical evolutions and over all sources satisfying conditions 1 and 2 (see Section 2.3.2).

*Remark:* It should be stressed that only detectors characterised by  $\langle \hat{I}_i \rangle \neq 0$  should be considered in Equation (2.70), in order to keep  $\overline{G}$  well defined. This is not a significant restriction, because a detector which does not receive any light from the apparatus can be dropped, and perhaps used to monitor another output spatial mode. More



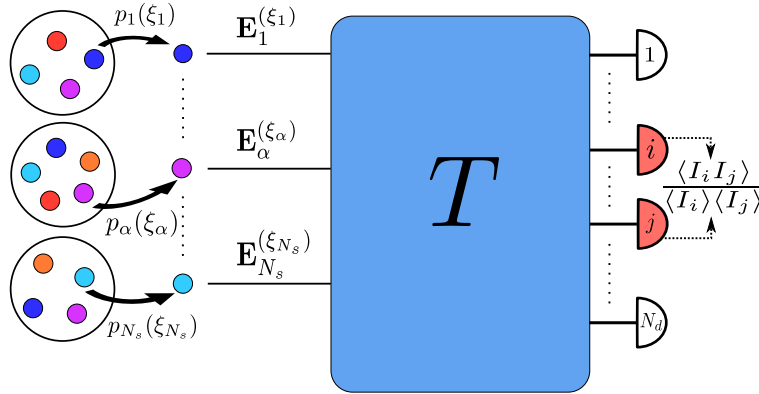


Figure 2.3: Sketch of the studied classical interferometric setup, with  $N_s$  sources and  $N_d$  detectors. The input coloured circles represent the light pulses with random phases  $\{\mathbf{E}_\alpha^{(\xi_\alpha)}(z, t)\}$ , which are emitted with probabilities  $\{p(\xi_\alpha)\}$  by independent sources. The evolution of the complex input fields through the interferometer is described by the complex transfer matrix  $T$ , and intensity measurements are performed in output.

importantly, as measurements of intensity and photon number are proportional for the quasi-monochromatic photons typically used in experiments, in order to calculate the correlation function in Equation (2.70) the detectors need to be able to resolve different photon numbers.

The remainder of this section is organised as follows. At first, in Section 2.4.1 we find a tight lower bound on the amount of intensity correlation in a classical framework, as measured by  $\overline{G}$ . After that, in Section 2.4.2 we show how this bound could be violated by exploiting quantum states of light, and we identify a necessary condition on the sources that needs to be satisfied in order for this to be possible. On top of the theoretical and fundamental interest of our results, in Section 2.4.3 we discuss some possible applications, concerning the possibility of characterising some aspects of the sources or of the interferometer from a measurement of  $\overline{G}$ . Then, as we anticipated, in Section 2.4.4 we formally prove that our classical lower bound on  $\overline{G}$  cannot be affected by the removal of the third assumption on the considered classical framework [see Section 2.3.2]. In other words, we prove that the minimal amount of correlations between output intensities is obtained when the sources emit pulses of light characterised by the same spectral-polarisation mode, travelling the same optical length before being detected.

### 2.4.1 Tight classical lower bound

The classical framework that we consider in this section is a straightforward multimode generalisation of the two-mode classical scenario previously discussed, and a pictorial representation can be seen in Figure 2.3. The light emitted by  $N_s$  independent and stochastic sources evolves through a linear optical interferometer and is monitored by  $N_d$  detectors performing intensity measurements. The interferometer generalises the beamsplitter of

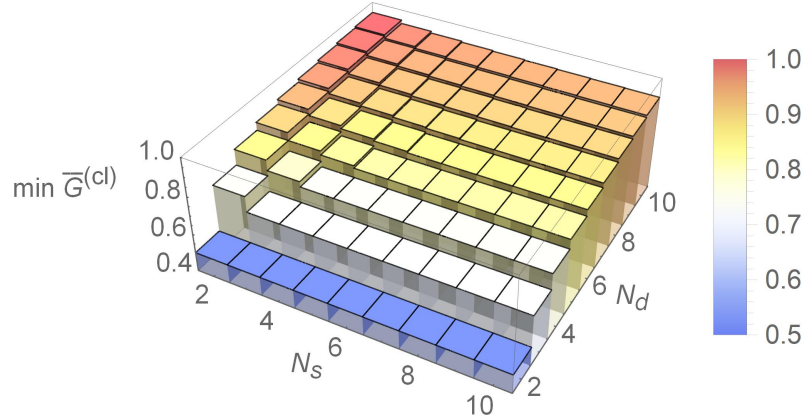


Figure 2.4: Classical lower bound on the amount of correlations among pairs of output intensities, measured by  $\overline{G}^{(cl)}$ .

the HOM setup, and is characterised by a complex matrix  $T$ , not necessarily unitary, which describes the evolution of the fields similarly to Equation (2.53). A special role will be played by the  $d$ -dimensional “Fourier interferometer”, defined by  $T = U_{FT(d)}$ , where

$$[U_{FT(d)}]_{i\alpha} \equiv \frac{1}{\sqrt{d}} \omega_d^{(i-1)(\alpha-1)}, \quad (2.71)$$

with  $\omega_d = e^{i\frac{2\pi}{d}}$ . The data measured by the detectors can be used to evaluate the classical analogous of the correlation function  $\overline{G}$ :

$$\overline{G}^{(cl)} = \frac{1}{\binom{N_d}{2}} \sum_{i < j} \frac{\langle I_i I_j \rangle}{\langle I_i \rangle \langle I_j \rangle}. \quad (2.72)$$

The multimode generalisation of the two-mode classical bound  $G_{12}^{(cl)} \geq 1/2$  is given by the following theorem.

**Theorem 2.1.** *In a classical framework satisfying the assumptions 1, 2, 3 discussed in Section 2.3.2, the amount of intensity correlation  $\overline{G}^{(cl)}$  found in output of a linear optical interferometer is lower bounded by a function of the number of sources ( $N_s$ ) and detectors ( $N_d$ ):*

$$\overline{G}^{(cl)} \geq \begin{cases} 1 - \frac{N_s - 1}{N_s(N_d - 1)} & \text{if } N_s < N_d, \\ 1 - \frac{1}{N_d} & \text{if } N_s \geq N_d. \end{cases} \quad (2.73)$$

Furthermore, this bound can be saturated by classical light fields with the same intensities, evolving through a  $N_d \times N_d$  Fourier interferometer (if  $N_s > N_d$  this requires to ignore  $N_s - N_d$  sources).

The remainder of this section will be completely devoted to prove this result. Before entering into the details of the proof, however, let us make a few comments.

- At first, notice that the lower bound is consistent with the two-mode case, because it reduces to  $1/2$  when  $N_s = N_d = 2$ . Another expected feature is the fact that the lower bound is smaller than 1, because  $\overline{G}^{(\text{cl})} = 1$  could be trivially obtained by considering non interfering light beams. For example, this condition could be realised by introducing large time delays between the light pulses emitted by different sources, with respect to the time-width of a single pulse. Needless to say, the detectors should remain in function until the energy of the last pulse has been completely measured.
- If the number of detectors increases, the minimum attainable value of intensity correlation becomes closer to 1, but an opposite behaviour arises if the number of sources increases. However, when  $N_s$  becomes larger than  $N_d$  the lower bound stabilises and is not affected by the addition of new sources. A plot of the lower bound given in Equation (2.73) can be observed in Figure 2.4. The monotonic dependence on  $N_s$  can be easily understood. Even if  $N_s$  increases, we can always recover a setup with a smaller number of sources  $N_s'$  simply by choosing vanishing intensities for the pulses emitted by  $N_s - N_s'$  sources. Therefore, the lower bound on  $\overline{G}^{(\text{cl})}$  has to be non-increasing with  $N_s$ . The same reasoning does not apply to the detectors: we cannot move some of them in a region where they do not receive any light, because  $\langle I_i \rangle$  should be non zero for any  $i = 1, \dots, N_d$ . Indeed, the monotonicity of the lower bound in  $N_d$ , and the stabilisation of its value for  $N_s > N_d$ , are new information provided by Theorem 2.1, that could not be easily deduced a priori by means of general considerations on the setup.
- The ideal interferometer, which can lead to the minimum value of  $\overline{G}^{(\text{cl})}$  in presence of sources with the same intensities, depends only on the number of detectors. A pictorial representation of the ideal setup can be observed in Figure 2.5 for a number of sources smaller or larger than  $N_d$ . In order to minimise the intensity correlation, for  $N_s > N_d$  it is more advantageous to completely ignore the light originating from  $N_s - N_d$  sources, rather than considering a larger interferometer with some unmonitored outputs. This can be seen from the  $N_d \times N_d$  dimensionality of the optimal interferometer: the light of  $N_s - N_d$  sources cannot evolve towards the detectors, and is effectively ignored by the configuration which saturates the bound.
- The reason why a Fourier interferometer is optimal can be intuitively understood by realising that it leads to a high degree of interference. This is because the input intensities are equally split among the output modes, and the phases applied to the fields are uniformly spaced between 0 and  $2\pi$ . We point out that this is not the only interferometer leading to the minimum given in Equation (2.73). For example, any other setting obtained by permuting the columns or the rows of a Fourier interferometer, or by adding a phase rotation to one of the input or output modes, leads to

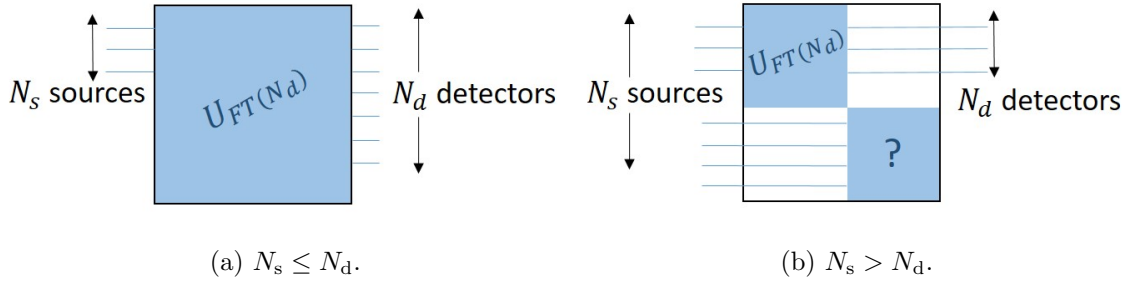


Figure 2.5: When the classical sources emit pulses of light with the same intensities and the same spectral-polarisation modes, an  $N_d$ -dimensional Fourier interferometer saturates the lower bound on  $\overline{G}^{(\text{cl})}$  found in Theorem 2.1 for any  $N_s, N_d$ . When there are more sources than detectors,  $N_s - N_d$  sources are effectively ignored.

the same result. Indeed, these phase shifts cannot contribute to  $\overline{G}^{(\text{cl})}$ , because the sources emit light with a random phase, and only intensity measurements are performed. However, the unicity of the optimal interferometer up to these permutations and phase transformations remains an open question at this stage.

### A short reminder on the notation

To ease the interpretation of the equations, we adopt the following convention. Greek letters (e.g.,  $\alpha, \beta = 1, \dots, N_s$ ) run over the spatial modes of the sources, while Roman indices (e.g.,  $i, j = 1, \dots, N_d$ ) identify the spatial modes of the detectors. We point out that with this notation  $\mathbf{E}_{\alpha=1}$  is the complex representation of the electric field emitted by the first source, while  $\mathbf{E}_{i=1}$  is the field detected by the first detector. The two are generally different from each other, because each detector receives the light coming from many different sources. Although it is important to keep in mind this distinction, we will never refer to a particular source or detector, so that no confusion should arise. As already implicitly done in Section 2.3.2, it is also convenient to identify each source with the spatial mode where it emits light. In this way, we can say that with probability  $p_\alpha(\xi_\alpha)$  the  $\alpha$ th source emits a pulse of light characterised by the complex electric field  $\mathbf{E}_\alpha^{(\xi_\alpha)}$ , defined exactly as in Equation (2.52), where  $\xi_\alpha$  is a random variable.

The multimode extension of a beamsplitter is a generic linear optical interferometer, that changes the spatial modes of the classical electric fields similarly to Equation (2.53):

$$\mathbf{E}_i^{(\boldsymbol{\xi})}(z, t) = \sum_{\alpha=1}^{N_s} T_{i\alpha} \mathbf{E}_\alpha^{(\xi_\alpha)}(z, t), \quad (2.74)$$

where  $T_{i\alpha}$  are generic complex numbers and the vector  $\boldsymbol{\xi} \equiv \{\xi_\alpha\}_\alpha$  contains all the information on the fields emitted by the sources in a particular realisation. Notice that no additional assumption is made on the matrix  $T$ : as it is not necessarily unitary, a lossy scenario where some of the emitted light is lost and never reaches the detectors is

automatically taken into account.

As in Section 2.3.2, we use the symbols  $z_\alpha$  and  $\bar{z}_i$  to indicate respectively the coordinates of the  $\alpha$ th source and  $i$ th detector, so that at time  $t$  the electric field on the  $i$ th detector is given by

$$\mathbf{E}_i^{(\xi)}(\bar{z}_i, t) = i \left( \frac{1}{2\epsilon_0 c \mathcal{A}} \right)^{\frac{1}{2}} \sum_{\alpha=1}^{N_s} T_{i\alpha} A_\alpha(\xi_\alpha) \boldsymbol{\zeta}_{g(\xi_\alpha)}(\bar{z}_i - z_\alpha, t). \quad (2.75)$$

As we can see, the component of the field reaching  $i$ th detector that originated from the  $\alpha$ th source has a temporal dependence given by  $\boldsymbol{\zeta}_{g(\xi_\alpha)}(\bar{z}_i - z_\alpha, t)$ . In other words, the mode  $|h_{i\alpha}(\xi_\alpha)\rangle_{\mathcal{M}}$  characterising its ability to interfere with other fields can be decomposed in spectral-polarisation coordinates as follows [see Equations (2.16) and (2.35)]:

$$\langle \omega, \lambda | h_{i\alpha}(\xi_\alpha) \rangle_{\mathcal{M}} = \langle \omega, \lambda | g(\xi_\alpha) \rangle_{\mathcal{M}} e^{+i\omega(\bar{z}_i - z_\alpha)/c}. \quad (2.76)$$

As we anticipated, for most of this chapter we will assume that the vectors  $|h_{i\alpha}(\xi_\alpha)\rangle_{\mathcal{M}}$  are all the same and equal to  $|h\rangle_{\mathcal{M}}$ , and we will find the minimum of  $\bar{G}^{(\text{cl})}$  under this hypothesis. Later on, in Section 2.4.4, we will relax this assumption and we will show that the minimum value of intensity correlation remains unaffected.

### Evaluation of the classical correlation function

The expectation values  $\langle I_i \rangle$ ,  $\langle I_j \rangle$  and  $\langle I_i I_j \rangle$  in the definition of  $\bar{G}^{(\text{cl})}$  can be evaluated as in Section 2.3.2, and the results are very similar:

$$\langle I_i \rangle = \frac{\mathcal{E}_h}{\mathcal{A}} \sum_{\alpha=1}^{N_s} |T_{i\alpha}|^2 \langle |A_\alpha|^2 \rangle, \quad \langle I_j \rangle = \frac{\mathcal{E}_h}{\mathcal{A}} \sum_{\beta=1}^{N_s} |T_{j\beta}|^2 \langle |A_\beta|^2 \rangle, \quad (2.77)$$

$$\begin{aligned} \langle I_i I_j \rangle &= \langle I_i \rangle \langle I_j \rangle + \frac{\mathcal{E}_h^2}{\mathcal{A}^2} \sum_{\alpha=1}^{N_s} |T_{i\alpha}|^2 |T_{j\alpha}|^2 \left( \langle |A_\alpha|^4 \rangle - \langle |A_\alpha|^2 \rangle^2 \right) \\ &\quad + \frac{\mathcal{E}_h^2}{\mathcal{A}^2} \sum_{\alpha \neq \beta}^{N_s} T_{i\alpha} T_{i\beta}^* T_{j\beta} T_{j\alpha}^* \langle |A_\alpha|^2 \rangle \langle |A_\beta|^2 \rangle, \end{aligned} \quad (2.78)$$

where  $\mathcal{E}_h$  is the energy carried by each pulse of light, as defined in Equation (2.27). Once again, the second term in Equation (2.78) is always positive, and we can safely ignore its contribution when looking for the minimum value of  $\bar{G}^{(\text{cl})}$ . This choice corresponds to a situation where the sources emit light pulses with a non fluctuating intensity, and we will keep track of this condition by adding a “nf” superscript (or subscript) to the quantities involved. Moreover, in this regime we can replace  $\langle |A_\alpha|^2 \rangle$  with  $|A_\alpha|^2$ .

Because of the size of the problem, parametrising the evolution matrix as in Equation (2.63) would not be feasible. However, this does not mean that we cannot be inspired

by the strategy used to perform that simple optimisation. In particular, in the two-mode case it has been useful to introduce new variables obtained by multiplying the components of the evolution matrix  $T$  with the square roots of the average intensities of the sources. Here we use the same idea, although with a different normalisation, and we introduce a set of  $N_d$  unit vectors  $|\psi_i\rangle \in \mathbb{C}^{N_s}$  with components

$$\langle \alpha | \psi_i \rangle_{\mathbb{C}^{N_s}} \equiv T_{i\alpha}^* \sqrt{\frac{\mathcal{E}_h |A_\alpha|^2}{\mathcal{A} \langle I_i \rangle}}, \quad (2.79)$$

where  $\{|\alpha\rangle_{\mathbb{C}^{N_s}}\}_{\alpha=1}^{N_s}$  is an orthonormal basis of  $\mathbb{C}^{N_s}$ . The Dirac notation is used only as a matter of convenience, and it should be stressed that there is nothing intrinsically quantum mechanical in these quantities. However, this notation will be extremely helpful in identifying the linear algebraic tools that will allow us to prove the desired lower bound. Thanks to Equation (2.79), we can rewrite the ratio  $\langle I_i I_j \rangle / [\langle I_i \rangle \langle I_j \rangle]$  for sources with fixed intensity as follows:

$$\left. \frac{\langle I_i I_j \rangle}{\langle I_i \rangle \langle I_j \rangle} \right|_{\text{nf}} = 1 + \sum_{\alpha \neq \beta}^{N_s} \langle \psi_j | \beta \rangle \langle \beta | \psi_i \rangle \langle \psi_i | \alpha \rangle \langle \alpha | \psi_j \rangle, \quad (2.80)$$

where we dropped the label  $\mathbb{C}^{N_s}$  for simplicity. We can now exploit the completeness of the set  $\{|\alpha\rangle_{\mathbb{C}^{N_s}}\}_{\alpha=1}^{N_s}$  and write the identity operator of  $\mathbb{C}^{N_s}$  as

$$\mathbb{1}_{\mathbb{C}^{N_s}} = \sum_{\alpha=1}^{N_s} |\alpha\rangle_{\mathbb{C}^{N_s}} \langle \alpha|. \quad (2.81)$$

By using this decomposition in Equation (2.80), and by averaging over all pairs of output modes,  $\overline{G}^{(\text{cl,nf})}$  can be written in a compact form as

$$\overline{G}^{(\text{cl,nf})} = 1 + \frac{1}{N_d(N_d - 1)} \sum_{i \neq j}^{N_d} \left[ |\langle \psi_i | \psi_j \rangle|^2 - \sum_{\alpha=1}^{N_s} |\langle \alpha | \psi_i \rangle|^2 |\langle \alpha | \psi_j \rangle|^2 \right]. \quad (2.82)$$

We will now proceed to prove that this quantity is larger than or equal to the right hand side of Equation (2.73), which is therefore proved because  $\overline{G}^{(\text{cl})} \geq \overline{G}^{(\text{cl,nf})}$ .

### Proof of the classical lower bound

We start by proving two linear algebraic inequalities.

**Lemma 2.1.** *Given a  $N_s \times N_s$  Hermitian matrix  $H$ , and the orthonormal basis with*

elements  $|\alpha\rangle_{\mathbb{C}^{N_s}}$  for  $\alpha = 1, \dots, N_s$ , the following inequality holds:

$$\sum_{\alpha=1}^{N_s} \langle \alpha | H | \alpha \rangle_{\mathbb{C}^{N_s}}^2 \leq \text{Tr}_{\mathbb{C}^{N_s}} [H^2]. \quad (2.83)$$

*Proof.* This follows from Equation (2.81) and from the expansion of the trace operator in basis  $\{|\alpha\rangle_{\mathbb{C}^{N_s}}\}_{\alpha=1}^{N_s}$ :

$$\begin{aligned} \sum_{\alpha=1}^{N_s} \langle \alpha | H | \alpha \rangle^2 &= \sum_{\alpha=1}^{N_s} \text{Tr} [|\alpha\rangle \langle \alpha| H |\alpha\rangle \langle \alpha| H] \\ &= \sum_{\alpha, \beta=1}^{N_s} \text{Tr} [|\alpha\rangle \langle \alpha| H |\beta\rangle \langle \beta| H] - \sum_{\alpha \neq \beta}^{N_s} \text{Tr} [|\alpha\rangle \langle \alpha| H |\beta\rangle \langle \beta| H] \\ &= \text{Tr} [H^2] - \sum_{\alpha \neq \beta}^{N_s} |\langle \alpha | H | \beta \rangle|^2 \leq \text{Tr} [H^2], \end{aligned} \quad (2.84)$$

where we dropped the label  $\mathbb{C}^{N_s}$  for simplicity.  $\square$

**Lemma 2.2.** *Given a Hermitian positive semidefinite matrix  $H$ , with rank  $r(H)$ , the following inequality holds:*

$$\text{Tr} [H^2] \geq \frac{\text{Tr} [H]^2}{r(H)}. \quad (2.85)$$

*Proof.* Let  $\rho$  be the density matrix defined as  $H/\text{Tr} [H]$ , and notice that Equation (2.85) can be rewritten as  $\text{Tr} [\rho^2] \geq 1/r(H)$ . We can interpret the rank of the matrix as the effective dimensionality of the system. In this way, a maximally mixed state  $\rho_{\text{mix}}$  defined in a  $r(H)$ -dimensional Hilbert space would have purity  $\text{Tr} [\rho_{\text{mix}}^2] = 1/r(H)$ . Therefore, the thesis of this lemma is equivalent to the well known fact that a completely mixed state, with density matrix multiple of the identity, is the quantum state with minimal purity.  $\square$

With these tools, we can manipulate Equation (2.82). At first, the sum over  $i \neq j$  can be split into two terms: one that runs over all  $i, j$ , and a negative contribution with  $i = j$ . This yields

$$\begin{aligned} N_d(N_d - 1) \left( \overline{G}^{(\text{cl}, \text{nf})} - 1 \right) &= \text{Tr} [H^2] - \sum_{\alpha}^{N_s} \langle \alpha | H | \alpha \rangle^2 - N_d \\ &\quad + N_d \sum_{\alpha=1}^{N_s} \sum_{i=1}^{N_d} \frac{1}{N_d} \left( \langle \alpha | [|\psi_i\rangle \langle \psi_i|] | \alpha \rangle \right)^2, \end{aligned} \quad (2.86)$$

where we introduced the  $N_s \times N_s$  Hermitian positive semidefinite matrix

$$H \equiv \sum_{i=1}^{N_d} |\psi_i\rangle \langle \psi_i|. \quad (2.87)$$

Then, we exploit the convexity of the square function in order to lower bound the last term appearing in Equation (2.86) as

$$\begin{aligned} N_d \sum_{\alpha=1}^{N_s} \sum_{i=1}^{N_d} \frac{1}{N_d} \left( \langle \alpha | [ |\psi_i\rangle \langle \psi_i| ] | \alpha \rangle \right)^2 &\geq N_d \sum_{\alpha=1}^{N_s} \left( \langle \alpha | \left[ \sum_{i=1}^{N_d} \frac{1}{N_d} |\psi_i\rangle \langle \psi_i| \right] | \alpha \rangle \right)^2 \\ &= \frac{1}{N_d} \sum_{\alpha=1}^{N_s} \langle \alpha | H | \alpha \rangle^2. \end{aligned} \quad (2.88)$$

The application of Lemma 2.1 to the sum of the right hand side obtained here with the second term appearing in Equation (2.86) yields

$$\overline{G}^{(\text{cl,nf})} \geq 1 + \frac{1}{N_d - 1} \left[ \frac{1}{N_d^2} \text{Tr} [H^2] - 1 \right]. \quad (2.89)$$

At this point, we can apply Lemma 2.2, and observe that  $\text{Tr} [H] = N_d$  and  $r(H) \leq \min\{N_s, N_d\}$ , as it can be seen from the definition of  $H$  in Equation (2.87). In conclusion, we are left with

$$\overline{G}^{(\text{cl,nf})} \geq 1 - \frac{\min\{N_d, N_s\} - 1}{\min\{N_d, N_s\}(N_d - 1)} = \begin{cases} 1 - \frac{N_s - 1}{N_s(N_d - 1)} & \text{if } N_s < N_d, \\ 1 - \frac{1}{N_d} & \text{if } N_s \geq N_d, \end{cases} \quad (2.90)$$

which is the desired result.

### Attainability of the lower bound

We are left with the task of showing that the lower bound on  $\overline{G}^{(\text{cl})}$  can be saturated when the conditions given in Theorem 2.1 are met. In particular, the intensities of the sources should not fluctuate from one realisation to another, in order to avoid the additional positive contribution in Equation (2.78) that originates from the intensities variances. Furthermore, the optimal interferometer should be characterised by  $T_{i\alpha} = [U_{\text{FT}(N_d)}]_{i\alpha} = \frac{1}{\sqrt{d}} \omega_{N_d}^{(i-1)(\alpha-1)}$ , where  $i, \alpha = 1, \dots, N_d$  and  $\omega_{N_d} = e^{i2\pi/N_d}$ . If  $N_s > N_d$ , this restriction means that the light emitted by  $N_s - N_d$  sources cannot enter the interferometer and is completely lost. We can effectively implement this by setting  $|A_\alpha| \equiv 0$  for any  $\alpha$  such that  $N_d < \alpha \leq N_s$ . Moreover, according to Theorem 2.1, the remaining sources should all have the same intensity  $|A_\alpha| \equiv |A|^2$ . All these conditions lead to the average output intensities

$$\langle I_i \rangle = \frac{\mathcal{E}_h \min\{N_s, N_d\}}{\mathcal{A} N_d} |A|^2, \quad (2.91)$$



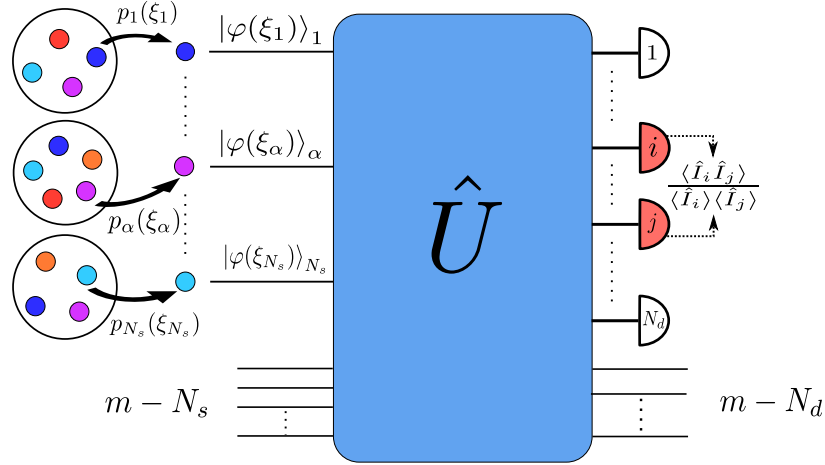


Figure 2.6: Sketch of the studied quantum interferometric setup, with  $N_s$  sources and  $N_d$  detectors. The input coloured circles represent the quantum states  $\{|\varphi(\xi_\alpha)\rangle_\alpha\}$ , which are emitted with probabilities  $\{p(\xi_\alpha)\}$  by independent sources with random phases. They evolve through an  $m$ -mode linear optical interferometer  $\hat{U}$ , with  $m \geq N_s, N_d$ : if these inequalities are strict there will be vacuum inputs and non monitored outputs.

and Equation (2.79) then yields

$$\langle \alpha | \psi_i \rangle_{\mathbb{C}^{N_s}} = \begin{cases} \frac{1}{\sqrt{\min\{N_s, N_d\}}} \omega_{N_d}^{-(i-1)(\alpha-1)} & \text{if } \alpha \leq N_d, \\ 0 & \text{if } \alpha > N_d. \end{cases} \quad (2.92)$$

Notice that  $\alpha$  could be larger than  $N_d$  only if the number of sources exceeds that of detectors. By inserting this expression into the definition of  $\overline{G}^{(\text{cl}, \text{nf})}$  given in Equation (2.82), we recover a value of intensity correlation which saturates the lower bound in Theorem 2.1.

### 2.4.2 Quantum violation of the classical threshold

Although the previous section was quite long and technical, the message it wants to convey should be clear: given a few conditions on the structure of the sources, the range of correlation values achievable in a classical optical framework is lower bounded by a known function of the number of sources and detectors. However, nothing precludes the possibility of violating this classical threshold by employing quantum states of light, as the original HOM effect highlights. In what follows, we show how some of the features of the latter can be generalised to a multimode setting. In particular, at first we describe the quantum setup analogous to the classical multimode framework just studied. Then, we identify a necessary condition that sources emitting photons with the same spectral-polarisation mode need to satisfy in order to violate the classical threshold previously identified, and we quantify the maximum violation that could be observed in a particular symmetric configuration.

## Quantum setting

The quantum setting we consider is generically represented in Figure 2.6, and we can see that it closely resembles its classical counterpart of Figure 2.3. As in the classical case, we consider independent sources that emit fields with a global random phase. Therefore, for  $\alpha = 1, \dots, N_s$ , with probability  $p_\alpha(\xi_\alpha)$  the  $\alpha$ th source emits the following quantum state in the  $\alpha$ th spatial mode:

$$|\varphi(\xi_\alpha)\rangle_\alpha = \sum_{n=0}^{\infty} \frac{\varphi_{\xi_\alpha}(n)}{\sqrt{n!}} \left[ e^{-i\phi(\xi_\alpha)} \hat{a}_{\alpha;g(\xi_\alpha)}^\dagger \right]^n |0\rangle, \quad (2.93)$$

where  $\hat{a}_{\alpha;g(\xi_\alpha)}^\dagger$  creates a photon with spectral-polarisation mode  $|g(\xi_\alpha)\rangle_{\mathcal{M}}$ , and  $\phi(\xi_\alpha)$  is the random phase of the field. Note that this must not be confused with the global phase of the quantum state, because  $e^{i\phi(\xi_\alpha)}$  multiplies the annihilation operator of the field, in analogy with the classical case where it multiplies the amplitude  $A_\alpha$ . This implies that the coherences between different photon numbers disappear when we average over this random phase  $\phi(\xi_\alpha)$ . For simplicity, we further assume that the modes  $|g(\xi_\alpha)\rangle_{\mathcal{M}}$  are all the same and identical to a given spectral-polarisation mode  $|h\rangle_{\mathcal{M}}$ , i.e., that all the considered quantum states are composed by indistinguishable photons. We remind the reader that our goal here is to show that the classical threshold on the amount of output intensity correlation can be violated in a quantum setting. It is, therefore, reasonable to explicitly consider a situation where all photons are indistinguishable, because this intuitively maximises their interference. With this choice, the average emitted state can be represented by the following density matrix:

$$\rho = \bigotimes_{\alpha=1}^{N_s} \sum_{n_\alpha} \frac{q_\alpha(n_\alpha)}{n_\alpha!} (\hat{a}_{\alpha;h}^\dagger)^{n_\alpha} |0\rangle \langle 0| (\hat{a}_{\alpha;h})^{n_\alpha}, \quad (2.94)$$

where  $q(\mathbf{n}) = q_1(n_1) \dots q_{N_s}(n_{N_s})$  is the effective probability distribution of the process, with

$$q_\alpha(n_\alpha) = \sum_{\xi_\alpha} p_\alpha(\xi_\alpha) |\varphi_{\xi_\alpha}(n_\alpha)|^2. \quad (2.95)$$

In order to characterise the evolution of the fields, we consider a generic  $m$ -mode linear optical interferometer, with  $m \geq N_s, N_d$ . The output state of the system can be written as

$$\rho_{\text{out}} = \hat{U} \rho \hat{U}^\dagger, \quad (2.96)$$

where  $\rho$  is the input state. Similarly to Equation (2.40), the operator  $\hat{U}$  associated with the interferometer alters the spatial mode of the creation operators while leaving their

spectral-polarisation mode unaltered

$$\hat{U} \hat{a}_{i;h}^\dagger \hat{U}^\dagger = \sum_{\alpha=1}^m U_{\alpha,i} \hat{a}_{\alpha;h}^\dagger, \quad (2.97)$$

where  $U$  is a  $m \times m$  unitary matrix. As we discussed at the end of Section 2.2, when the photons share the same spectral-polarisation mode  $|h\rangle_{\mathcal{M}}$ , any intensity operator  $\hat{I}_i$  can be substituted by the photon number  $\hat{a}_{i;h}^\dagger \hat{a}_{i;h}$  up to a proportionality factor<sup>1</sup>, and we can write for  $i \neq j$ :

$$\langle \hat{I}_i \rangle = \frac{\mathcal{E}_h}{\mathcal{A}} \text{Tr} \left[ \rho_{\text{out}} \hat{a}_{i;h}^\dagger \hat{a}_{i;h} \right], \quad \langle \hat{I}_i \hat{I}_j \rangle = \frac{\mathcal{E}_h^2}{\mathcal{A}^2} \text{Tr} \left[ \rho_{\text{out}} \hat{a}_{i;h}^\dagger \hat{a}_{j;h}^\dagger \hat{a}_{i;h} \hat{a}_{j;h} \right], \quad (2.98)$$

where we exploited the commutation relation  $[\hat{a}_{i;h}, \hat{a}_{j;h}^\dagger] = 0$ .

We have now all the tools needed to study the quantum correlation function  $\overline{G}$ . At first, explicit expressions for  $\langle \hat{I}_i \rangle$  and  $\langle \hat{I}_i \hat{I}_j \rangle$  are evaluated and compared with their classical counterparts. As in the two-mode HOM setup [62], this analysis reveals that the classical threshold can be violated only if some of the input states are characterised by sub-Poissonian photon number statistics. Then, we find the minimum value that  $\overline{G}$  could assume in a quantum framework, under specific assumptions on the inputs. In particular, we are able to analytically solve this minimisation for a symmetric setup in which states with the same sub-Poissonian photon number statistics  $q_1(n) = \dots = q_{N_s}(n)$  are injected in every input port of the interferometer, i.e., when  $N_s = m$ .

### Sub-Poissonianity is necessary in order to violate the classical threshold

The evaluation of  $\langle \hat{I}_i \rangle$  and  $\langle \hat{I}_i \hat{I}_j \rangle$  follows the steps of other studies (see, e.g., Refs. [47, 68, 69], and it is best performed by applying the operators  $\hat{U}, \hat{U}^\dagger$  characterising the linear optical evolution to the creation and annihilation operators in Equation (2.98). This can be done by exploiting the cyclic property of the trace, and by noticing that from Equation (2.97) it follows that

$$\hat{U}^\dagger \hat{a}_{i;h} \hat{U} = \sum_{\alpha=1}^m U_{i,\alpha} \hat{a}_{\alpha;h}. \quad (2.99)$$

From now on, we drop the subscript  $h$  for the sake of simplicity. Straightforward calculations yield

$$\langle \hat{I}_i \rangle = \frac{\mathcal{E}_h}{\mathcal{A}} \sum_{\alpha=1}^{N_s} |U_{i\alpha}|^2 \langle \hat{n}_\alpha \rangle_q, \quad \langle \hat{I}_j \rangle = \frac{\mathcal{E}_h}{\mathcal{A}} \sum_{\beta=1}^{N_s} |U_{j\beta}|^2 \langle \hat{n}_\beta \rangle_q, \quad (2.100)$$

---

<sup>1</sup>For the interested reader, general expressions for the average intensities  $\langle \hat{I}_i \rangle$  and  $\langle \hat{I}_i \hat{I}_j \rangle$ , without the assumption  $|g_{(\xi_\alpha)}\rangle_{\mathcal{M}} = |h\rangle_{\mathcal{M}}$ , can be found in the supplementary material of Ref. [9].

$$\begin{aligned}
\langle \hat{I}_i \hat{I}_j \rangle = & \langle \hat{I}_i \rangle \langle \hat{I}_j \rangle + \frac{\mathcal{E}_h^2}{\mathcal{A}^2} \sum_{\alpha=1}^{N_s} |U_{i\alpha}|^2 |U_{j\alpha}|^2 \left( \langle \hat{n}_\alpha^2 \rangle_q - \langle \hat{n}_\alpha \rangle_q^2 - \langle \hat{n}_\alpha \rangle_q \right) \\
& + \frac{\mathcal{E}_h^2}{\mathcal{A}^2} \sum_{\alpha \neq \beta}^{N_s} U_{i\alpha} U_{i\beta}^* U_{j\beta} U_{j\alpha}^* \langle \hat{n}_\alpha \rangle_q \langle \hat{n}_\beta \rangle_q,
\end{aligned} \tag{2.101}$$

where  $\langle \hat{n}_\alpha^k \rangle_q = \text{Tr} [\rho (\hat{a}_\alpha^\dagger \hat{a}_\alpha)^k]$ , and the subscript  $q$  reminds us of the effective probability distribution that fully characterises the input state  $\rho$  via Equation (2.94). These expressions share many similarities with their classical counterparts, given in Equations (2.77) and (2.78). The entries of the unitary matrix  $U$  take the place of those of the evolution matrix  $T$ , and the average photon numbers  $\langle \hat{n}_\alpha \rangle_q$  play the role of the classical average intensities  $\langle |A_\alpha|^2 \rangle$ . The main difference lies in the presence of a negative term in  $\langle \hat{I}_i \hat{I}_j \rangle$ , linear in  $\hat{n}_\alpha$ . This contribution is a direct consequence of the quantum commutation relations, because it originates when evaluating the expectation value of  $\hat{a}_\alpha^\dagger \hat{a}_\alpha^\dagger \hat{a}_\alpha \hat{a}_\alpha = \hat{n}_\alpha^2 - \hat{n}_\alpha$ . Differently than in the classical case, the second term appearing in Equation (2.101) can now be negative, and this gives the possibility of obtaining a value of intensity correlation that is smaller than the classical threshold. This observation immediately shows that a *necessary* condition to violate the classical bound is the presence of sources characterised by a sub-Poissonian photon number statistics, i.e., such that:

$$\exists \alpha : \quad \langle \hat{n}_\alpha^2 \rangle_q - \langle \hat{n}_\alpha \rangle_q^2 \leq \langle \hat{n}_\alpha \rangle_q. \tag{2.102}$$

For example, this immediately excludes the possibility of violating the classical threshold by injecting in the interferometer quantum states obtained by squeezing the vacuum<sup>2</sup>, because these are super-Poissonian. This is interesting and somehow unexpected, as squeezed states are intrinsically non-classical, and are known to offer advantages in many information processing tasks. However, the nonclassical character of squeezed states lies in the fact that the variance of one of their quadratures is smaller than in the vacuum, at the expenses of increasing the variance of the other one. This feature is lost once we average over all possible phases of the field, or equivalently over all directions in phase space. The remaining state thus behaves like a noisy classical source, and cannot lead to a nonclassical signature in  $\overline{G}$ .

It is also important to stress that the presence of a sub-Poissonian source is not, by itself, a sufficient condition to observe values of intensity correlation that cannot be explained classically. This is because the remaining sources could be very noisy, with  $\langle \hat{n}_\alpha^2 \rangle_q - \langle \hat{n}_\alpha \rangle_q^2 \gg \langle \hat{n}_\alpha \rangle_q$ , and lead to an overall positive second term in Equation (2.101), once the average over all pairs  $i \neq j$  is considered. On the other hand, in certain conditions a single sub-Poissonian source could be enough to violate the classical bound on intensity correlations. For example, this is the case if coherent states with the same average photon

<sup>2</sup>We will extensively discuss squeezing in Chapter 3. The reader unfamiliar with the subject can find there a short introduction and further references.

number of the sub-Poissonian source are injected in all other ports of a  $N_d \times N_d$  Fourier transform interferometer. Indeed, this situation leads to a value of  $\overline{G}$  equal to the classical threshold minus some negative contribution due to the sub-Poissonian source. Both the observations made in this paragraph are already known to apply to the case with two sources and two detectors [62]. The similarity can be expected, because the correlators  $\langle I_i I_j \rangle$  have the same structure of the quantity  $\langle I_1 I_2 \rangle$  studied in the past.

Before moving on, we briefly discuss the case of input coherent states, for which the quantum expectation value  $\langle \hat{I}_1 \hat{I}_2 \rangle$  is expected to reduce to its classical counterpart of Equation (2.78). Formally, this is enforced by the fact that coherent states are eigenvectors of the electric field operators, as was shown in Equation (2.22). From a more practical perspective, we can calculate  $\langle \hat{n}^2 \rangle_q - \langle \hat{n} \rangle_q^2 - \langle \hat{n} \rangle_q$  for a pure coherent state  $|A\rangle$ . From its definition in Equation (2.19), the latter is characterised by a Poissonian statistics  $q(n) = e^{-|A|^2} |A|^{2n} / n!$ , and one has

$$\langle \hat{n}^2 \rangle_q - \langle \hat{n} \rangle_q^2 - \langle \hat{n} \rangle_q = |A|^4 - |A|^2. \quad (2.103)$$

The expression for  $\langle I_i I_j \rangle$  given in Equation (2.78) is thus recovered by taking into account the possibility of injecting mixtures of different coherent states.

### Minimisation of $\overline{G}$ in a symmetric quantum framework

Having discussed the possibility of violating the classical threshold on intensity correlation, it is natural to wonder by “how much” this could be done. For a fixed pair  $(N_s, N_d)$ , one should minimise  $\overline{G}$  over all possible quantum states of the form of Equation (2.94) and over all linear optical interferometers with  $m \geq N_s, N_d$  modes. One idea is to use the same strategy that was adopted in Section 2.4.1 for the classical case, by introducing  $N_d$  complex vectors  $|\tilde{\psi}_i\rangle \in \mathbb{C}^{N_s}$  with components

$$\langle \alpha | \tilde{\psi}_i \rangle_{\mathbb{C}^{N_s}} \equiv U_{i\alpha}^* \sqrt{\frac{\mathcal{E}_h \langle \hat{n}_\alpha \rangle_q}{\mathcal{A} \langle \hat{I}_i \rangle}}. \quad (2.104)$$

However, the linear term in  $\hat{n}_\alpha$  appearing in Equation (2.101) makes this minimisation much harder than its classical counterpart, because we cannot express the correlation function  $\overline{G}$  only in terms of these vectors. Despite this problem, an analytical minimum can still be found with this method under some additional constraints.

For example, here we consider a symmetric setting, in which states with the *same* sub-Poissonian photon number statistics are injected in *every port* of a linear optical interferometer. The correlation function obtained in this scenario will be labelled by  $\overline{G}^{(\text{sym})}$ . Although not general, this situation is still of interest in the study of many-particle interference effects, because its symmetry can simplify the calculation and the interpretation of the results. Indeed, particular instances of this more intuitive and balanced framework,

for example with a single photon per input mode, have been extensively studied in the literature (e.g., see Refs. [23, 25, 33]). When states with the same sub-Poissonian photon number statistics are injected in every port of a linear optical interferometer, all the information on the sources are summarised by the  $\alpha$ -independent coefficient

$$0 \leq \eta(q) = -\frac{\langle \hat{n}_\alpha^2 \rangle_q - \langle \hat{n}_\alpha \rangle_q^2}{\langle \hat{n}_\alpha \rangle_q} \leq 1, \quad (2.105)$$

whose positivity is enforced by sub-Poissonianity. Its maximum value is 1, which is reached by single-photon sources. Under these conditions, the amount of correlation among output intensities can be lower bounded by the following theorem.

**Theorem 2.2.** *Consider a symmetric quantum framework, in which  $m$  identical states with sub-Poissonian photon number statistics with parameter  $\eta(q) \in [0, 1]$  are injected in the input ports of a  $m$ -mode linear optical interferometer. The amount of correlation among the output intensities detected in the first  $N_d \leq m$  modes can be lower bounded by*

$$\overline{G}^{(sym)} \geq 1 - \frac{1 + \eta(q)}{N_d}. \quad (2.106)$$

Furthermore, this inequality can be saturated by using a linear optical interferometer characterised by a block diagonal evolution matrix  $U = U_{FT(N_d)} \oplus U_0$ , where  $U_{FT(N_d)}$  is a  $N_d$ -dimensional Fourier transform matrix and  $U_0$  is a generic unitary with dimension  $m - N_d$ .

Before going into the details of the proof, we shortly comment on the result.

- For single-photon sources in a two-mode interferometer whose outputs are both monitored, we recover the HOM result  $\overline{G}^{(sym)} = 0$ , as expected.
- In a symmetric setting we are always dealing with a number of detectors that is smaller than the number of sources. The linear optical setup that leads to the minimum value of  $\overline{G}^{(sym)}$  is the same optical interferometer that appeared in the classical case for  $N_s \geq N_d$ , represented in Figure 2.5. In particular,  $N_d$  sources are connected to the detectors via a Fourier transform interferometer, while the light originating from the remaining  $N_s - N_d$  sources is ignored. A Fourier transform interferometer can be built in waveguides by using a number of beamsplitters that scales efficiently with the dimensionality [70]. Because of its symmetry, this setup often appears in theoretical and experimental investigations. For example, in Ref. [44] it was proposed in the context of boson sampling certification, as a tool to distinguish real bosonic interference from semiclassical imitations. Our findings show that it represents the optimal interferometric choice in order to certify that the interference pattern obtained in a quantum experiment cannot be explained by classical optics, at least under the assumption of symmetric inputs.

- Because of the sub-Poissonianity of the sources, the coefficient  $\eta(q)$  is always positive, and the minimum value achievable by  $\overline{G}^{(\text{sym})}$  is strictly smaller than the classical threshold of  $1 - 1/N_d$  found in Theorem 2.1. Single-photon sources lead to the maximum violation obtainable in a symmetric setup, but even in this case the gap with the classical threshold reduces with the number of detectors as  $1/N_d$ . However, it is possible that smaller values of intensity correlation could be obtained by removing the hypothesis on symmetric inputs. Despite the difficulty in dealing analytically with the general form of  $\overline{G}$ , a numerical study might be possible. This investigation is left open for future research.

*Proof of Theorem 2.2.* Importantly, we can use the relation in Equation (2.105) in order to write  $\overline{G}$  as a quadratic function of the average photon numbers  $\{\langle \hat{n}_\alpha \rangle_q\}_{\alpha=1}^{N_s}$ . The resulting expression can then be effectively parametrised by the vectors  $\{|\tilde{\psi}_i\rangle\}_{i=1}^{N_d}$ . The result is

$$\overline{G}^{(\text{sym})} = 1 + \frac{1}{N_d(N_d - 1)} \sum_{i \neq j}^{N_d} \left[ |\langle \tilde{\psi}_i | \tilde{\psi}_j \rangle|^2 - [1 + \eta(q)] \sum_{\alpha=1}^{N_s} |\langle \alpha | \tilde{\psi}_i \rangle|^2 |\langle \alpha | \tilde{\psi}_j \rangle|^2 \right]. \quad (2.107)$$

This expression closely resembles the classical one reported in Equation (2.82), which could be recovered by setting  $\eta(q) = 0$ . Thanks to the symmetric framework that we are considering, and the unitarity of  $U$ , the average output intensities are all given by  $\langle \hat{I}_i \rangle = \mathcal{E}_h \langle \hat{n} \rangle_q / \mathcal{A}$  and the set  $\{|\tilde{\psi}_i\rangle\}_{i=1}^{N_d}$  is composed of  $N_d \leq N_s = m$  orthonormal vectors:

$$\langle \tilde{\psi}_i | \tilde{\psi}_j \rangle = \frac{\mathcal{E}_h \langle \hat{n} \rangle_q}{\mathcal{A} \sqrt{\langle \hat{I}_i \rangle \langle \hat{I}_j \rangle}} \sum_{\alpha=1}^{N_s} U_{i\alpha} U_{j\alpha}^* = \delta_{ij}. \quad (2.108)$$

We stress that these simplifications are possible only because there is one source per input mode, otherwise the sum in  $\alpha$  would not run over all the columns of the unitary matrix  $U$ . The orthogonality relation that we just proved allows us to take the factor  $[1 + \eta(q)]$  in Equation (2.107) outside the square brackets, and we recover the same term that was appearing in the classical case. Therefore, we can lower bound  $\overline{G}^{(\text{sym})}$  through the minimum value of classical intensity correlations:

$$\overline{G}^{(\text{sym})} \geq 1 + [1 + \eta(q)] \left( \min_{\{|\psi_i\rangle\}_{i=1}^{N_d}} \overline{G}^{(\text{cl})} - 1 \right). \quad (2.109)$$

Furthermore, the above inequality can be saturated by choosing each  $|\tilde{\psi}_i\rangle$  as in the optimal classical case of Equation (2.92). In a quantum symmetric setup this could be obtained by letting  $N_d$  modes evolve through a  $N_d$ -dimensional Fourier transform interferometer with monitored outputs, while the light emitted by the remaining  $m - N_d$  sources independently evolves towards the unmonitored outputs. As the number of sources equals the number of interferometric modes,  $N_s$  is necessarily larger than the number of detectors.



Hence, according to Theorem 2.1, the classical minimum is  $1 - 1/N_d$ , and the smallest value of intensity correlation that could be possibly obtained in a quantum setting under symmetric conditions is:

$$\min \overline{G}^{(\text{sym})} = 1 - \frac{1 + \eta(q)}{N_d}. \quad (2.110)$$

□

### 2.4.3 Two practical applications

So far, we have been mostly concerned with the problem of nonclassicality certification, and we showed that under certain conditions the result of a multimode linear optical experiment cannot be explained by the classical theory of light. In this section, instead, we discuss two more practical applications of the formalism developed so far. Given full control on the sources, we show that a measurement of  $\overline{G}$  can yield information on the interferometer characterising the evolution, and vice-versa.

In the first case, we develop a criterion that can certify the impossibility of dividing an interferometer into independent sub-blocks. In other words, we can guarantee that the unitary matrix  $U$ , characterising the evolution of the creation operators, cannot be written as direct sum  $U_1 \oplus U_2$  for some unitary matrices  $U_1, U_2$ . In the second case, we address the issue of quantify the average distinguishability between the spectral-polarisation modes characterising the fields emitted by the sources.

#### Interferometer divisibility

Consider an  $m \times m$  “divisible” interferometer, i.e., composed by two independent smaller interferometers with respectively  $m_1$  and  $m_2 = m - m_1$  modes, as in Figure 2.7a. Its evolution matrix  $U$  can thus be divided into two blocks, as  $U_1 \oplus U_2$ . We now show that, in a symmetric framework, the amount of intensity correlation detected in output of a divisible interferometer has to be larger than a certain threshold  $\min \overline{G}^{(\text{sym,div})}$ , and cannot reach the minimum value of Equation (2.110). This gives us a way to certify the impossibility of dividing an interferometer into independent sub-blocks, whenever the measured value of  $\overline{G}^{(\text{sym})}$  is smaller than  $\min \overline{G}^{(\text{sym,div})}$ .

In a symmetric setup  $m$  identical input states of light are used, with photon-number statistics described by the coefficient  $\eta \geq 0$  defined as in Equation (2.105). When every output port is monitored by a detector, the correlation function  $\overline{G}^{(\text{sym,div})}$  for a divisible interferometer can be split into three parts. Two of these involve the correlations between the intensities found in output of the same smaller interferometer, characterised by  $U_1$  or  $U_2$ , while the last term contains all cross-correlations. This observation, together with the fact that  $\langle \hat{I}_i \hat{I}_j \rangle = \langle \hat{I}_i \rangle \langle \hat{I}_j \rangle$  whenever  $i$  and  $j$  label the output ports of two independent



interferometers, yields:

$$\overline{G}^{(\text{sym}, \text{div})} = \frac{1}{\binom{m}{2}} \left[ \binom{m_1}{2} \overline{G}_1^{(\text{sym})} + \binom{m_2}{2} \overline{G}_2^{(\text{sym})} + m_1 m_2 \right], \quad (2.111)$$

where  $\overline{G}_1^{(\text{sym})}$  and  $\overline{G}_2^{(\text{sym})}$  are the correlation functions that would be obtained by considering the two smaller interferometers separately [see the definition of  $\overline{G}$  in Equation (2.70)]. By applying the result of Theorem 2.2 to  $\overline{G}_1^{(\text{sym})}$  and  $\overline{G}_2^{(\text{sym})}$ , we obtain the bound:

$$\overline{G}^{(\text{sym}, \text{div})} \geq 1 - [1 + \eta] \frac{m - 2}{m(m - 1)}, \quad (2.112)$$

which is saturated if  $U_k = U_{\text{FT}(m_k)}$ , for  $k = 1, 2$ . Note that this value does not depend on the dimensions of the two sub-blocks, but only on the total number of modes  $m$  of the interferometer.

For any  $\eta \geq 0$  and number of modes  $m \geq 2$ , the right-hand side of Equation (2.112) is larger than the global minimum value of  $\overline{G}^{(\text{sym})}$  that could be obtained by a fully connected interferometer, which according to Theorem 2.2 is  $1 - [1 + \eta]/m$ . Therefore, a violation of the bound in Equation (2.112) can certify the impossibility of dividing the interferometer into two independent blocks. The range of correlation values that, if measured, can trigger this certification has the following dependence upon  $m$  and  $\eta$ :

$$\Delta \overline{G}^{(\text{sym})} \equiv \min \overline{G}^{(\text{sym}, \text{div})} - \min \overline{G}^{(\text{sym})} = \frac{1 + \eta}{m(m - 1)}, \quad (2.113)$$

which is plotted in Figure 2.7b. Unfortunately, even in the optimal case of single-photon sources, this gap closes very quickly with the size of the interferometer, as it can be seen from the figure.

### Measure of average distinguishability

One of the most typical applications of the original HOM effect consists in using the dip in the probability of detecting coincident events, shown in Figure 2.2, in order to deduce the distinguishability of the photons that are injected into the beamsplitter. This is quantified by the squared overlap between the vectors describing the spectral-polarisation modes of the two photons, i.e., by  $|\langle g|f \rangle_{\mathcal{M}}|^2$  in Equation (2.44). In the same spirit, we now show that if the  $\alpha$ th source emits pulses of light with spectral-polarisation modes  $\{|g_\alpha\rangle_{\mathcal{M}}\}_\alpha$  for  $\alpha = 1, \dots, N_s$ , it is possible to recover information about the average pairwise distinguishability of the sources through a measurement of  $\overline{G}$ , as long as we have control over the interferometric setup and the photon number statistics of the sources.

It should be stressed that in our formalism, differently than in the discussion that led to Equation (2.44), we are dealing with intensity measurements rather than with photon-number measurements. The two pictures are the same, up to a proportionality factor,

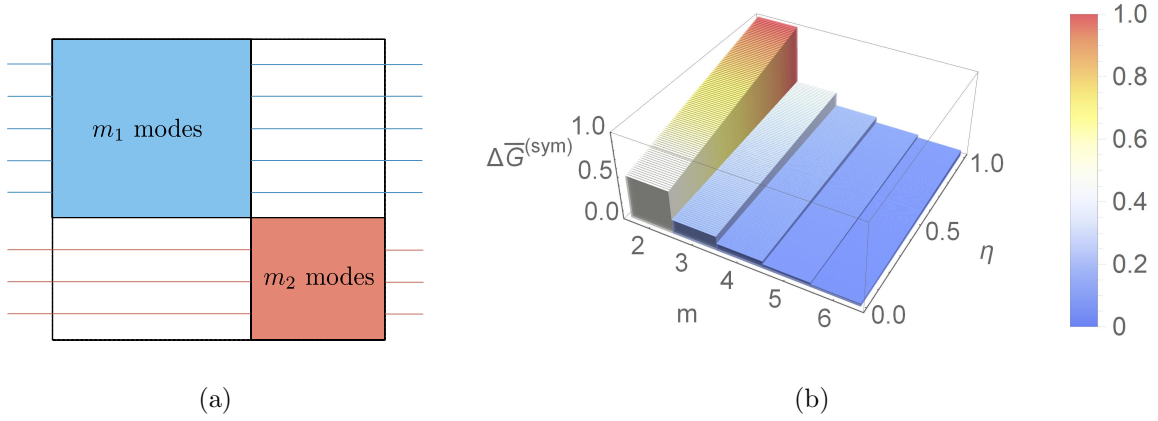


Figure 2.7: Certification of the impossibility of dividing an interferometer into smaller independent blocks. Panel (a) schematically shows the structure of a divisible interferometer. In panel (b) we plot the gap in intensity correlations between the minimum values that could be possibly achieved with a divisible or generic interferometer with  $m$  modes, when  $m$  indistinguishable single-photon sources are used.

as long as all the photons have the same spectral-polarisation mode, but are slightly different in general. The reason can be simply understood by looking at the definition of the operator  $\hat{I}$  in Equation (2.30), and by remembering that the operator  $\hat{a}_{\omega,\lambda}$  brings a factor  $g(\omega, \lambda)$  via Equation (2.21) whenever it is applied on a quantum state with mode  $|g\rangle_{\mathcal{M}}$ . If the quantum states on which we are evaluating the expectation value of intensity operators are composed by photons with different spectral-polarisation modes, say  $|g\rangle_{\mathcal{M}}$  and  $|f\rangle_{\mathcal{M}}$ , the following complex overlaps appear in the expression of  $\bar{G}$ :

$$\sum_{\lambda=1,2} \int_0^{+\infty} d\omega \hbar\omega g(\omega, \lambda)^* f(\omega, \lambda), \quad (2.114)$$

in addition to the energies  $\mathcal{E}_g$  and  $\mathcal{E}_f$  defined as in Equation (2.27). Therefore, the only distinguishability statements that we can make by measuring intensities concern overlaps of this form, rather than of the form  $\langle g|f\rangle_{\mathcal{M}}$ . An argument similar to the one that follows appeared in Ref. [69], where the authors studied the possibility of using correlators of the form  $\langle \hat{n}_i \hat{n}_j \rangle - \langle \hat{n}_i \rangle \langle \hat{n}_j \rangle$  to obtain information on the average distinguishability of the sources.

Consider an  $m$ -mode Fourier transform interferometer with one detector per output port, and input quantum states of light with the same sub-Poissonian photon-number statistics characterised by the coefficient  $\eta \geq 0$  [see Equation (2.105)], but with source-dependent spectral-polarisation modes. For simplicity, we further assume that each  $|g_\alpha\rangle_{\mathcal{M}}$  carries the same amount of energy, that is  $\mathcal{E}_{g_\alpha} = \mathcal{E}$ . For example, starting from a given mode  $|g\rangle_{\mathcal{M}}$ , one could obtain different modes with the same energy by rotating the initial polarisation vector or by delaying in time the light beam, so that  $g(\omega, \lambda)$  would go to

$e^{-i\omega t}g(\omega, \lambda)$ . With this setting, we now describe a procedure to measure the average pairwise overlap

$$\frac{1}{\binom{m}{2}} \sum_{\alpha \neq \beta}^m |\langle \tilde{g}_\alpha | \tilde{g}_\beta \rangle_{\mathcal{M}}|^2, \quad (2.115)$$

where we defined for each  $\alpha = 1, \dots, m$ :

$$\langle \omega, \lambda | \tilde{g}_\alpha \rangle_{\mathcal{M}} \equiv \sqrt{\frac{\hbar\omega}{\mathcal{E}}} \langle \omega, \lambda | g_\alpha \rangle_{\mathcal{M}}. \quad (2.116)$$

Indeed, the expectation values of output intensity operators can be written as

$$\langle \hat{I}_i \rangle = \frac{\mathcal{E}}{\mathcal{A}} \sum_{\alpha=1}^m |U_{i\alpha}|^2 \langle \hat{n}_\alpha \rangle, \quad (2.117)$$

and

$$\begin{aligned} \langle \hat{I}_i \hat{I}_j \rangle &= \langle \hat{I}_i \rangle \langle \hat{I}_j \rangle - \frac{\mathcal{E}^2}{\mathcal{A}^2} \eta \sum_{\alpha=1}^m |U_{i\alpha}|^2 |U_{j\alpha}|^2 \langle \hat{n}_\alpha \rangle^2 \\ &+ \frac{\mathcal{E}^2}{\mathcal{A}^2} \sum_{\alpha \neq \beta}^m |\langle \tilde{g}_\alpha | \tilde{g}_\beta \rangle_{\mathcal{M}}|^2 U_{i\alpha} U_{i\beta}^* U_{j\beta} U_{j\alpha}^* \langle \hat{n}_\alpha \rangle \langle \hat{n}_\beta \rangle, \end{aligned} \quad (2.118)$$

where by hypothesis  $\langle \hat{n}_\alpha \rangle = \langle \hat{n} \rangle \forall \alpha$ , because the sources have the same photon-number statistics. Note that the squared overlap  $|\langle \tilde{g}_\alpha | \tilde{g}_\beta \rangle_{\mathcal{M}}|^2 \leq 1$  reduces the amount of interference between sources  $\alpha$  and  $\beta$ . By taking into account the form of the Fourier transform interferometer of Equation (2.71), the correlation function  $\overline{G}^{(\text{sym})}$  can be written as:

$$\overline{G}^{(\text{sym})} = 1 - \frac{1}{m} \left[ \eta + \frac{1}{\binom{m}{2}} \sum_{\alpha \neq \beta}^m |\langle \tilde{g}_\alpha | \tilde{g}_\beta \rangle_{\mathcal{M}}|^2 \right], \quad (2.119)$$

from which one can deduce the average pairwise distinguishability of the sources as

$$\frac{1}{\binom{m}{2}} \sum_{\alpha \neq \beta}^m |\langle \tilde{g}_\alpha | \tilde{g}_\beta \rangle_{\mathcal{M}}|^2 = m(1 - \overline{G}^{(\text{sym})}) - \eta. \quad (2.120)$$

*Remark:* Note that if Equation (2.120) takes the value 1, then the vectors  $|\tilde{g}_\alpha\rangle_{\mathcal{M}}$  are all the same. As a consequence of Equation (2.116), also the original vectors  $|g_\alpha\rangle_{\mathcal{M}}$  have to be the same, and the photons are perfectly indistinguishable. However, it is important to keep in mind that the pairwise distinguishabilities of three or more photons are in general not enough to fully characterise their evolution, because the phases of terms like  $\langle g|f\rangle \langle f|h\rangle \langle h|g\rangle$  are also relevant. For further details about this issue we refer the reader to Ref. [71].

### 2.4.4 Weakening the assumptions on the classical framework

Most of this chapter has been developed under the assumption that the pulses of light emitted by the sources can maximally interfere at the detectors' locations. In the classical framework, this was enforced by the third condition mentioned in Section 2.3.2. Because of the importance of limiting as little as possible the studied classical picture, in this section we relax the hypothesis on the spectral-polarisation modes emitted by the sources, and on the optical paths travelled by the light beams to reach the detectors. We will show that, even in this more general setting, the classical values of intensity correlation cannot drop below the threshold found in Theorem 2.1. The reader not interested in details can safely jump to Section 2.5 for a discussion on the results presented so far.

#### A convenient formalism

The generic expression for the electric field at the detectors' positions, for a given realisation  $\xi$  of the sources, has been given in Equation (2.54). As it is rather long, we are going to break it down into smaller terms, composed by vectors and operators defined on the Hilbert space  $\mathcal{M} = L^2 \otimes \mathbb{C}^2$  describing the frequency-polarisation degrees of freedom of the light. We start by defining a vector  $|\mathcal{E}_g\rangle_{\mathcal{M}}$ , with components

$$\langle\omega, \lambda|\mathcal{E}_g\rangle_{\mathcal{M}} \equiv \sqrt{\hbar\omega} g(\omega, \lambda). \quad (2.121)$$

It resembles the vector  $|\tilde{g}\rangle_{\mathcal{M}}$  defined in the previous paragraph, when we discussed the possibility of deducing the average distinguishability of the sources from a measurement of intensity correlation. However, we point out that  $|\mathcal{E}_g\rangle_{\mathcal{M}}$  does not have unit norm, because  $\langle\mathcal{E}_g|\mathcal{E}_g\rangle_{\mathcal{M}} = \mathcal{E}_g$  represents the energy associated with the spectral-polarisation mode  $|g\rangle_{\mathcal{M}}$ . If the distances between different sources and detectors are not the same, each spectral component of the light beams will acquire a path-dependent phase. Therefore, it will be convenient to introduce a phase shift operator diagonal in the frequency-polarisation basis:

$$\mathcal{P}(\Delta) \equiv \sum_{\lambda=1,2} \int_0^{+\infty} d\omega |\omega, \lambda\rangle_{\mathcal{M}} \langle\omega, \lambda| e^{i\Delta\omega}. \quad (2.122)$$

Moreover, we introduce another complete basis in the space of squared integrable functions, with elements  $|t\rangle_{L^2}$  such that  $\langle t|\omega\rangle_{L^2} \equiv e^{-i\omega t}/\sqrt{2\pi}$ . With all these definitions, we can rewrite the complex electric field at the  $i$ th detector as follows:

$$\mathbf{E}_i^{(\xi)}(\bar{z}_i, t) = i \left( \frac{1}{2\epsilon_0 c \mathcal{A}} \right)^{\frac{1}{2}} \sum_{\alpha=1}^{N_s} T_{i\alpha} A_{\alpha}(\xi_{\alpha}) \int_{L^2} \langle t|\mathcal{P}(\Delta_{i\alpha})|\mathcal{E}_{g(\xi_{\alpha})}\rangle_{\mathcal{M}}, \quad (2.123)$$

where  $\Delta_{i\alpha} \equiv (\bar{z}_i - z_{\alpha})/c$ , and  $\xi$  contains all the information about the particular realisation of the emitted pulses of light, up to the random phase of  $A_{\alpha}(\xi_{\alpha})$ . The integrated intensity

reaching the  $i$ th detector can then be obtained from the definitions in Equations (2.25) and (2.26), and reads

$$I_i^{(\xi)} = \frac{1}{\mathcal{A}} \sum_{\alpha, \beta=1}^{N_s} T_{i\alpha} T_{i\beta}^* A_\alpha(\xi_\alpha) A_\beta^*(\xi_\beta) \langle \mathcal{E}_{g(\xi_\beta)} | \mathcal{P}(\Delta_{i\alpha} - \Delta_{i\beta}) | \mathcal{E}_{g(\xi_\alpha)} \rangle_{\mathcal{M}}. \quad (2.124)$$

Having obtained this compact expression for the measured intensity, we now have to average  $I_i^{(\xi)}$  and  $I_i^{(\xi)} I_j^{(\xi)}$  over all possible realisations  $\xi$ . This task is simplified by the randomness of the phases characterising the complex amplitudes  $\{A_\alpha(\xi_\alpha)\}_\alpha$ . For example, when we compute  $\langle I_i \rangle$ , only the terms with  $\alpha = \beta$  survive. This yields

$$\langle I_i \rangle = \sum_{\alpha=1}^{N_s} |T_{i\alpha}|^2 \langle I_\alpha \rangle, \quad (2.125)$$

where  $\langle I_\alpha \rangle$  is the average intensity emitted by the  $\alpha$ th source, i.e.

$$\langle I_\alpha \rangle = \sum_{\xi_\alpha} p_\alpha(\xi_\alpha) I_\alpha^{(\xi_\alpha)} \equiv \sum_{\xi_\alpha} p_\alpha(\xi_\alpha) \frac{1}{\mathcal{A}} |A_\alpha(\xi_\alpha)|^2 \mathcal{E}_{g(\xi_\alpha)}. \quad (2.126)$$

In a similar manner, when we evaluate the average of  $I_i^{(\xi)} I_j^{(\xi)}$  we are left with three terms. Say that  $I_i^{(\xi)}$  is expanded as in Equation (2.124), and that  $I_j^{(\xi)}$  has an analogous expression as sum of terms depending on  $A_\gamma(\xi_\gamma) A_\delta^*(\xi_\delta)$ . The first two contributions in  $I_i^{(\xi)} I_j^{(\xi)}$  are obtained respectively when: (i)  $\alpha \neq \gamma$ ,  $\alpha = \beta$ , and  $\gamma = \delta$ , (ii)  $\alpha = \beta = \gamma = \delta$ . Their sum can be written as

$$\langle I_i \rangle \langle I_j \rangle + \sum_{\alpha=1}^{N_s} |T_{i\alpha}|^2 |T_{j\alpha}|^2 [\langle I_\alpha^2 \rangle - \langle I_\alpha \rangle^2], \quad (2.127)$$

and generalises what has been previously obtained in the first line of Equation (2.78). The variance of  $I_\alpha^{(\xi_\alpha)}$  that appears between squared brackets is always positive, and can become zero without altering the remainder of the considered expression simply by setting the variance of the input intensities equal to zero. We can thus ignore its contribution because we are looking for the minimum amount of correlation between output intensities by optimizing over all possible sources. With the notation previously adopted, the last term appearing in  $I_i^{(\xi)} I_j^{(\xi)}$  is obtained when  $\alpha \neq \beta$ ,  $\alpha = \delta$ ,  $\beta = \gamma$ , and it is due to the interference of light beams coming from the same source but reaching different detectors. After some straightforward manipulations, by ignoring the second term in Equation (2.127) we are left with

$$\langle I_i I_j \rangle \geq \langle I_i \rangle \langle I_j \rangle + \sum_{\alpha \neq \beta}^{N_s} T_{i\alpha} T_{i\beta}^* T_{j\beta} T_{j\alpha}^* \text{Tr}_{\mathcal{M}} [\mathcal{E}_\alpha \mathcal{P}(\Delta_{j\beta} - \Delta_{j\alpha}) \mathcal{E}_\beta \mathcal{P}(\Delta_{i\alpha} - \Delta_{i\beta})] \langle I_\alpha \rangle \langle I_\beta \rangle, \quad (2.128)$$

where  $\mathcal{E}_\alpha$  is a positive semidefinite operator with  $\text{Tr}_{\mathcal{M}}[\mathcal{E}_\alpha] = 1$ :

$$\mathcal{E}_\alpha \equiv \frac{1}{\mathcal{A}\langle I_\alpha \rangle} \sum_{\xi_\alpha} p_\alpha(\xi_\alpha) |A_\alpha(\xi_\alpha)|^2 |\mathcal{E}_{g(\xi_\alpha)}\rangle_{\mathcal{M}} \langle \mathcal{E}_{g(\xi_\alpha)}|, \quad (2.129)$$

which characterises the average pulse of light emitted by the  $\alpha$ th source.

The trace in Equation (2.128) reduces the interference ability of the pulses of light coming from sources  $\alpha$  and  $\beta$ , once they reach detectors  $i$  and  $j$ . The different optical paths travelled by the light are relevant because wave-packets identical at the sources might become shifted once they reach the detectors. This effect is taken into account by the operators  $\mathcal{P}$ , and vanishes when  $\Delta_{i\alpha} \equiv \Delta$ , because  $\mathcal{P}(0)$  is the identical operator of the Hilbert space  $\mathcal{M}$ . In this simpler case, sources  $\alpha$  and  $\beta$  can maximally interfere when  $\text{Tr}_{\mathcal{M}}[\mathcal{E}_\alpha \mathcal{E}_\beta] = 1$ , i.e., if the vectors  $|\mathcal{E}_{g(\xi_\alpha)}\rangle_{\mathcal{M}}$ , and thus the modes  $|g(\xi_\alpha)\rangle_{\mathcal{M}}$ , are the same for all sources and realisations. This formally justifies the intuition that condition 3 of Section 2.3.2 maximises the interference among the sources.

### The lower bound of Theorem 2.1 still holds

The right hand side of Equation (2.128) is linear in each of the operators  $\mathcal{E}_\alpha$ , and so is  $\overline{G}^{(\text{cl})}$ . Therefore, its minimum value has to be reached when  $\mathcal{E}_\alpha$  is a projector  $|\mathcal{E}_{g_\alpha}\rangle_{\mathcal{M}} \langle \mathcal{E}_{g_\alpha}|$  onto a pure state. This condition corresponds to a situation in which each source emits pulses of light always characterised by the same spectral-polarisation mode, even though possibly different from the modes emitted by the other sources. As we have already done in previous sections, it is convenient to introduce auxiliary vectors  $|\psi_i\rangle \in \mathbb{C}^{N_s}$ . The generalisation of Equation (2.79) to the current framework is:  $\langle \alpha | \psi_i \rangle \equiv T_{i\alpha}^* \sqrt{\langle I_\alpha \rangle / \langle I_i \rangle}$ . The whole correlation function can thus be lower bounded as

$$\overline{G}^{(\text{cl})} \geq 1 + \min \frac{1}{N_d(N_d - 1)} \sum_{i \neq j} \sum_{\alpha \neq \beta}^{N_d} \text{Tr}_{\mathcal{M} \otimes \mathbb{C}^{N_s}} \left[ |\mathcal{E}_{g_\alpha}, \alpha\rangle \langle \mathcal{E}_{g_\alpha}, \alpha| \tilde{\mathcal{P}}^{(j)} |\mathcal{E}_{g_\beta}, \beta\rangle \langle \mathcal{E}_{g_\beta}, \beta| \tilde{\mathcal{P}}^{(i)} \right], \quad (2.130)$$

where  $|\mathcal{E}_{g_\alpha}, \alpha\rangle = |\mathcal{E}_{g_\alpha}\rangle_{\mathcal{M}} \otimes |\alpha\rangle_{\mathbb{C}^{N_s}}$ , and  $\tilde{\mathcal{P}}^{(j)}$  (similarly for  $i$ ) is an operator acting in  $M \otimes \mathbb{C}^{N_s}$  defined by adsorbing the phases  $\{e^{-i\omega\Delta_{j\alpha}}\}_{\alpha=1}^{N_s}$  into  $N_d$   $\omega$ -dependent vectors  $|\Psi_j(\omega)\rangle \in \mathbb{C}^{N_s}$ :

$$\tilde{\mathcal{P}}^{(j)} \equiv \sum_{\lambda=1,2} \int_0^{+\infty} d\omega |\omega, \lambda\rangle_{\mathcal{M}} \langle \omega, \lambda| \otimes |\Psi_j(\omega)\rangle_{\mathbb{C}^{N_s}} \langle \Psi_j(\omega)|, \quad (2.131)$$

where

$$\langle \alpha | \Psi_j(\omega) \rangle_{\mathbb{C}^{N_s}} \equiv e^{-i\omega\Delta_{j\alpha}} \langle \alpha | \psi_j \rangle_{\mathbb{C}^{N_s}}. \quad (2.132)$$

The form of Equation (2.130) might not be so intuitive at first sight, but it should be relatively easy to check its validity by noticing that

$${}_{\mathbb{C}^{N_s}}\langle\alpha|\tilde{\mathcal{P}}^{(j)}|\beta\rangle_{{}_{\mathbb{C}^{N_s}}} = \mathcal{P}(\Delta_{j\beta} - \Delta_{j\alpha}) \langle\alpha|\psi_j\rangle \langle\psi_j|\beta\rangle. \quad (2.133)$$

The  $N_s$  projectors  $|\mathcal{E}_{g_\alpha}, \alpha\rangle \langle\mathcal{E}_{g_\alpha}, \alpha|$  are orthogonal and have rank 1, because  $\langle\alpha|\beta\rangle = \delta_{\alpha\beta}$ . Therefore, the trace in Equation (2.130) actually acts on a space whose dimension is still effectively  $N_s$ , spanned by the basis elements

$$|a_\alpha\rangle \equiv |\mathcal{E}_{g_\alpha}, \alpha\rangle, \quad (2.134)$$

for  $\alpha = 1, \dots, N_s$ . Hence, Equation (2.130) can be rewritten as

$$\overline{G}^{(\text{cl})} \geq 1 + \min \frac{1}{N_d(N_d - 1)} \sum_{i \neq j}^{N_d} \sum_{\alpha \neq \beta}^{N_s} \text{Tr} \left[ |a_\alpha\rangle \langle a_\alpha| \tilde{\mathcal{P}}^{(j)} |a_\beta\rangle \langle a_\beta| \tilde{\mathcal{P}}^{(i)} \right], \quad (2.135)$$

where  $\tilde{\mathcal{P}}^{(j)}$  can be interpreted as a mixed density matrix in the aforementioned subspace of  $\mathcal{M} \otimes C^N$ , because it is self-adjoint, positive semidefinite [see Equation (2.131)], and with  $\text{Tr} [\tilde{\mathcal{P}}^{(j)}] = 1$ . This last statement can be easily verified by using Equation (2.133), the fact that  $\mathcal{P}(0)$  is the identity operator on  $\mathcal{M}$ , and the normalisation  $\langle\mathcal{E}_\alpha|\mathcal{E}_\alpha\rangle_{\mathcal{M}} = 1$ :

$$\begin{aligned} \text{Tr} [\tilde{\mathcal{P}}^{(j)}] &= \sum_{\alpha=1}^{N_d} \langle a_\alpha| \tilde{\mathcal{P}}^{(j)} |a_\alpha\rangle = \sum_{\alpha=1}^{N_s} \langle\mathcal{E}_\alpha, \alpha| \tilde{\mathcal{P}}^{(j)} |\mathcal{E}_\alpha, \alpha\rangle \\ &= \langle\mathcal{E}_\alpha| \mathcal{P}(0) |\mathcal{E}_\alpha\rangle \sum_{\alpha=1}^{N_s} |\langle\alpha|\psi_j\rangle|^2 = \text{Tr}_{\mathbb{C}^{N_s}} [|\psi_j\rangle \langle\psi_j|] = 1. \end{aligned} \quad (2.136)$$

Therefore, similarly to what has been done in order to substitute the operators  $\{\mathcal{E}_\alpha\}_\alpha$  with the projectors  $\{|\mathcal{E}_\alpha\rangle \langle\mathcal{E}_\alpha|\}_\alpha$ , a linearity argument allows us to state that the minimum value of the right-hand side of Equation (2.135) has to be reached when

$$\tilde{\mathcal{P}}^{(j)} = |\psi'_j\rangle \langle\psi'_j|, \quad (2.137)$$

for all  $j = 1, \dots, N_d$ , where  $|\psi'_j\rangle$  are pure states on the space spanned by  $\{|a_\alpha\rangle\}_\alpha$ . The lower bound on  $\overline{G}^{(\text{cl})}$  obtained by enforcing this condition on Equation (2.135), has the same form of the expression minimised in Section 2.4.1, up to the substitution  $|\psi'\rangle \leftrightarrow |\psi\rangle$  and  $|a\rangle \leftrightarrow |\alpha\rangle$ , as can be easily seen by averaging Equation (2.80) over all pairs of detectors. Therefore, the result of Theorem 2.1 does not need the extra hypothesis provided by condition 3 of Section 2.3.2, but it only requires independent sources emitting light with a random phase.

## 2.5 Discussion

After having presented our results, we now comment on the extent of their applicability, and on their connection to other forms of nonclassicality. This will also be a good occasion to mention possible future directions.

### 2.5.1 Can we further weaken the classical assumptions?

We have been able to prove the classical lower bound on intensity correlations of Theorem 2.1 under weak assumptions on the structure of the sources, i.e., independence and phase randomness. We also considered a specific form of light evolution, by studying a linear optical setup that does not change the input spectral-polarisation modes of light, hypothesis that can be expected to experimentally hold up to a reasonable degree of accuracy. Nonetheless, it is interesting trying to understand to what extent we can further relax the assumptions that have been made without changing our results.

Intuitively, any change on the light modes of the fields introduced by the interferometer cannot increase the interference over the ideal conditions studied in this chapter. Therefore, we do not expect the classical threshold identified here to be affected by non-ideal linear optical settings. A particular example that is worth discussing is an evolution with frequency-dependent matrix entries  $T_{i\alpha}(\omega)$ . Their phases could be taken into account along the lines of Section 2.4.4, by suitably changing the definition of the unitary operator  $\mathcal{P}$  in a path-dependent way. On the other hand, a frequency dependent absolute value  $|T_{i\alpha}(\omega)|$  cannot be straightforwardly included in our proof, and new methods should be developed in order to formally take this possibility into account.

Another interesting line of thought deals with the possibility of removing the condition of independence of the sources. Correlated input light beams have been considered in the past for this sort of studies, but only for some specific configurations. For example, Ref. [62] deals with two-mode light beams, created through parametric down-conversion, impinging on a beamsplitter. Finding a threshold for the amount of intensity correlations obtainable with correlated classical sources, at first in a two-mode, and then in a multi-mode setup, should be the first step in this analysis. This would open up the possibility of comparing the classical result with the correlations obtainable by generic quantum sources, potentially entangled with each other<sup>3</sup>.

The idea of removing the assumption on the random phases of the fields emitted by the sources, instead, should be abandoned, at least as long as one wants to study lower bounds on correlation functions of the form  $\langle I_i I_j \rangle$ . Indeed, we can easily show that independent sources of coherent states could yield  $\langle I_1 I_2 \rangle = 0$  in output of a balanced beamsplitter, if the phases of the fields are properly controlled. Say that one of the sources always emits a

---

<sup>3</sup>We point out that, after the phase averaging procedure, even the light emitted by entangled sources can be described as a separable state. However, this might be a way to engineer non-trivial photon number probability distributions in the sources.



coherent state  $|\alpha\rangle$ , while the other one emits either  $|+\alpha\rangle$  or  $|-\alpha\rangle$  with equal probabilities. In this scenario both output intensities are nonzero on average, but in every run of the experiment one of the two vanishes. The minimum of  $\langle I_1 I_2 \rangle$  can thus reach zero even in a classical picture if the assumption on random phases is removed.

Finally, it is worth mentioning that the classical bound of Theorem 2.1 is completely unaffected by losses, because the evolution matrix  $T$  is not required to be unitary. However, this is not true in the quantum case. For example, Theorem 2.2 only holds for symmetric setups, where no vacuum state can be used in input of an interferometer characterised by the unitary matrix  $U$ . Rigorously speaking, this is possible only for a lossless setup. However, the same quantum results would be obtained in presence of balanced losses, which are independent of the path taken by the light inside the interferometer. Indeed, under this assumption the relevant  $m \times m$  sub-matrix characterising the evolution of the accessible annihilation operators would become proportional to the original  $m \times m$  unitary  $U$ . This proportionality factor would then cancel in the ratio defining the correlation function  $\overline{G}$ , so that the same result found for the evolution matrix  $U$  in Theorem 2.2 would be recovered. We point out that balanced losses can be expected to arise with good approximation if the interferometer is symmetrically built, for example by using the scheme recently proposed in Ref. [72].

### 2.5.2 Relation to other forms of nonclassicality

In quantum optics, the most common and comprehensive definition of classicality dates back to the seminal works by Glauber and Sudarshan [58, 73]. In general, any single-mode density matrix  $\rho$  can be written as

$$\rho = \int \frac{d^2\alpha}{\pi} P(\alpha) |\alpha\rangle \langle \alpha|, \quad (2.138)$$

as long as the so-called “P-function”  $P(\alpha)$  is allowed to become a generic, and non necessarily positive, distribution (see Ref. [74] for a recent discussion about P-function irregularities). According to Glauber’s and Sudarshan’s definition, a state of light is considered “classical” if it can be written as convex mixture of coherent states, i.e., if it can be described by a P-function that is a probability density in phase space. This condition can be equivalently stated [75] by saying that the expectation value  $\langle : \hat{f}^\dagger \hat{f} : \rangle$  has to be non-negative for any function  $\hat{f} \equiv f(\hat{a}, \hat{a}^\dagger)$  of the annihilation and creation operators, where  $: \hat{O} :$  is the normal ordered version of  $\hat{O}$ . The multimode generalisation of these considerations is straightforward. For what we have just said, a quantum state is nonclassical as soon as an operator  $\hat{f}$  such that  $\langle : \hat{f}^\dagger \hat{f} : \rangle < 0$  is identified. Other specific forms of nonclassicality, such as sub-Poissonian photon number statistics [76] and single-mode squeezing [75], can be recovered from this definition by considering respectively  $\hat{f}_{\text{sub}} \equiv \hat{a}^\dagger \hat{a} - \langle \hat{a}^\dagger \hat{a} \rangle$  and  $\hat{f}_{\text{sq}} \equiv \hat{x}_\phi - \langle \hat{x}_\phi \rangle$ , where  $\hat{x}_\phi = (e^{-i\phi} \hat{a} + e^{+i\phi} \hat{a}^\dagger)/\sqrt{2}$ . By

exploiting this line of thoughts, in recent years several nonclassicality criteria have been developed [77], which are written in terms of measurement outcomes obtained from on/off detectors [78–81] or POVMs with a generic number of outputs [82, 83].

The results presented in this chapter share some similarities with the framework detailed in the previous paragraph. Indeed, we considered classical sources emitting states with positive P-functions, and we obtained an inequality which, if violated, acts as a nonclassicality witness. However, there are two main differences, which prevent each approach from being a straightforward consequence of the other. First, we are taking additional hypotheses on the structure of classical states (independence and phase-randomness of the sources), so that our results only apply to this experimentally relevant setting. Second, the classical lower bound on  $\overline{G}$  cannot be easily deduced from an inequality of the form  $\langle : \hat{f}^\dagger \hat{f} : \rangle \geq 0$ . For example, let us write the correlation function  $\overline{G}$  in terms of photon number operators rather than intensities, and consider the reasonable guess  $\hat{f}_g = \sum_i^{N_d} \frac{\hat{n}_i}{\langle \hat{n}_i \rangle}$ , so that:

$$N_d(N_d - 1) \overline{G} = \sum_{i \neq j}^{N_d} \frac{\langle \hat{n}_i \hat{n}_j \rangle}{\langle \hat{n}_i \rangle \langle \hat{n}_j \rangle} = \langle : \hat{f}_g^\dagger \hat{f}_g : \rangle - \sum_{i=1}^{N_d} \frac{\langle : \hat{n}_i^2 : \rangle}{\langle \hat{n}_i \rangle^2}. \quad (2.139)$$

Even if we make use of the positivity of  $\langle : \hat{f}_g^\dagger \hat{f}_g : \rangle$ , we are still left with a sum that is difficult to upper bound. In fact, thanks to the result of Theorem 2.1, we know that the number of sources would have to appear in a non-trivial way.

Another nonclassical feature that is ubiquitous in quantum information is the presence of entanglement, i.e., the impossibility of decomposing a quantum state as convex mixture of tensor products of pure local states. We can, therefore, wonder whether the nonclassicality discussed in this chapter could be somehow connected to the entanglement paradigm. A first observation is that there cannot be any entanglement between the modes of the states injected in the interferometer, because of the random phase characterising the light emitted by each source [see Equation (2.94)]. Moreover, as any input coherent state is sent to another coherent state by a linear optical transformation, we can also conclude that there is no entanglement in the state found as output of the interferometer, when classical sources are used. This argument does not follow through when generic quantum sources are considered, and in this case the output state can be entangled. For example, from Equation (2.43) it follows that, after an HOM experiment with indistinguishable photons, the output state is proportional to  $(\hat{a}_1^{\dagger 2} - \hat{a}_2^{\dagger 2}) |0\rangle$ . This is related to the well-known fact that nonclassicality at the P-function level can be converted into mode entanglement by a linear optical evolution [84–96]. However, we are not aware of any direct connection between the value of  $\overline{G}$  and the amount of output entanglement.

Finally, we should comment on the fact that the nonclassicality detection discussed in this chapter cannot be straightforwardly connected with the problem of boson sampling.

On the one hand, suppose that a certain experiment yields a value of  $\overline{G}$  within the classical region. Although there is an optical setup that can reproduce the same value of correlations by using only classical states of light, it is not possible in general to recover the probability of each detection event, or to sample from this distribution. On the other hand, here we are dealing only with the issue of comparing the predictions of quantum mechanics with a precise classical framework: classical electrodynamics. The research on boson sampling, instead, aims at proving a gap between the computational power allowed by quantum mechanics, and *any* algorithm that could possibly run on a classical computer. Although any evolution of classical light through linear optical setups can be efficiently simulated by a classical computer, the opposite seems unlikely.

## 2.6 Conclusions

In this chapter, we addressed the issue of detecting and certifying a non-classical behaviour in a typical interferometric setup, where the light emitted by independent sources evolves through a linear optical interferometer, before being measured by intensity-sensitive detectors. We introduced and studied a normalised quantifier of correlations among pairs of output intensities, and found a tight lower bound on the values that it can assume when the light emitted by the sources can be fully described by classical electric fields. The presence of sub-Poissonian quantum sources has been identified as a necessary condition to observe a violation of the classical threshold, and the maximal extent of this violation has been characterised under the assumption of symmetric input conditions. Moreover, if the interferometric setup or the input sources are suitably chosen, we have shown that a measurement of  $\overline{G}$  can yield information respectively on the average distinguishability of the light beams or on the structure of the interferometer.

This study confirms the importance and usefulness of low-order correlation functions in the study of many-particle interference effects. However, it goes without saying that the full set of probabilities for output events will typically contain much more information than what can be extracted from any combination of low-order correlators. Therefore, we feel there is still room for improvement, for example by attempting to study higher order correlation functions. By comparing quantum predictions with classical electromagnetic theory, our results give a new perspective on the issue of multi-particle interference, that has recently received a lot of attention. We hope that this study can inspire further research on the subject, and that our approach might find further applications in the future. On a more practical note, the possibility of obtaining information on sources or interferometers by measuring intensity correlations can be of interest for experimental groups, as possible tools for characterising their setups.



# Chapter 3

## Versatile probes for squeezing estimation

### 3.1 Introduction

In this chapter we focus on a very specific form of nonclassicality that can arise in quantum optics, known as “squeezing”. Due to the uncertainty principle, whenever two incompatible observables are being measured on a quantum system, the variances of the results cannot be made arbitrarily small at the same time. Quantum states of light make no exception to this rule, when two incompatible observables characterising the oscillations of the light field, known as quadratures, are measured. Perhaps unexpectedly, the vacuum is one of those states with minimal uncertainty, and one could be led to believe that no other state could yield smaller variances. However, this is not completely true, because the variance affecting the measurement of a given quadrature could be “squeezed” below the vacuum level, as long as the opposite process affects the other one. This simple, yet powerful, idea has been extensively studied and exploited for various tasks. Not only it can lead to enhanced estimations of optical phases [97], but it is also commonly used as a tool in order to approximate single-photon states [98], or as a resource in common information processing protocols [99]. Moreover, the same mathematical formalism is successfully employed in condensed matter physics, where it goes under the name of Bogoliubov transformations, or in relativistic frameworks, as it happens for the Unruh effect [100–102].

Part of the success of squeezed light is due to the fact that it is not difficult to generate experimentally, typically by exploiting non-linear properties of some medium. We will not further discuss practical implementations, but the reader interested in non-linear optics is referred to some of the many books on the subject. Instead, here we take an abstract approach, and model a single-mode squeezing device as a box that yields in output the squeezed version of any state that it receives in input. This evolution is characterised by two parameters: the squeezing “direction”, which loosely speaking defines which are the

quadratures most affected by the process, and the squeezing “strength”, which quantifies by how much the variances of these quadratures are respectively reduced or amplified. In this chapter, we are supposedly given one of these boxes, with the goal of extracting information about its squeezing strength without prior knowledge on its direction. In particular, we focus on the problem of determining the best probing system, by assuming that any measurement allowed by quantum mechanical laws could be performed on the state that is retrieved from the box. In doing this, we will restrict our analysis to the set of Gaussian states, because of their simplicity and experimental relevance (more on this on the following).

In the past, the problem of squeezing estimation has been investigated either by looking at its effect on the Hilbert space of the radiation field [103,104], or by using Gaussian [105–107] or non-Gaussian probes obtained via Kerr interactions [108], but never by considering limited knowledge on its direction of application. As our discussion will be naturally framed from a metrological perspective, in the remainder of this introduction we will provide further details on quantum metrology. Particular emphasis will be given to its “black-box” version, in which some of the details of the probe evolution are unknown. Almost free from technicalities, this overview aims at providing a summary of some of the recent developments in this research field, as well as the main ideas that one should keep in mind when reading this chapter.

The following sections, then, are organised as follows. In Section 3.2 we review the formalism of estimation theory and of Gaussian states in continuous variable systems. After that, in Section 3.3 we introduce the “average quantum Fisher information” (AvQFI), an original quantifier for probe versatility that will be used throughout the chapter. Our results are presented in Section 3.4, where we apply all the tools previously introduced to the problem of squeezing estimation. First, under the assumption of noiseless conditions, we analytically evaluate the AvQFI for single-mode Gaussian probes, and compare their performances with a paradigmatic class of two-mode correlated probes. Then, we take photon losses into account and we numerically repeat our analysis, showing in this way that the advantage of using correlations becomes more important in realistic conditions. Our conclusions and future perspectives conclude the chapter in Section 3.5.

### 3.1.1 From black-box metrology to squeezing estimation

Within the branches of quantum information theory, quantum metrology is one of the most developed fields, which has been investigated by several authors during the last decade [19,20,109–113]. Its main objective is the characterisation of the ultimate precision allowed by quantum mechanics for the problem of parameter estimation. Because the need for precise measurements and observations is ubiquitous in science, it is possible to understand why the possibility of improving the estimation precision over the best classical strategy is very appealing. Currently, quantum metrology found application in

many physical situations, in which accurate information need to be extracted from a certain system or process [114–118]. The most common metrological approach can be split into three steps: (i) preparation of the probe, (ii) interaction between probe and system of interest, and (iii) readout measurement on the evolved state of the probe. By comparing the output state with the probe, our knowledge about the evolution process can be improved. The quality of the overall estimation protocol depends on how well these three steps work together. For example, no information could possibly be extracted if the probing system cannot be significantly altered during the evolution, or if the readout measurement is not sensitive to the changes. As we anticipated, in this chapter we are going to focus only on the issue of probe selection, by optimising over all measurements allowed by quantum mechanics.

In most scenarios, the dynamics affecting the evolution of the probe is completely known in advance, up to the value of a certain parameter that needs to be estimated. For example, this could be the case if we want to measure the phase shift acquired by a spin system interacting with an external field of known direction but unknown intensity. Indeed, in this case the Hamiltonian characterising the unitary evolution is known up to the desired multiplicative factor. For this situation, it has been shown that the optimal performance can be achieved by selecting a probe which is maximally coherent in the eigenbasis of the Hamiltonian [19]. However, there are many situations where the ability of matching the probing system to the details of the actual evolution is limited by a larger degree of ignorance. For example, this might be due to a noisy environment that can randomly affect the probe evolution, thus greatly reducing the estimation precision. From this perspective, it is interesting to investigate the possibility of engineering a versatile probing system, which can perform reasonably well over a broad range of encoding operations, and to characterise the resources responsible for this behaviour.

A common way of approaching this problem is to consider a so called “black-box” metrology setting [119], which can be formulated on an abstract level as follows. At first, the set of possible probe-system interactions is fixed and communicated to us. Then, we are asked to prepare many copies of a probing system of our choice, which will be used for *all* the following experiments. Only at this point, a particular realisation of the interaction is chosen and communicated to us, so that we can exploit this information in designing the optimal readout measurement. Finally, the estimation experiment is performed, by letting the probes originally prepared sequentially interact with the system of interest, and by measuring their final state according to the optimal strategy.

For unitary encodings in finite-dimensional systems, this framework has been recently studied in the case of evolution Hamiltonians characterised by a set of fixed and known eigenvalues, but completely unknown eigenvectors [119–122]. The authors were looking for a particular “kind” of versatility, by maximising the *worst-case* estimation precision over the whole set possible encodings, i.e., of isospectral Hamiltonians. According to this figure

of merit, the performance of each probe was linked to the presence of nonclassical “discord-like” correlations [123, 124] between the probe itself and an ancillary system, which was kept for reference and measured together with the probe. Therefore, although maximally entangled states guarantee the largest minimal precision, interestingly non zero worst-case performances can be obtained also with separable input states [125–128]. The same approach has been used also in estimation problems involving continuous variable systems, by considering additional realistic constraints on the probe and its evolution. These include finite energy requirements, as well as the restriction to a Gaussian framework (more on this in the following). Once again, quantum discord has been identified as the main resource for these discrimination tasks [129–131].

Another approach to versatility in a “black-box” metrology setting consists in looking for the probe that guarantees the best *average* performance, rather than minimal, and has been pioneered by the authors of Ref. [132]. In this case, in addition to nonclassical correlations, also local purity plays an important role. This second approach to versatility is much more appealing than the first one, because it is more closely related to a physically relevant situation in which the probe-system interaction is free to fluctuate from one probe to the next, within the set of the allowed possibilities [133]. In this case one would really have no choice other than looking for versatile probes, whereas in a proper black-box metrology setting it would have been a priori possible to change the probing states after the specific interaction becomes known.

In this chapter we introduce a figure of merit for probe versatility directly connected with the average estimation precision obtainable in a metrological setting with limited initial knowledge. This can be either a proper black-box framework, as the one previously described, or a scenario where the interaction is free to fluctuate in time, but the information on each specific realisation becomes available to the experimenter only at the measurement stage. Then, we apply this formalism to the particular case of squeezing estimation without prior knowledge on its direction. Notice that, similarly to what has been previously discussed, all possible interaction Hamiltonians considered in this problem have the same spectrum, as they can be obtained from a single “seed” Hamiltonian (i.e., squeezing along a fixed direction) by applying phase rotation operators. For this reason, our work can be considered an extension of Ref. [132] to a continuous variables setting.

## 3.2 Preliminary notions

The discussion of this chapter strongly relies on concepts of estimation theory, as the notion of quantum Fisher information (QFI) [134], and on the symplectic formalism commonly used to represent Gaussian states and operations of continuous variable systems [99, 135, 136]. In this section we shortly review these topics, and provide a list of



references that could be used to gather further information. We introduce quantum estimation theory and the Gaussian framework respectively in Sections 3.2.1 and 3.2.2, then in Section 3.2.3 we join the two by providing explicit expressions for the QFI of one- and two-mode Gaussian states. In Section 3.2.4, instead, we discuss the method that will be used in the following to uniformly sample from the set of single-mode Gaussian states.

### 3.2.1 Quantum estimation theory

The problem of parameter estimation has a long history, and was already studied in a classical framework. Typically, this involves a parameter dependent distribution  $p(y|\mu)$ , where  $y$  is the random variable being repeatedly and independently sampled with the goal of obtaining the best possible estimation of  $\mu$ . The quality of any estimator  $\tilde{\mu}$  can be quantified by its Root Mean Square Error (RMSE):

$$\delta\tilde{\mu} \equiv \sqrt{\sum_{\mathbf{y}} p(\mathbf{y}|\mu) (\tilde{\mu}(\mathbf{y}) - \mu)^2}, \quad (3.1)$$

where  $\mathbf{y}$  is a  $K$ -dimensional vector containing all the sampled values of  $y$ . One of the main results in classical estimation theory is the Cramér-Rao bound [137], stating that any unbiased estimator  $\tilde{\mu}$ , i.e. such that its expectation value is equal to  $\mu$ , satisfies the following inequality:

$$\delta\tilde{\mu} \geq \frac{1}{\sqrt{KF_\mu}}, \quad (3.2)$$

where  $K$  is the number of performed samplings and  $F_\mu$  is the *Fisher information* of the process, defined by

$$F_\mu = \int dy p(y|\mu) [\partial_\mu \log p(y|\mu)]^2. \quad (3.3)$$

It is important to point out that the bound in Equation (3.2) can always be asymptotically saturated in the limit  $K \rightarrow \infty$ , for example by choosing  $\tilde{\mu}$  to be the maximum likelihood estimator  $\tilde{\mu}_{\text{ML}} = \text{argmax}_\mu p(\mathbf{y}|\mu)$ . For this reason, the Fisher information  $F_\mu$  provides a good indication of the estimation precision obtainable by measuring the variable  $y$ .

Having discussed classical estimation theory, we now turn to its quantum counterpart. In this new scenario the parameter  $\mu$  is encoded in a quantum state  $\rho_\mu$ , which is typically obtained by applying a linear completely positive and trace-preserving (CPT) map  $\Phi_\mu$  to a known input probe  $\rho$ , so that

$$\rho_\mu = \Phi_\mu[\rho]. \quad (3.4)$$

Information about the encoding device, and thus on  $\mu$ , can be extracted by performing a positive operator valued measurement (POVM) on the encoded state  $\rho_\mu$ . This kind of measurement is characterised by a set of positive operators  $\{E_y\}_y$  such that  $\sum_y E_y = \mathbb{1}$ , where  $\mathbb{1}$  is the identity operator. The variable  $y$  labels the outcome of the measurement,

and it is obtained with probability

$$p|_{\{E_y\}_y}(y|\mu) = \text{Tr} [\rho_\mu E_y]. \quad (3.5)$$

At this stage the quantum problem has been mapped to a classical one, and we can compute the Fisher information  $F_\mu|_{\{E_y\}_y}$  associated with the chosen POVM through Equation (3.3). In order to obtain the ultimate precision allowed by quantum mechanics for any unbiased estimation of  $\mu$ , one can still optimise over all POVMs, obtaining in this way the *quantum Cramér-Rao bound* [109, 134]:

$$\delta\tilde{\mu} \geq \frac{1}{\sqrt{KH_\mu(\rho)}}. \quad (3.6)$$

The quantity  $H_\mu$  is the quantum Fisher information (QFI) associated with the encoded state  $\rho_\mu$ , which is obtained from  $\rho$  via Equation (3.4), and can be written as

$$H_\mu[\rho] = \text{Tr} [\rho_\mu L_\mu^2], \quad (3.7)$$

where  $L_\mu$  is an Hermitian operator known as “symmetric logarithmic derivative” (SLD), which satisfies the relation

$$\rho_\mu L_\mu + L_\mu \rho_\mu = 2 \partial_\mu \rho_\mu. \quad (3.8)$$

In order to obtain the ultimate precision for the estimation of the parameter  $\mu$  characterising the CPT encoding map  $\Phi_\mu$ , an optimisation has to be performed over the probe  $\rho$ . This maximal QFI is often called quantum Fisher information of the channel  $\Phi_\mu$  [138].

The implicit dependence of  $L_\mu$  upon  $\rho_\mu$  often makes Equation (3.7) difficult to use. Fortunately, it has been shown [139, 140] that in most cases the QFI can also be obtained from the second order expansion of the Bures distance [141], or equivalently of the Uhlmann fidelity  $\mathcal{F}(\rho_1, \rho_2) = (\text{Tr} [\sqrt{\sqrt{\rho_1} \rho_2 \sqrt{\rho_1}}])^2$  [142], as

$$H_\mu[\rho] = 8 \lim_{d\mu \rightarrow 0} \frac{1 - \sqrt{\mathcal{F}(\rho_\mu, \rho_{\mu+d\mu})}}{d\mu^2}. \quad (3.9)$$

Technically, this last equality only holds as long as the rank of  $\rho_{\mu+d\mu}$  does not change for  $d\mu = 0$ . If this is not the case, it has been recently shown that Equation (3.9) should be corrected by adding a term composed by second derivatives of the vanishing eigenvalues [143]. This fact has important consequences, and the QFI might even become discontinuous at those pathological points. Fortunately, we can safely ignore this issue in all the cases studied in the remainder of this chapter, because the rank of  $\rho_{\mu+d\mu}$  will not depend on  $d\mu$ .

It is worth to conclude this introduction to quantum estimation theory with some remarks. To begin with, we point out that a Fisher information equal to the QFI can

always be obtained by performing a projective measurement on the eigenbasis of the SLD operator  $L_\mu$  [134]. This fact guarantees that the quantum Cramér-Rao bound of Equation (3.6) is a priori tight, so that we can interpret the QFI as the ultimate precision bound allowed by quantum mechanics, even though this optimal readout strategy might not necessarily be easily implemented experimentally. Even in this situation, however, the QFI can be considered a good figure of merit for the probing state. Indeed, a large QFI value certifies that the probe is very susceptible to small changes of the parameter  $\mu$  characterising the considered evolution. As a second comment, we point out that in general the QFI depends on the parameter  $\mu$  itself. This is the reason why the QFI is known to characterise the ultimate precision attainable in a *local* estimation, where one already has some previous knowledge on  $\mu$ , and is only interested in finding small fluctuations around this approximately known value [144]. For example, this situation could be achieved through an adaptive strategy, where one uses a small fraction of the probes in order to select (a good approximation of) the optimal measurement, that will be then used to extract information from the remaining large number of encoded states. This situation is much simpler in the case of unitary encodings, for which  $\Phi_\mu[\rho]$  can be written as  $\hat{U}_\mu \rho \hat{U}_\mu^\dagger$ , where  $\hat{U}_\mu = e^{i\hat{\mathcal{H}}\mu}$  is a unitary operator in which the parameter  $\mu$  is a global phase multiplying a parameter-independent Hamiltonian  $\hat{\mathcal{H}}$  [145]. In this case,  $H_\mu(\rho)$  becomes independent from  $\mu$ , and the same can be said for the ultimate precision attainable by the probe  $\rho$  [134].

### 3.2.2 Gaussian states and operations

Differently from systems composed by quantum spins, a quantum description of light fields necessarily deals with an infinite number of degrees of freedom. Indeed, even if we consider a field composed by only one mode, its states can contain an arbitrary number of photons, not necessarily fixed. Because of this complexity, it is usually very difficult to study a quantum optical problem in full generality, i.e., by taking into account any possible state defined in the Fock space. For this reason, in addition to some physically relevant constraints (such as finite energy), one often limits the analysis to a particular subset of all quantum states, which is deemed to be both tractable and relevant. One of the most common choices is the set of Gaussian states, that we are going to introduce in this section. In addition to being amenable to analytical calculations, Gaussian states encompass most of the quantum states of light that can be experimentally generated and manipulated. In what follows we introduce some of the theoretical tools usually employed to easily characterise this class of states, but we will limit our discussion to the notions that are needed for the purposes of this chapter. Further details can be found on the many available reviews on the subject (e.g., see [99, 135, 136, 146]).

### Symplectic formalism

We will consider continuous variable systems composed by one or two bosonic modes, described by the annihilation operators  $\hat{a}$  and  $\hat{b}$ , and we will label by  $A$  and  $B$  their respective systems. Despite this choice, everything that we will say in this overview could be straightforwardly generalised to an arbitrary large number of modes. Together with their adjoints,  $\hat{a}$  and  $\hat{b}$  satisfy the canonical commutation relations

$$[\hat{a}, \hat{a}] = [\hat{b}, \hat{b}] = [\hat{a}, \hat{b}] = [\hat{a}, \hat{b}^\dagger] = 0, \quad [\hat{a}, \hat{a}^\dagger] = [\hat{b}, \hat{b}^\dagger] = 1, \quad (3.10)$$

where  $[\cdot, \cdot]$  represents the commutator, and they can be combined to form the following quadratures:

$$\hat{x}_A = \frac{\hat{a} + \hat{a}^\dagger}{\sqrt{2}}, \quad \hat{p}_A = \frac{\hat{a} - \hat{a}^\dagger}{i\sqrt{2}}, \quad \hat{x}_B = \frac{\hat{b} + \hat{b}^\dagger}{\sqrt{2}}, \quad \hat{p}_B = \frac{\hat{b} - \hat{b}^\dagger}{i\sqrt{2}}. \quad (3.11)$$

These can be conveniently organised within a single vector  $\hat{\mathbf{r}} = (\hat{x}_A, \hat{p}_A, \hat{x}_B, \hat{p}_B)^\top$ , which can be used to compactly express the properties of the system. For example, the commutation relations can be written in a compact form as  $[\hat{\mathbf{r}}, \hat{\mathbf{r}}^\top] = i\Omega$ , where  $\Omega$  is the direct sum of two (one per mode, in general)  $2 \times 2$  standard symplectic forms  $\Omega_1$ :

$$\Omega \equiv \bigoplus_{k=1}^2 \Omega_1 \equiv \bigoplus_{k=1}^2 \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (3.12)$$

Any real squared matrix  $T$  such that  $T\Omega T^\top = \Omega$  is said to be “symplectic”, and we write the set of these matrices as  $Sp$ . In the theory of continuous variable systems, and especially in that of Gaussian states, a special role is played by the operators that preserve the structure of these commutation relations when applied on  $\hat{\mathbf{r}}$ . These can be divided in two classes: displacement operators  $\hat{D}_{\mathbf{z}}$ , and Gaussian unitary operators  $\hat{U}_S$ . The connection between the latter and the set of Gaussian states will shortly be clear. The displacement operator is defined as  $\hat{D}_{\mathbf{z}} \equiv e^{-i\mathbf{z}^\top \Omega \hat{\mathbf{r}}}$ , but often it is enough to just consider its action on  $\hat{\mathbf{r}}$ :

$$\hat{D}_{\mathbf{z}}^\dagger \hat{\mathbf{r}} \hat{D}_{\mathbf{z}} = \hat{\mathbf{r}} + \mathbf{z}. \quad (3.13)$$

Similarly, for a Gaussian unitary one has

$$\hat{U}_T^\dagger \hat{\mathbf{r}} \hat{U}_T = T \hat{\mathbf{r}}, \quad (3.14)$$

for a symplectic matrix  $T \in Sp$ .

Other useful definitions are those of *displacement vector*  $\boldsymbol{\xi}$  and *covariance matrix*  $\Gamma$ , that can be defined for any quantum state  $\rho$  as

$$\boldsymbol{\xi} = \text{Tr} [\rho \hat{\mathbf{r}}], \quad \Gamma = \text{Tr} [\rho \{\hat{\mathbf{r}} - \boldsymbol{\xi}, \hat{\mathbf{r}} - \boldsymbol{\xi}^\top\}_+], \quad (3.15)$$

where  $\{\cdot, \cdot\}_+$  represents the anti-commutator. In order to be physically acceptable, any covariance matrix must satisfy the Robertson-Schrödinger uncertainty relation

$$\Gamma + i\Omega \geq 0, \quad (3.16)$$

while  $\xi$  can be any real vector with the same dimension of  $\hat{\mathbf{r}}$ . It is also worthwhile to introduce the quantum maps  $\mathcal{U}_T$  and  $\mathcal{D}_{\mathbf{z}}$ , which act on a state  $\rho$  as

$$\mathcal{U}_T[\rho] = \hat{U}_T \rho \hat{U}_T^\dagger, \quad \mathcal{D}_{\mathbf{z}}[\rho] = \hat{D}_{\mathbf{z}} \rho \hat{D}_{\mathbf{z}}^\dagger. \quad (3.17)$$

As a consequence, from the definitions in Equation (3.15) it follows that these evolutions change the displacement vector and covariance matrix of  $\rho$  as:

$$\Gamma \xrightarrow{\mathcal{U}_T} T \Gamma T^\top, \quad \xi \xrightarrow{\mathcal{U}_T} T \xi, \quad \Gamma \xrightarrow{\mathcal{D}_{\mathbf{z}}} \Gamma, \quad \xi \xrightarrow{\mathcal{D}_{\mathbf{z}}} \xi + \mathbf{z}. \quad (3.18)$$

The representation of infinite-dimensional quantum states through their expansion in the Fock basis is often inconvenient, and for this reason one typically turns to a phase space representation. In this approach the state of a quantum system is represented by a function of the classical counterparts of the quadratures defined in Equation (3.11). Although we will not enter in the details of this formalism, it is important to keep in mind that these functions are generally irregular, and cannot be considered proper probability distributions. Within the “zoo” of phase space representations, a special role is played by the characteristic function  $\chi_\rho(\mathbf{z})$  of a quantum state, defined as follows:

$$\chi_\rho(\mathbf{z}) = \text{Tr} \left[ \rho \hat{D}_{\mathbf{z}} \right], \quad (3.19)$$

which can be used to generate all other representations if desired [135]. We have now all the tools to define and characterise the set of Gaussian states [146].

**Definition 3.1** (Gaussian state). *A state  $\rho$  of a continuous variable system is said to be Gaussian if its characteristic function  $\chi_\rho(\mathbf{z})$  is the inverse Fourier transform of a Gaussian function:*

$$\chi_\rho(\mathbf{z}) = e^{-\frac{1}{4} \mathbf{z}^\top \Omega \Gamma \Omega^\top \mathbf{z} - i \mathbf{z}^\top \Omega \xi}, \quad (3.20)$$

for  $\xi$  and  $\Gamma$  related to  $\rho$  as in Equation (3.15).

A useful tool for the study of Gaussian states is Williamson’s theorem [147].

**Theorem 3.1** (Williamson). *Any covariance matrix  $\Gamma$  can be decomposed into the form*

$$\Gamma = T I T^\top. \quad (3.21)$$

Here,  $T$  is a matrix of the real symplectic group  $Sp$ , and  $I = \bigoplus_j \nu_j \mathbb{1}_2$  is a block-diagonal matrix whose blocks are multiples of the  $2 \times 2$  identity matrix. The number of blocks

corresponds to the number of modes in the system.

These proportionality coefficients  $\{\nu_j\}_j$  are called *symplectic eigenvalues* of the Gaussian state, they are also constrained by Equation (3.16) to be larger than or equal to 1, and can be found as the absolute values of the regular eigenvalues of  $i\Omega\Gamma$ . A pure Gaussian state is characterised by  $\nu_j \equiv 1$ , and we refer to a state with  $T = \mathbb{1}$  as to a “thermal state”. The reason behind this choice is that they can be written as thermal states of harmonic Hamiltonians. Let us show this through a simple example, which could be straightforwardly generalised to any number of modes.

**Example 3.1** (Single-mode thermal state). *Let  $\rho^{(\text{th})}$  be a thermal state for the single-mode harmonic Hamiltonian  $\hat{a}^\dagger\hat{a}$ , i.e.*

$$\rho^{(\text{th})} = \frac{e^{-\beta\hat{a}^\dagger\hat{a}}}{\text{Tr}[e^{-\beta\hat{a}^\dagger\hat{a}}]} = (1 - \beta) \sum_{n=0}^{\infty} \beta^n |n\rangle \langle n|, \quad (3.22)$$

for some  $\beta > 0$ . By explicitly computing its characteristic function  $\chi_{\rho^{(\text{th})}}$  through Equation (3.19), it can be shown that for this state  $\boldsymbol{\xi} = 0$  and  $\Gamma = \nu(\beta)\mathbb{1}_2$ , where

$$\nu(\beta) = \frac{1 + e^{-\beta}}{1 - e^{-\beta}} \geq 1, \quad (3.23)$$

with  $\nu \rightarrow 1$  when the state becomes pure in the limit  $\beta \rightarrow \infty$ .

Thanks to Williamson’s decomposition, it is possible to prove the following theorem on the structure of any Gaussian state.

**Theorem 3.2.** (Structure of a Gaussian state) *Any Gaussian state  $\rho_G$  can be obtained from a thermal state  $\rho^{(\text{th})}$  with the same number of modes by applying a Gaussian unitary map and a displacement map:*

$$\rho_G = \mathcal{D}_{\boldsymbol{\xi}} \circ \mathcal{U}_T[\rho^{(\text{th})}]. \quad (3.24)$$

If  $\{\nu_j\}_j$  are the symplectic eigenvalues of  $\rho^{(\text{th})}$ , the displacement vector and covariance matrix of  $\rho_G$  are respectively given by  $\boldsymbol{\xi}$  and  $\Gamma = T \left( \bigoplus_j \nu_j \mathbb{1}_2 \right) T^\top$ .

*Proof.* A sketch of the proof goes as follows. From the action of  $\hat{U}_T$  and  $\hat{D}_{\boldsymbol{\xi}}$  on the operator vector  $\hat{\mathbf{r}}$  described in Equations (3.13) and (3.14), one can easily prove that the quantum state  $\rho_G$  is Gaussian [i.e., that its characteristic function has the form of Equation (3.20)]. In order to show that its displacement vector and covariance matrix are those stated in the theorem, we can use Equation (3.18) to reduce the problem to the evaluation of  $\boldsymbol{\xi}$  and  $\Gamma$  for a thermal state, which are known from Example 3.1. Finally, Williamson’s decomposition guarantees that all Gaussian states can be obtained in this way.  $\square$

It is also important to mention the so-called “Euler decomposition” [135].

**Theorem 3.3.** (*Euler decomposition*) Any symplectic matrix  $T$  can be written as  $O_1 S O_2$ , where  $O_1, O_2 \in Sp \cap SO$ , and  $S$  is direct sum of  $2 \times 2$  diagonal matrices (not necessarily equal) of the form  $\text{Diag}(x, 1/x)$ , with  $x > 0$ .

As a consequence of this result, any Gaussian unitary  $\hat{U}_T$  can be split into three simpler terms:  $\hat{U}_T = \hat{U}_{O_1} \hat{U}_S \hat{U}_{O_2}$ , and similarly any Gaussian map  $\mathcal{U}_T$  can be decomposed as  $\mathcal{U}_{O_1} \circ \mathcal{U}_S \circ \mathcal{U}_{O_2}$ . For example, this can be used in Theorem 3.2 in order to construct any Gaussian state by sequentially applying simpler unitary evolutions on a thermal state. We will shortly describe the physical meaning of these simple evolutions for a single-mode field. In general, we stress that it is extremely convenient to describe (evolution of) Gaussian states only in terms of (changes in) their covariance matrices and displacement vectors. This symplectic formalism can always be translated into the actions of unitary operators on quantum states via Equation (3.18).

### Gaussian states and operations of interest for this chapter

After a general introduction to the symplectic formalism used to describe Gaussian states and operations, we now turn our attention to the specific examples that will be considered in this chapter. In particular, we will extensively deal with Gaussian unitary maps that can squeeze or rotate a single-mode Gaussian state in phase space. These can be respectively written as  $\mathcal{S}_\alpha$ ,  $\mathcal{R}_\theta$ , and the Gaussian unitary operators associated to them as in Equation (3.17) are:

$$\hat{U}_{S_\alpha} = e^{-\frac{\alpha}{2}(\hat{a}^{\dagger 2} - \hat{a}^2)}, \quad \hat{U}_{R_\theta} = e^{-i\theta \hat{a}^\dagger \hat{a}}. \quad (3.25)$$

In turn, the symplectic matrices describing the action of these operators on the vector  $\hat{\mathbf{r}}$  via Equation (3.14) are:

$$S_\alpha = \begin{pmatrix} e^\alpha & 0 \\ 0 & e^{-\alpha} \end{pmatrix}, \quad R_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}, \quad (3.26)$$

which also play the role of  $T$  in Equation (3.18). By inspection, we see that  $S_\alpha$  squeezes the quadrature  $p_A$ , at the expenses of  $x_A$ , whose variance becomes larger, while  $R_\theta$  corresponds to a simple rotation in phase space. When the first of these Gaussian unitary operations is applied to the vacuum state, we obtain a single-mode squeezed state, that in the following will be labelled by  $\rho^{(\text{sq})}$ .

At this point, we can also provide a compact expression for a generic single-mode Gaussian state  $\rho_A$  that will become useful in the following. By applying Euler decomposition to the symplectic matrix  $T$  appearing in Equation (3.24), for a single-mode state one has

$$\rho_A = \mathcal{D}_\xi \circ \mathcal{U}_{R_\phi} \circ \mathcal{U}_{S_\alpha} [\rho^{(\text{th})}(\nu)], \quad (3.27)$$

where  $R_\phi$  and  $S_\alpha$  play respectively the roles of  $O_1$  and  $S$  in Euler decomposition. Although Theorem 3.3 would require an additional map  $\mathcal{U}_{O_2}$ , we can omit it because it would leave any single-mode thermal state unaltered. Intuitively, this expansion is telling us that any single-mode Gaussian state can be obtained by taking the vacuum, adding some thermal noise, squeezing it along  $x$  and  $p$ , and finally by rotating and displacing it in phase space.

If we move to a two-mode scenario, there is a paradigmatic class of states that can be obtained by applying a particular Gaussian unitary to the vacuum state, known as “two-mode squeezing”:

$$\hat{U}_{S_r^{(2)}} = e^{r(\hat{a}^\dagger \hat{b}^\dagger - \hat{a} \hat{b})}. \quad (3.28)$$

This operator is characterised by the symplectic matrix

$$S_r^{(2)} = \left( \begin{array}{c|c} \cosh r \mathbb{1} & \sinh r \sigma_z \\ \hline \sinh r \sigma_z & \cosh r \mathbb{1} \end{array} \right), \quad (3.29)$$

with  $\sigma_z = \text{Diag}(+1, -1)$ . An explicit calculation yields the following expansion for the two-mode squeezed vacuum state [135]:

$$\rho_{AB}^{(\text{ts})} = |\psi_r^{(\text{ts})}\rangle_{AB} \langle \psi_r^{(\text{ts})}|, \quad |\psi_r^{(\text{ts})}\rangle_{AB} = \hat{U}_{S_r^{(2)}} |0\rangle = \frac{1}{\cosh r} \sum_{n=0}^{\infty} \tanh(r)^n |nn\rangle_{AB}, \quad (3.30)$$

which shows that the two modes are highly correlated. Indeed, in the limit of infinite squeezing  $|\psi_r^{(\text{ts})}\rangle_{AB}$  tends to a uniform superposition of Fock elements of the form  $|nn\rangle_{AB}$ , and it is therefore considered the counterpart of a maximally entangled state in continuous variables systems.

Of course, there are CPT maps that preserve the Gaussian character of a state even without being Gaussian unitary evolutions [148]. One example of this sort that will be considered in the following is the lossy channel  $\mathcal{L}_\eta$ , parametrised by  $\eta \in [0, 1]$ . When applied on the first mode of a two-mode state, we can describe its action on the whole two-mode input covariance matrix  $\Gamma$  and displacement vector  $\xi$  as:

$$\Gamma \xrightarrow{\mathcal{L}_\eta^{(A)}} K_\eta \Gamma K_\eta^\top + N_\eta N_\eta^\top, \quad \xi \xrightarrow{\mathcal{L}_\eta^{(A)}} K_\eta \xi, \quad (3.31)$$

where  $K_\eta$  and  $N_\eta$  are the following block-diagonal matrices

$$K_\eta = \left( \begin{array}{c|c} \sqrt{\eta} \mathbb{1}_2 & \\ \hline & \mathbb{1}_2 \end{array} \right), \quad N_\eta = \left( \begin{array}{c|c} \sqrt{1-\eta} \mathbb{1}_2 & \\ \hline & \end{array} \right), \quad (3.32)$$

with empty blocks being composed only by zeros.



### 3.2.3 Quantum Fisher information for Gaussian states

In this section we bridge the previous discussions on estimation theory and Gaussian states. Because of their practical relevance and simple structure, the latter have been extensively studied. In particular, many information theoretical quantifiers have been computed and expressed as function of their covariance matrices and displacement vectors. Here we focus our attention on the QFI. By expanding the Bures distance up to the second order as in Equation (3.9), the QFI has been explicitly evaluated for single-mode [149], two-mode [106] and recently for generic multi-mode Gaussian states [150]. Thanks to these tools, several authors have been able to calculate the performance reached by Gaussian states in different estimation tasks [107, 151].

As we will deal with two modes at most, we report here the expression for the two-mode QFI obtained and studied in Ref. [106], which will play a central role in the remainder of this chapter

$$\begin{aligned}
H_\mu[\rho_{AB}] = & \frac{1}{2(|\tilde{M}| - 1)} \left\{ |\tilde{M}| \text{Tr} \left[ (\tilde{M}^{-1} \dot{\tilde{M}})^2 \right] + \sqrt{|\mathbb{1} + \tilde{M}^2|} \text{Tr} \left[ \left( (\mathbb{1} + \tilde{M}^2)^{-1} \dot{\tilde{M}} \right)^2 \right] \right\} \\
& + \frac{4}{2(|\tilde{M}| - 1)} (\tilde{\nu}_1^2 - \tilde{\nu}_2^2) \left( -\frac{\dot{\tilde{\nu}}_1^2}{\tilde{\nu}_1^4 - 1} + \frac{\dot{\tilde{\nu}}_2^2}{\tilde{\nu}_2^4 - 1} \right) + 2\dot{\tilde{\xi}}^\top \tilde{\Gamma}^{-1} \dot{\tilde{\xi}}.
\end{aligned} \tag{3.33}$$

Here the determinant of a matrix is represented by the symbol  $|\cdot|$ , the dot corresponds to the derivative with respect to the squeezing parameter  $\mu$  being estimated, and we defined  $M \equiv i\Omega\Gamma$ . We also remind the reader that  $\xi$ ,  $\Gamma$ , and  $\{\nu_i\}_{i=1}^2$  are respectively the displacement vector, the covariance matrix, and the symplectic eigenvalues of a two-mode Gaussian state, and that  $\Omega$  is the symplectic form. We added a tilde on top of all the involved quantities in order to remind us to compute them on the encoded state  $\Phi_\mu[\rho_{AB}]$ , that is obtained after the input probe  $\rho_{AB}$  completed its evolution. This formula is rather involved, but it can be split in three contributions. The first one takes into account the variation of the covariance matrix as a whole, the second one explicitly depends on the variation of the symplectic eigenvalues, while the third and last one is due to a change in the displacement vector. Special care has to be used in dealing with the second term, because it might lead to irregular behaviours when one or more symplectic eigenvalues become equal to 1 for the parameter  $\mu$  being investigated. We will not further discuss this issue, because this pathological condition will not be realised in the physical scenario we consider<sup>1</sup>, but more details can be found in Refs. [106, 143].

At a certain point, when studying squeezing estimation in presence of losses, it will become useful to restrict some of our analysis to single-mode Gaussian states. It is therefore important to discuss how Equation (3.33) can be simplified in this regime.

---

<sup>1</sup>As a matter of fact, we will investigate either: (i) unitary encodings, which do not change the symplectic eigenvalues, or (ii) mixed encoded states with symplectic eigenvalues strictly larger than 1.

We should say that an expression for the single-mode QFI has been known for a few years [149], but discussing the limit of the two-mode expression previously reported is going to be useful nonetheless. The conclusion of this considerations will be the only original result presented in this preliminary section.

Intuitively, the QFI quantifies the susceptibility of a probe under small changes in the value of the parameter which characterises its evolution. In the scenario that will be discussed in next section, only system  $A$  will be actively altered by the encoding channel. System  $B$ , instead, will be left untouched and kept as a reference. This situation can be formally expressed by saying that the encoding evolution can be written in a factorised form as

$$\Phi_\mu = \Phi_\mu^{(A)} \otimes \mathbb{1}_B. \quad (3.34)$$

However, this does not mean that system  $B$  cannot play a role in the estimation, because of the correlations that it could share with the “active” probe  $A$ . On the other hand, it can be easily proven that if the two-mode probe can be written in a factorised form  $\rho_{AB} = \rho_A \otimes \rho_B$ , then the QFI for the channel  $\Phi_\mu$  applied to  $\rho_{AB}$  is equal to the QFI of the single-mode channel  $\Phi_\mu^{(A)}$  acting on  $\rho_A$ . Perhaps surprisingly, the way in which Equation (3.33) has been written is hiding this property behind some non obvious relation among the quantities involved. Therefore, we can use this general argument of independence upon  $\rho_B$  as a trick to recover a relation between the quantities appearing in the single-mode QFI. This new relation can then be used to rewrite the single-mode QFI in a equivalent, but slightly different, way with respect to the expression found in Ref. [149].

After this rather long declaration of intents, let us apply the aforementioned trick by evaluating Equation (3.33) for a specific factorised probe  $\rho_A \otimes \rho_{th}(\nu_B)$ , where  $\rho^{(th)}(\nu_B)$  is a thermal Gaussian state with symplectic eigenvalue  $\nu_B$ . Because the channel we are considering has a factorised structure as in Equation (3.34), all the derivatives are applied only on subsystem  $A$ , and after some straightforward manipulation we can see that  $H_\mu(\rho_A \otimes \rho^{(th)}(\nu_B))$  is independent from  $\nu_B$  only if

$$\frac{1}{|\tilde{M}_A|} \left( \frac{d|\tilde{M}_A|}{d\mu} \right)^2 = |\tilde{M}_A| \text{Tr} \left[ \left( \tilde{M}_A^{-1} \dot{\tilde{M}}_A \right)^2 \right] - (1 - |\tilde{M}_A|)^2 \text{Tr} \left[ [(\mathbb{1} + \tilde{M}_A^2)^{-1} \dot{\tilde{M}}_A]^2 \right], \quad (3.35)$$

where  $\tilde{M}_A$  is evaluated on the encoded single-mode state  $\Phi_\mu^{(A)}(\rho_A)$ . Although this equality has been obtained by considering a two-mode state  $\rho_A \otimes \rho^{(th)}(\nu_B)$ , in the following we will use it only in order to simplify the QFI expression of single-mode states. Indeed, the single-mode QFI depends on the three terms:

$$\frac{d|\tilde{M}_A|}{d\mu}, \quad \text{Tr} \left[ (\tilde{M}_A^{-1} \dot{\tilde{M}}_A)^2 \right], \quad \text{Tr} \left[ [(\mathbb{1} + \tilde{M}_A^2)^{-1} \dot{\tilde{M}}_A]^2 \right]. \quad (3.36)$$

The non trivial relation between  $\tilde{M}_A$  and its derivative, that has been found in Equ-

tion (3.35) via the aforementioned trick, can now be used to rewrite either one of these components as a function of the other two. This allows us to keep only the two contributions that are easier to evaluate. For example, if one removes the explicit dependence on  $\text{Tr} \left[ (\tilde{M}_A^{-1} \dot{\tilde{M}}_A)^2 \right]$ , the full expression for the single-mode QFI, which is labelled by  $H_\mu^{(1)}$ , becomes

$$H_\mu^{(1)}(\rho_A) = \frac{|\tilde{M}_A| - 1}{2} \text{Tr} \left[ [(\mathbb{1} + \tilde{M}_A^2)^{-1} \dot{\tilde{M}}_A]^2 \right] + \frac{1}{2(|\tilde{M}_A|^2 - 1)} \left( \frac{d|\tilde{M}_A|}{d\mu} \right)^2 + 2\dot{\tilde{\boldsymbol{\xi}}}_A^\top \tilde{\Gamma}_A^{-1} \dot{\tilde{\boldsymbol{\xi}}}_A. \quad (3.37)$$

On the other hand, the expression provided in Ref. [149] can be recovered by removing the explicit dependence on the last term appearing in Equation (3.36). In our case it will be convenient to work with Equation (3.37), because  $\text{Tr} \left[ [(\mathbb{1} + \tilde{M}_A^2)^{-1} \dot{\tilde{M}}_A]^2 \right]$  will be easier to compute than  $\text{Tr} \left[ (\tilde{M}_A^{-1} \dot{\tilde{M}}_A)^2 \right]$  (see Sec. 3.4.2).

### 3.2.4 Uniform sampling of single-mode Gaussian states

In order to gain a preliminary understanding about the properties of a given set, a simple idea consists in acquiring data by randomly sampling from its elements. If the size of the sample is large enough, one would be tempted to infer conclusions about the set as a whole. Although this is not completely rigorous, in many cases this sort of approach works reasonably well, as long as the samples are uniformly distributed. In our specific case, we would like to have a way to uniformly sample from the set of single-mode Gaussian states with the same energy. In the following, we will use this numerical approach whenever we need to randomly sample single-mode Gaussian states. For example, this will be done in Sec. 3.4.2 in order to check that in lossy conditions pure single-mode states yield the best estimation performance. We now introduce a sampling method based on the Haar measure defined on the group of Gaussian unitary operators. Further details on this subject, or multi-mode generalisations, can be found in Ref. [152] or in references therein.

Let us start by discussing pure single-mode Gaussian states, which can be written as

$$|\psi_G^{(1)}\rangle\langle\psi_G^{(1)}| = \mathcal{U}_G^{(A)} [|0\rangle\langle 0|], \quad (3.38)$$

for a generic single-mode Gaussian map  $\mathcal{U}_G^{(A)}$ . According to Williamson's theorem and Euler's decomposition, this can be expanded as

$$\mathcal{U}_G^{(A)} = \mathcal{D}_{\boldsymbol{\xi}_A}^{(A)} \circ \mathcal{R}_\phi^{(A)} \circ \mathcal{S}_\alpha^{(A)} \circ \mathcal{R}_{\phi'}^{(A)}. \quad (3.39)$$

By writing the displacement vector in polar coordinates

$$\boldsymbol{\xi}_A = (|\boldsymbol{\xi}| \cos \psi, |\boldsymbol{\xi}| \sin \psi)^\top, \quad (3.40)$$

the Haar invariant measure on the group of 1-mode Gaussian unitary maps is then given by [152]

$$d(\mathcal{U}_G^{(A)}) = \frac{\mathcal{N}_1}{2} d(\cosh \alpha) d\phi d\phi' d(|\xi|^2) d\psi, \quad (3.41)$$

up to a normalisation constant  $\mathcal{N}_1$ . As we already mentioned, the phase  $\phi'$  has no effect on the vacuum state because of its rotational symmetry. The invariant measure induced on the set of single-mode pure states is thus

$$d(\psi_G^{(1)}) = \frac{\mathcal{N}_1}{2} d(\cosh \alpha) d\phi d(|\xi|^2) d\psi. \quad (3.42)$$

For a mixed single-mode Gaussian state, instead, we can first sample a pure two-mode state, with an approach similar to the one just described, and then apply a partial trace over one of the two subsystems. This is a standard approach in generating random mixed quantum states (see e.g., Ref. [153]). Following Ref. [152], a pure two-mode Gaussian state can be written as

$$|\psi_G^{(2)}\rangle\langle\psi_G^{(2)}| = \mathcal{U}_G^{(A)} \otimes \mathcal{U}_G^{(B)} [|\psi^{(\text{ts})}(\nu_A)\rangle_{AB}\langle\psi^{(\text{ts})}(\nu_A)|], \quad (3.43)$$

where  $|\psi^{(\text{ts})}(\nu_A)\rangle_{AB}$  is a two-mode squeezed vacuum state

$$|\psi^{(\text{ts})}(\nu_A)\rangle_{AB} = \sqrt{\frac{2}{\nu_A + 1}} \sum_{j=0}^{\infty} \left(\frac{\nu_A - 1}{\nu_A + 1}\right)^{j/2} |j, j\rangle_{AB}. \quad (3.44)$$

Note that this parametrisation is different from, although equivalent to, the one given in Equation (3.30). Here, the parameter  $\nu_A$  corresponds to the symplectic eigenvalue of the reduced single-mode thermal state that is obtained by tracing away mode  $B$ . According to Ref. [152], the invariant measure on the manifold of pure two-mode Gaussian states is

$$d(\psi_G^{(2)}) = \frac{\mathcal{N}_2}{3} d(\nu_A^3) d\mathcal{U}_G^{(A)} d\mathcal{U}_G^{(B)}. \quad (3.45)$$

By applying partial trace on mode  $B$ , the invariant measure for a mixed single-mode Gaussian state becomes:

$$d(\rho_G^{(1)}) = \frac{\mathcal{N}_2}{3} d(\nu_A^3) d\mathcal{U}_G^{(A)} = \frac{\mathcal{N}_2 \mathcal{N}_1}{6} d(\nu_A^3) d(\cosh \alpha) d(|\xi|^2) d\phi d\psi, \quad (3.46)$$

where we used Equation (3.41) without the irrelevant angle  $\phi'$ , which has no effect on a thermal state.

From this result we see that in order to uniformly sample from the set of single-mode Gaussian states we should sample the parameters  $\nu_A^3, \cosh \alpha, |\xi|^2, \phi$  and  $\psi$ . However, in order to do so, we should first limit the region in which the first three parameters can take values. In particular, we choose the physically motivated approach of considering

states with the same energy, i.e., composed by the same average number of photons  $n_A$ . An explicit calculation yields:

$$n_A = \text{Tr} \left[ \hat{a}^\dagger \hat{a} \rho_G^{(1)} \right] = \frac{\nu_A \cosh(2\alpha) - 1 + |\xi|^2}{2}. \quad (3.47)$$

This constraint can be formally taken into account by adding to the right-hand side of Equation (3.46) the Dirac delta function

$$\delta \left( n_A - \frac{\nu_A \cosh(2\alpha) - 1 + |\xi|^2}{2} \right). \quad (3.48)$$

Practically, in order to uniformly sample a mixed single-mode Gaussian state with  $n_A$  photons with respect to this invariant measure, we need to: (i) uniformly sample  $\nu_A^3$ ,  $\cosh(\alpha)$  and  $|\xi|^2$  within the region of  $\mathbb{R}^3$  allowed by the energy constraint, and (ii) uniformly sample  $\phi, \psi$  in  $[0, 2\pi]$ . In the following, we will use this method in order to numerically study the set of isoenergetic single-mode Gaussian states.

### 3.3 Quantification of probe versatility

The problem of quantifying the versatility of a given probing system lies at the heart of this chapter. However, we can interpret the concept of versatility in many different ways, and in its quantification we should take into account the degree of uncertainty that we have on the parameter-encoding evolution. In order to fix the ideas on a particular scenario, we start this section by introducing the details of the physical problem under investigation. Then we take a step back, and in Section 3.3.2 we summarise the relevant assumptions characterising the problem, and show that within this framework a natural quantifier for probe versatility can be identified in the average value of its QFI. The additional features that emerge when this quantifier is applied to the problem of squeezing estimation are summarised in Section 3.3.3.

#### 3.3.1 The model

In the metrological problem under investigation we aim at estimating the strength  $\mu$  of the squeezing applied on the probing state, in a regime where we do not know in advance the direction  $\theta$  in which the squeezing operation is applied. For the sake of simplicity, we will focus on the case where  $\theta$  is uniformly distributed in  $[0, 2\pi]$ . This situation could arise in two ways. Either the angle  $\theta$  is fixed, but initially unknown when we have to prepare *all* the probes that will be used for the following estimations, or it can fluctuate randomly from one interaction to the next. In absence of noise, the overall encoding CPT

map acting on mode  $A$  for a particular squeezing direction can be written as:

$$\Phi_{\mu,\theta}^{(A)} \equiv \mathcal{S}_{\mu}^{(A)} \circ \mathcal{R}_{\theta}^{(A)}, \quad (3.49)$$

where the Gaussian unitary maps  $\mathcal{S}_{\mu}^{(A)}, \mathcal{R}_{\theta}$  have been previously introduced just before Equation (3.25). In our analysis we will also be interested in seeing whether the presence of input correlations can help in the estimation. In this case, we will consider a two-mode probe, where subsystem  $A$  is sent to the squeezing device, whereas the  $B$  component is kept as reference. The total encoding channel in this more general case will be

$$\Phi_{\mu,\theta} = \Phi_{\mu,\theta}^{(A)} \otimes \mathbb{1}^{(B)}. \quad (3.50)$$

The core assumption that we make is that the value of  $\theta$  characterizing the evolution is revealed to the experimenter just before the measurement stage, but it is unknown when the probes are being created. Therefore, although the presence of this phase rotation cannot be compensated by opportunely changing the probe, one can still take the information into account when designing the optimal readout measurement. It is because of this assumption that the average quantum Fisher information becomes a figure of merit for the protocol, as we will formally discuss in the next section. Indeed, if the phase  $\theta$  was to remain unknown for the whole estimation process, one would have no other choice than trying to evaluate the quantum Fisher information of the average output state: a difficult task because of the non Gaussianity of the latter. We can think at two scenarios in which the parameter  $\theta$  is not known when the probes are being prepared. In the first one, the experimenter might be required to prepare all probes at once, before the experimental apparatus has been chosen. In the second one, instead, the parameter  $\theta$  can change from one probe to another. We will provide more details about these two situations in the next section.

To be more concrete, we now discuss how a CPT map as in Equation (3.49) could be realised in an optical setup, schematically represented in Figure 3.1. In this example, the uncertainty on  $\theta$  can be due to an unknown (or fluctuating) distance from the squeezing device, which applies the map  $\mathcal{S}_{\mu}^{(A)}$  to any state reaching its position. When travelling to and from the squeezer, the light will be affected by the free evolution  $\mathcal{R}_{\theta}^{(A)}$ , so that the overall probe evolution will be given by  $\mathcal{R}_{\theta}^{(A)} \circ \mathcal{S}_{\mu}^{(A)} \circ \mathcal{R}_{\theta}^{(A)}$ . In this example, the value of  $\theta$  could be inferred by measuring the travel time of the light. Alternatively, if the setup has been properly calibrated and  $\theta$  does not change throughout the experiments, its value could be independently estimated once the versatile probes have been created, and the estimation is about to begin. Notice that the form of the encoding map can be effectively changed from  $\mathcal{R}_{\theta}^{(A)} \circ \mathcal{S}_{\mu}^{(A)} \circ \mathcal{R}_{\theta}^{(A)}$  to the form written in Equation (3.49), because upon receiving the encoded state it is possible to exploit the knowledge acquired on  $\theta$  and apply the unitary correction  $\mathcal{R}_{-\theta}^{(A)}$ . This will not change the quantum Fisher information of

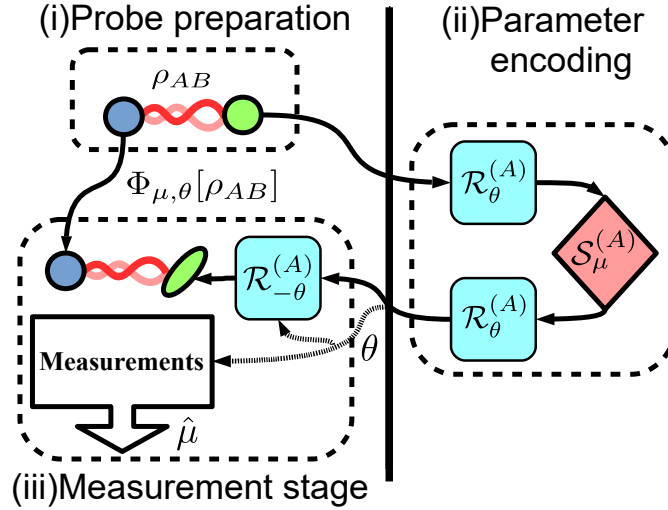


Figure 3.1: Sketch of a physical scenario in which a squeezing operation along an initially unknown direction is applied on the probing system. Stage (i): versatile single-mode or two-mode probes are prepared, without knowing the angle  $\theta$  that will characterise their evolution. Stage (ii): the probing systems are sent to the squeezing device, and in each one-way trip they acquire a certain phase  $\theta$ . Stage (iii): upon recollection of the probes the angle  $\theta$  becomes known, and can be used to devise the best measurement allowed by quantum mechanics, from which an estimator  $\hat{\mu}$  is recovered. Note that the knowledge of  $\theta$  can also be used to apply the unitary correction  $\mathcal{R}_{-\theta}^{(A)}$  on the final state of the probes.

the output state, because the unitary correction is known and the optimal measurement process can be changed accordingly. We point out that for the exact same reason we could equivalently change the probe evolution in Equation (3.49) into  $\mathcal{R}_{-\theta}^{(A)} \circ \mathcal{S}_{\mu}^{(A)} \circ \mathcal{R}_{\theta}^{(A)}$ , simply by applying a correction  $\mathcal{R}_{-\theta}^{(A)}$  on the probe after its round trip to and from the squeezing device. Therefore, for the purposes of evaluating the AvQFI, the CPT map in Equation (3.49) is equivalent to a map which squeezes the input state along a  $\theta$ -dependent direction in phase space.

### 3.3.2 Average quantum Fisher information

With this physical model in mind, we now take a step back and rephrase the same problem on a more abstract level. This allows us to show that the average value of the QFI (AvQFI in the following) can be considered a meaningful quantifier of the ultimate average precision attainable by each probe, and therefore of its versatility. The important facts are that with a certain probability  $p(\theta)$  the encoding map  $\Phi_{\mu, \theta}$  is applied on the probe and that the chosen value of  $\theta$  is then communicated to the experimenter, who can use this information in order to perform the best possible readout POVM with the aim of estimating  $\mu$ .

At first, let us analyse a scenario of black-box metrology, where  $\theta$  is initially unknown but fixed. According to the quantum Cramér-Rao bound (3.6), in this case the RSME of



any unbiased estimator obtained by performing the optimal measurement on the evolution of  $K$  copies of a probing state  $\rho$  will depend on  $\theta$

$$\delta\hat{\mu}(\theta) \geq \frac{1}{\sqrt{K H_{\mu}^{(\theta)}(\rho)}}, \quad (3.51)$$

with  $H_{\mu}^{(\theta)}(\rho)$  being the QFI for  $\mu$  evaluated on  $\Phi_{\mu,\theta}[\rho]$ . The versatility of an input state  $\rho$  can thus be assessed by looking at the average among all possible values of QFI, i.e., its AvQFI

$$\overline{H}_{\mu}(\rho) \equiv \int_0^{2\pi} d\theta p(\theta) H_{\mu}^{(\theta)}(\rho). \quad (3.52)$$

This is because from Equation (3.51) and from the convexity of the function  $1/\sqrt{x}$ , it follows that  $\overline{H}_{\mu}(\rho)$  sets a lower bound on the average value of the attainable RMSEs, i.e.

$$\overline{\delta\hat{\mu}} \equiv \int_0^{2\pi} d\theta p(\theta) \delta\hat{\mu}(\theta) \geq \frac{1}{\sqrt{K \overline{H}_{\mu}(\rho)}}. \quad (3.53)$$

In the second scenario of interest, the evolution of each probe is characterized by a fixed value of  $\theta$ , but this value can change from one experiment to another. However, just before the measurement stage the information on the value of  $\theta$  that affected the evolution of the probe becomes available, and can be used to choose a suitable POVM  $\{E_x^{(\theta)}\}_x$  tailored to the specific realisation of the encoding map  $\Phi_{\mu,\theta}$ . This measurement, applied on  $\rho_{\mu,\theta} = \Phi_{\mu,\theta}[\rho]$ , yields an outcome  $x$  with probability  $p(x|\mu, \theta) = \text{Tr} [\rho_{\mu,\theta} E_x^{(\theta)}]$ . After repeating this procedure for each of the  $K$  probes, we are left with a classical problem in which  $\mu$  needs to be estimated from the knowledge of  $K$  pairs  $(x, \theta)$ , sampled according to the distribution

$$p(x, \theta|\mu) = p(x|\mu, \theta)p(\theta). \quad (3.54)$$

The Fisher information associated with this probability distribution is given by:

$$F_{\mu} = \int dx d\theta p(x, \theta|\mu) [\partial_{\mu} \log p(x, \theta|\mu)]^2 = \int d\theta p(\theta) F_{\mu}^{(\theta)}, \quad (3.55)$$

where  $F_{\mu}^{(\theta)}$  is the classical Fisher information of the probability distribution  $p(x|\mu, \theta)$  obtained for fixed  $\theta$ . Therefore, the classical Cramér-Rao bound [see Equation (3.2)] yields the following lower bound on the RMSE of any unbiased estimator  $\hat{\mu}_F$ :

$$\delta\hat{\mu}_F \geq \frac{1}{\sqrt{K \int d\theta p(\theta) F_{\mu}^{(\theta)}}}, \quad (3.56)$$

where the subscript “ $F$ ” reminds us of the fluctuating regime that we are considering.



At last, the right-hand side can be minimised by suitably choosing, for every  $\theta$ , the POVM maximising the Fisher information  $F_\mu^{(\theta)}$ . This optimisation allows us to substitute  $F_\mu^{(\theta)}$  with the quantum Fisher information  $H_\mu^{(\theta)}(\rho)$ , so that we are left with the following ultimate bound on the variance of any unbiased estimator:

$$\delta\hat{\mu}_F \geq \frac{1}{\sqrt{K\overline{H}_\mu(\rho)}}, \quad (3.57)$$

which is characterised once again by the AvQFI of the probe.

As a final remark, we point out that the AvQFI is a convex function of the input state, because it inherits this property from the QFI.

### 3.3.3 AvQFI for squeezing estimation

Differently than in the previous section, in the remainder of this chapter we will always consider a uniform distribution for  $p(\theta)$ , so that for a two-mode probe  $\rho_{AB}$  the AvQFI takes the form

$$\overline{H}_\mu(\rho_{AB}) \equiv \int_0^{2\pi} \frac{d\theta}{2\pi} H_\mu^{(\theta)}(\rho_{AB}). \quad (3.58)$$

We will also explicitly label by  $\overline{H}_\mu^{(1)}$  the AvQFI associated with a single-mode probing system.

The AvQFI has a few interesting properties, which can be exploited in order to simplify the analyses in the following sections. In particular, it is invariant under phase rotations on mode  $A$  and generic unitary evolutions  $\mathcal{U}^{(B)}$  applied on mode  $B$ , i.e.,

$$\overline{H}_\mu \left[ \mathcal{R}_\phi^{(A)} \otimes \mathcal{U}^{(B)}[\rho_{AB}] \right] = \overline{H}_\mu[\rho_{AB}]. \quad (3.59)$$

In order to see this, at first notice that  $\mathcal{U}^{(B)}$  commutes with the encoding channel  $\Phi_{\mu,\theta}$ . Since any known map applied after the encoding channel can be absorbed in the measurement process, the QFI is left unaltered by the application of  $\mathcal{U}^{(B)}$ . Then, notice that

$$\Phi_{\mu,\theta} \circ \mathcal{R}_\phi^{(A)}[\rho_{AB}] = \mathcal{R}_\phi^{(A)} \circ \Phi_{\mu,\theta+\phi}[\rho_{AB}]. \quad (3.60)$$

Once again, the map  $\mathcal{R}_\phi^{(A)}$  on the right-hand side can be included in the readout process, while the shift in  $\theta$  does not affect the AvQFI value because of the uniform average appearing in Equation (3.58). Importantly, the symmetry of the AvQFI shown in Equation (3.59) can be exploited in order to significantly simplify the structure of the set of two-mode probes that have to be considered in our search for the optimal average performance. In particular, it is enough to consider Gaussian states written in “standard form” as detailed in the following theorem, whose proof can be found at the end of the section. Should the reader be familiar with the topic of quantum correlations in Gaussian

states, we point out that the standard form used in this chapter is different from the one typically used in the studies of Gaussian entanglement. This is because generic Gaussian unitary operations on mode  $A$  do not leave the AvQFI unaltered.

**Theorem 3.4.** (*Standard form of a two-mode Gaussian state*) *For any two-mode Gaussian probe there exists a Gaussian state in “standard form” with the same AvQFI. The latter is characterised by a covariance matrix and a displacement vector with the following structure:*

$$\Gamma^{(\text{std})} = \left( \begin{array}{cc|cc} a_x & a_{xp} & c & 0 \\ a_{xp} & a_p & 0 & d \\ \hline c & 0 & b & 0 \\ 0 & d & 0 & b \end{array} \right), \quad \boldsymbol{\xi}^{(\text{std})} = (\xi_x, \xi_p, 0, 0)^\top. \quad (3.61)$$

The parametrisation in Equation (3.61) is very convenient for two-mode input states, but not particularly useful whenever single-mode probes are explicitly considered. In this second case, a better approach consists in using Equation (3.27), from which it follows that the covariance matrix and the displacement vector of a generic single-mode Gaussian state  $\rho_A$  can be parametrised as

$$\Gamma_A = \nu_A R_\phi S_{2\alpha} R_\phi^\top, \quad \boldsymbol{\xi}_A = |\xi| \begin{pmatrix} \cos \psi \\ \sin \psi \end{pmatrix}, \quad (3.62)$$

where  $\nu_A$  characterises its thermal excitations,  $\alpha$  and  $\phi$  respectively quantify the amount and the direction of squeezing, while  $\psi$  fixes the displacement direction. When a phase rotation  $\mathcal{R}_\theta^{(A)}$  is applied on a state with the form of Equation (3.62), the parameters  $\phi'$  and  $\psi'$  of the evolved state  $\mathcal{R}_\theta^{(A)}[\rho_A]$  are respectively given by  $\phi' = \phi + \theta$  and  $\psi' = \psi - \theta$ . As the average over  $\theta$  in the AvQFI can be equivalently performed over  $\theta + \phi$ , only the sum  $\phi + \psi$  can influence the average estimation precision obtained with the probe  $\rho_A$ . Therefore, without loss of generality we can always set  $\phi = 0$  in Equation (3.62) when single-mode probes are being studied.

In proving Equation (3.59), and therefore all other results in this section, we explicitly considered the noiseless unitary encoding of Equation (3.50). Nonetheless, the same reasoning can also be applied for the noisy evolution considered in Section 3.4.2, because photon loss commutes with phase rotations [see Equation (3.31)].

*Proof of Theorem 3.4.* The covariance matrix and the displacement vector of any two-mode probe state can be grouped in the following blocks:

$$\Gamma = \left( \begin{array}{c|c} \Gamma_A & \Gamma_{OFF} \\ \hline \Gamma_{OFF}^\top & \Gamma_B \end{array} \right), \quad \boldsymbol{\xi} = (\boldsymbol{\xi}_A, \boldsymbol{\xi}_B)^\top. \quad (3.63)$$

By applying Williamson decomposition [see Theorem 3.1] to subsystem  $B$ , one can find

a local Gaussian unitary map  $\mathcal{T}^{(B)}$  which changes

$$\Gamma_B \rightarrow T^{-1}\Gamma_B T^{-1\tau} = b\mathbb{1}_2, \quad \text{with } b \geq 1, \quad (3.64)$$

as well as

$$\Gamma_{OFF} \rightarrow \Gamma_{OFF} T^{-1\tau}, \quad \boldsymbol{\xi}_B \rightarrow \mathbf{0}, \quad (3.65)$$

without altering subsystem  $A$ . Notice that  $b$  is not one of the symplectic eigenvalues of the total matrix  $\Gamma$ , but rather the symplectic eigenvalue of the reduced covariance matrix  $\Gamma_B$ . By applying the singular value decomposition to  $\Gamma_{OFF} T^{-1\tau}$ , one can write

$$\Gamma_{OFF} T^{-1\tau} = R_1 \text{Diag}(c, d) R_2^\tau, \quad (3.66)$$

for some  $R_1, R_2 \in SO(2)$  and real (non necessarily positive) parameters  $c, d$ . Now let  $\mathcal{R}_1^{(A)}, \mathcal{R}_1^{(B)}$  be the single-mode rotations associated respectively to  $R_1^{-1}, R_2^{-1}$  via Equation (3.18). The Gaussian unitary that transforms the input probe to an equivalent one with standard form is  $(\mathcal{R}_1^{(A)} \otimes \mathcal{R}_1^{(B)}) \circ \mathcal{T}^{(B)}$ . By construction, this simpler Gaussian state has the same AvQFI of the original state, because of Equation (3.59).  $\square$

## 3.4 Results on squeezing estimation

In this section, we evaluate the AvQFI for squeezing estimation for a broad class of Gaussian input probes, parametrised as in Equation (3.61) or (3.62). In doing this, we will make use of the QFI expressions given in Equations (3.33) and (3.37). The discussion will be divided in two big sections: at first we consider a noiseless evolution, and then we take photon losses into account. In both cases, we start by characterising the average estimation precision associated with single-mode states, and then we compare it with a paradigmatic class of correlated probes: two-mode squeezed vacuum states. In our analysis the energy of the probes (i.e., their average photon number) will be kept fixed. This choice is physically meaningful because any estimation precision could be trivially enhanced by simply sending more photons into the device being investigated. Therefore, unbounded precisions could be obtained if probes with an arbitrary amount of energy were to be simultaneously taken into account.

In a noiseless setup, we find that pure single-mode squeezed states maximise the AvQFI, at the cost of an estimation precision that strongly depends on the squeezing direction. By considering two-mode squeezed probes, it is possible to bring this variance to zero, but the AvQFI value cannot be increased in the same way. The situation changes when the effects of photon losses are considered. On the one hand, when the amount of loss exceeds a certain threshold the optimal single-mode probe becomes a simple coherent state. On the other hand, the presence of quantum correlations in the probing system can

yield larger AvQFI values, in addition to a vanishing variance in the precision attainable for different squeezing directions.

### 3.4.1 Noiseless evolution

In order to evaluate the AvQFI of a two-mode Gaussian state, we first need to find its  $\theta$ -dependent QFI through Equation (3.33). Note that the contribution coming from the derivatives of the symplectic eigenvalues can be ignored, because the encoding evolution  $\Phi_{\mu,\theta}$  defined in Equation (3.50) is a unitary map. For probes in the standard form of Theorem 3.4, the other terms can be computed by writing the matrix  $\tilde{M}$  as a function of the input covariance matrix  $\Gamma$

$$\text{Tr} \left[ \left( \tilde{M}^{-1} \dot{\tilde{M}} \right)^2 \right] = 2 \text{Tr} \left[ \Gamma^{-1} V_\theta \Gamma V_\theta + V_\theta^2 \right], \quad (3.67)$$

$$\text{Tr} \left[ \left( (1 + \tilde{M}^2)^{-1} \dot{\tilde{M}} \right)^2 \right] = - \text{Tr} \left[ (O V_\theta \Gamma + O \Gamma V_\theta)^2 \right], \quad (3.68)$$

where  $V_\theta = R_\theta^\text{T} (S_\mu^{-1} \dot{S}_\mu) R_\theta$  and  $O = (\mathbb{1} - \Omega \Gamma \Omega \Gamma)^{-1} \Omega$ . In this way it is possible to obtain an analytical, although quite involved, expression for the AvQFI, which can be found in Appendix A.1. By setting  $c = d = 0$ , however, we can obtain a much simpler expression independent from  $b$ , which corresponds to the AvQFI of a single-mode probe. If  $\Gamma_A$  represents the input covariance matrix, it can be written as

$$\overline{H}^{(1)}[\rho_A] = \frac{\text{Tr} [\Gamma_A]^2 + 4 \det \Gamma_A}{2(1 + \det \Gamma_A)} + \frac{|\xi|^2 \text{Tr} [\Gamma_A]}{\det \Gamma_A}. \quad (3.69)$$

As anticipated, we first comment the single-mode result, and then discuss the effects of input correlations.

#### Single-mode probes

A first remarkable fact is that Equation (3.69) is independent from the displacement direction represented by the phase  $\psi$  appearing in Equation (3.62), so that only the absolute value  $|\xi|$  is relevant. This is a peculiarity of the noiseless evolution, that will disappear when losses are considered in Section 3.4.2. It is also worthwhile to note that the single-mode AvQFI of Equation (3.69) can be also obtained by taking the average of the single-mode QFI for fixed squeezing direction found in Ref. [106]. We can show this by rewriting their expression for the  $\theta$ -dependent QFI in our notation, for input states

parametrised as in Equation (3.62) with  $\phi = 0$ :

$$H_\theta^{(1)} = \frac{2|\xi|^2}{\nu_A} (\cosh(2\alpha) + \cos[4\theta - 2\psi] \sinh(2\alpha)) \\ + \frac{4\nu_A^2}{\nu_A^2 + 1} \left( \cosh^4 \alpha + \sinh^4 \alpha - \frac{1}{2} \cos[4\theta] \sinh^2(2\alpha) \right). \quad (3.70)$$

An average over  $\theta$  then yields Equation (3.69). Intuitively, this is because an average over all possible squeezing directions can be equivalently represented by an average over all rotations of the input probe in phase space.

As we already mentioned, in looking for the optimal single-mode probe we compare states with fixed average photon number  $n_A$ , whose expression can be found in Equation (3.47). The AvQFI for randomly generated single-mode input Gaussian states is plotted against  $n_A$  in Figure 3.2a. It is possible to see how pure undisplaced squeezed states ( $|\xi| = 0$ ,  $\nu_A = 1$ ) are the optimal choice, while the smallest estimation precision is obtained for undisplaced thermal states ( $|\xi| = 0$ ,  $\alpha = 0$ ). Their AvQFIs can be respectively written as:

$$\overline{H}^{(1)} [\rho_A^{(\text{sq})}] = 4n_A^2 + 4n_A + 2, \quad (3.71)$$

$$\overline{H}^{(1)} [\rho_A^{(\text{th})}] = 4 \frac{(2n_A + 1)^2}{1 + (2n_A + 1)^2}, \quad (3.72)$$

while coherent states ( $\alpha = 0$ ,  $\nu_A = 1$ ) have an intermediate scaling

$$\overline{H}^{(1)} [\rho_A^{(\text{coh})}] = 2(1 + 2n_A). \quad (3.73)$$

Formal proofs of these bounds can be found in Appendix A.2.

In the simple case of single-mode probes, we can also study how sensitive the QFI written in Equation (3.70) is with respect to variations in  $\theta$ . In order to do so, we consider its variance

$$\text{VAR}(H_\theta) = \frac{V_1^2 + V_2^2 - 2V_1V_2 \cos(2\psi)}{2}, \quad (3.74)$$

where

$$V_1 = \frac{2\nu_A^2}{\nu_A^2 + 1} \sinh^2(2\alpha), \quad V_2 = \frac{2|\xi|^2}{\nu_A} \sinh(2\alpha). \quad (3.75)$$

We can see how it is possible to have a vanishing variance when  $\alpha = 0$ , or when the probe displacement is opportunely chosen in order to have  $V_1 = V_2$ , with  $\psi$  an integer multiple of  $\pi$ . From this observation we conclude that the QFIs of  $\rho_A^{(\text{th})}$  and  $\rho_A^{(\text{coh})}$  do not depend on  $\theta$ , because these probes are characterised by  $\alpha = 0$ . However, the situation is rather different for pure undisplaced squeezed states, for which the variance increases with  $\alpha$ , and thus with  $n_A$ . Hence, the single-mode probes with the optimal AvQFI are affected

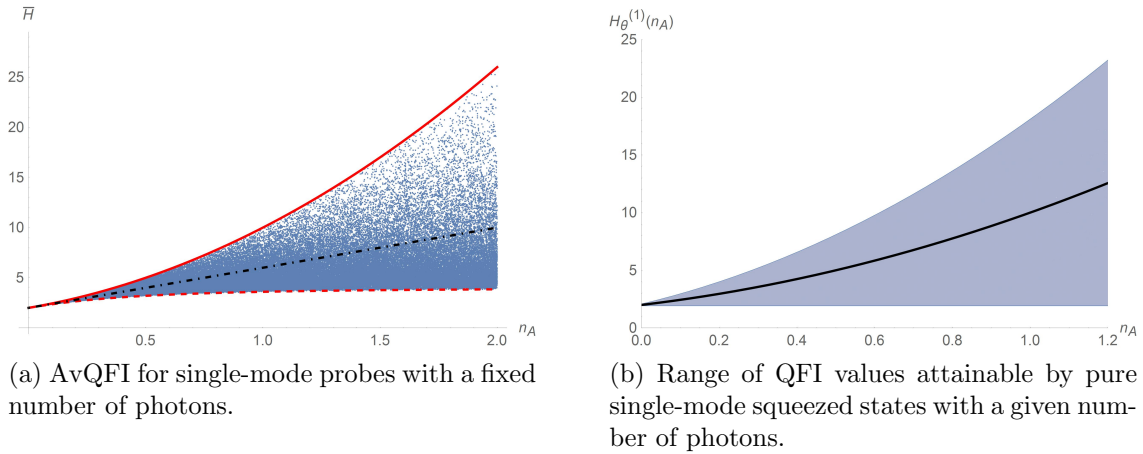


Figure 3.2: (a) Noiseless AvQFI for single-mode Gaussian probes with respect to their photon-number  $n_A$ . Blue dots:  $10^5$  single-mode Gaussian state uniformly sampled with the method of Section 3.2.4; red top solid line: pure squeezed states; red bottom dashed line: thermal states; black dot-dashed line: coherent states. (b) Range of QFI values that could be achieved by pure and undisplaced squeezed single-mode probes when  $\theta$  is varied in  $[0, 2\pi]$ , for any fixed number of photons  $n_A$ . The corresponding AvQFI is plotted as a black line for comparison.

by strong fluctuations in precision when  $\theta$  is varied. In particular, an unlucky choice of  $\theta$  can bring the QFI to its absolute minimum value of 2, as can be seen in Figure 3.2b, so that in this case the squeezed probe is no better than the vacuum.

## Two-mode probes

Instead of simply sending a single-mode probe to the squeezer, another option is to use a two-mode probing system, in which the first mode actively goes through the squeezing device, while the second one is kept as reference. Due to the presence of input correlations, it is possible that the average estimation precision could be improved in this way. In looking for two-mode probes able to improve over the best single-mode input state, an idea would be to correlate the latter with an ancillary mode, without changing its local state. However, this cannot be done, because the optimal single-mode state  $\rho_A^{(\text{sq})}$  is pure, and as such cannot be correlated with any other system. This fact suggests the presence of a trade-off between pure local squeezing and two-mode correlations. This is consistent with similar results obtained for finite-dimensional systems<sup>2</sup> [132], where local purity was identified as an important factor affecting the average estimation ability of a probing system. Instead, as paradigmatic example of correlated probes, we study the two-mode squeezed vacuum states  $\rho_{AB}^{(\text{ts})}$  previously introduced in Equation (3.30). These states are

<sup>2</sup>To be precise, the authors of Ref. [132] use the average skew information as figure of merit [154, 155], rather than the QFI.

already in the standard form of Theorem 3.4, with parameters  $\xi_x = \xi_p = a_{xp} = 0$  and

$$a_x = a_p = b = \cosh(2r), \quad c = -d = \sinh(2r), \quad (3.76)$$

where  $r$  is the squeezing parameter characterising the probe  $\rho_{AB}^{(\text{ts})}$ . By substituting these values in Equation (A.1), we can obtain

$$\overline{H}[\rho_{AB}^{(\text{ts})}] = 4 \sinh(r)^4 + 4 \sinh(r)^2 + 2. \quad (3.77)$$

When comparing two-mode probes with single-mode ones, we can either fix the number of photons in mode  $A$ , labelled as  $n_A$ , or fix the total number of photons  $N$ . We start with the first of these options. With this choice, a preliminary observation is that any two-mode Gaussian state cannot have smaller AvQFI values than its reduced state on mode  $A$ , i.e.,  $\rho_A = \text{Tr}_B[\rho_{AB}]$ . The reason behind this is that the QFI can only decrease under partial trace, so that

$$\overline{H}[\rho_{AB}] \geq \overline{H}^{(1)}[\rho_A] \geq \min_{\rho_A} \overline{H}^{(1)}[\rho_A]. \quad (3.78)$$

More importantly, for the particular example of  $\rho_{AB}^{(\text{ts})}$  the number of photons in mode  $A$  is equal to  $\sinh(r)^2$ , so that  $\rho_{AB}^{(\text{ts})}$  yields the same average performance of the optimal single-mode probe  $\rho_A^{(\text{sq})}$  with the same  $n_A$ :

$$\overline{H}[\rho_{AB}^{(\text{ts})}] = \overline{H}^{(1)}[\rho_A^{(\text{sq})}] = \max_{\rho_A} \overline{H}^{(1)}[\rho_A]. \quad (3.79)$$

This is different from what has been recently found for the aforementioned finite dimensional setting [132], where entanglement between subsystems  $A$  and  $B$  was necessary in order to obtain the maximum average precision.

Moreover, the QFI associated with a two-mode squeezed probe  $\rho_{AB}^{(\text{ts})}$  does not depend on the squeezing direction  $\theta$ , because the terms depending on this phase in Equation (3.67) and Equation (3.68) cancel, due to the symmetric structure of the covariance matrix of  $\rho_{AB}^{(\text{ts})}$ . Therefore, although  $\rho_A^{(\text{sq})}$  and  $\rho_{AB}^{(\text{ts})}$  have the same average precision, the fluctuations in performance can be avoided by using a correlated probe, at the cost of doubling the total number  $N$  of input photons. This is beneficial whenever the possibility of obtaining a predictable precision is preferable to the risk of dealing with a fluctuating performance. At this stage one could wonder whether other two-mode Gaussian states with the same  $n_A$  could yield larger AvQFI values than  $\rho_{AB}^{(\text{ts})}$ . In order to answer this question we numerically sampled up to  $10^6$  of these states, but all of them had smaller AvQFI values. Although we could not find a formal proof, we can still conjecture the following inequality, among

Gaussian states with the same number of photons in the local subsystem  $A$ :

$$\overline{H}[\rho_{AB}] \leq \overline{H}[\rho_{AB}^{(\text{ts})}]. \quad (3.80)$$

Finally, if two-mode squeezed states are compared to the best single-mode probes with the same total number of photons  $N$ , the presence of a reference beam cannot improve the AvQFI. Intuitively, this is because the average estimation precision reached by  $\rho_{AB}^{(\text{ts})}$  can match that of  $\rho_A^{(\text{sq})}$  only at the cost of doubling the total number of photons. In order to explicitly prove this fact, let us write as  $\rho_{AB}|_{n_A, N}$  a two-mode Gaussian state with  $n_A$  photons in mode  $A$  and  $N$  photons in total, and as  $\rho_A|_{n_A}$  a single-mode state with  $n_A$  photons. From Equations (3.80), (3.79), and the monotonicity of  $H[\rho_A^{(\text{sq})}]$  in  $n_A$ , one has:

$$\overline{H}[\rho_{AB}|_{n_A, N}] \leq \overline{H}[\rho_{AB}|_{n_A, 2n_A}] = \overline{H}[\rho_A^{(\text{sq})}|_{n_A}] \leq H[\rho_A^{(\text{sq})}|_N], \quad (3.81)$$

thus proving that in a noiseless regime the presence of correlations cannot improve the AvQFI when the total number of photons is kept constant.

### 3.4.2 Lossy evolution

Instead of considering the ideal and noiseless evolution  $\Phi_{\mu, \theta}$ , which has been defined in Equation (3.50), in this section we move to a more realistic scenario by adding some noise. In particular, we take into account the effects of having non-zero photon losses affecting the propagation of the probes to and from the squeezing device. The overall quantum channel acting on subsystem  $A$  can thus be written as:

$$\tilde{\Phi}_{\mu, \eta, \theta}^{(A)} = \mathcal{L}_{\eta}^{(A)} \circ \Phi_{\mu, \theta}^{(A)} \circ \mathcal{L}_{\eta}^{(A)}, \quad (3.82)$$

where the action of the lossy channel  $\mathcal{L}_{\eta}^{(A)}$  has been detailed in Equation (3.31), and the loss parameter  $\eta$  is considered to be fixed and known. Note that now the physical map that encodes the parameter  $\mu$  on the probing system is not unitary any more, fact which has two main consequences. One the one hand, the QFI will generally depend on the squeezing parameter  $\mu$  being estimated. For this reason, often in the following we will restrict our analysis to the particular example of  $\mu = 1$ , but qualitatively similar results could be obtained for different values of  $\mu$ . On the other hand, the symplectic eigenvalues of the probe can change during the evolution, thus leading to an additional contribution in the QFI [third line of Equation (3.33)]. As the resulting expression for the QFI becomes rather involved, in this section we will often have to recur to numerical analyses.

Once again we start by studying single-mode probes, and we show that, when the amount of noise is larger than a certain  $n_A$ -dependent threshold, coherent states becomes optimal. After that, we compare the performance of the best single-mode probe with the AvQFI of two-mode squeezed states. For fixed  $n_A$ , the latter always yield larger average



estimation precisions. However, if the pair  $(N, \eta)$  lies within a certain region a similar result holds even if the comparison is performed by fixing the *total* number of photons  $N$  of the probe state.

### Optimal single-mode probes

For single-mode probes, a closed expression for the QFI in presence of losses can be explicitly computed. In order to do this, it is convenient to use the formula found in Equation (3.37), because when we apply the channel  $\tilde{\Phi}_{\mu, \eta, \theta}^{(A)}$  to a probe with standard covariance matrix as in Equation (3.61),  $(\mathbb{1} + \tilde{M}_A^2)$  becomes a multiple of the identity. For this reason, Equation (3.37) is much easier to compute than the equivalent expression given in Ref. [149]. Unfortunately, the explicit formula for the  $\theta$ -dependent QFI is too cumbersome to be reported here, but it can be found in Appendix A.3 for any values of  $\eta$  and  $\mu > 0$ . As the average over  $\theta$  in  $[0, 2\pi]$  cannot be evaluated analytically, after some preliminary considerations we will study the AvQFI numerically.

A first noticeable feature is that a dependence on the displacement direction  $\psi$  is generally retained even after performing the average over the squeezing direction  $\theta$ . This is in contrast with the noiseless case, where only the absolute value of the displacement was relevant [see Equation (3.69)]. By fixing without loss of generality  $\phi = 0$  in the state parametrisation of Equation (3.62), we find that the maximum of  $\overline{H}_{\mu, \eta}$  is reached when  $\psi = \pm\pi/2$ , for any fixed values of  $n_A, \mu, \eta$  and  $\alpha$ . In phase space this corresponds to a displacement in the direction of the quadrature with the smallest variance. The formal proof is rather technical and not particularly instructive, but it can be found in Appendix A.4.

After having identified the optimal choice for  $\psi$ , we can continue with the task of identifying the single-mode states with the largest AvQFI, for a given pair  $(n_A, \eta)$ . By exploiting the constraint on the input energy given in Equation (3.47), we can write the parameter  $\alpha$  characterising the probe squeezing as function of the remaining parameters:  $n_A, \nu$  and  $|\xi|^2$ . We have strong numerical evidences that pure states (i.e.,  $\nu = 1$ ) seem to reach the maximum AvQFI value<sup>3</sup> allowed for any fixed value of  $n_A$  and  $\eta$ . For example, in Figure 3.3a we can see a typical plot showing the dependence of AvQFI on  $\nu$  and  $|\xi|^2/(2n_A)$ , for  $\mu = 1, n_A = 5$  and  $\eta = 0.95$ . The ratio  $|\xi|^2/(2n_A)$  is a useful parameter because it quantifies the percentage of photons that should be used to displace the state, rather than to squeeze it.

As a consequence, in searching for the optimal single-mode state we also fix  $\nu = 1$ , and numerically look for the ratio  $|\xi|^2/(2n_A)$  that yields the largest AvQFI. The results of this optimisation can be seen in Figure 3.3b, for  $10^4$  pairs  $(n_A, \eta) \in [0, 10] \times [0, 1]$  in the

<sup>3</sup>We point out that convexity of AvQFI does not allow us to prove the optimality of pure Gaussian probes, because of the constraint on the average photon number. Indeed, convexity would be conclusive if a generic mixed Gaussian state could be decomposed as mixture of single-mode Gaussian states with the same average photon-number, but we are not aware of any such decomposition.

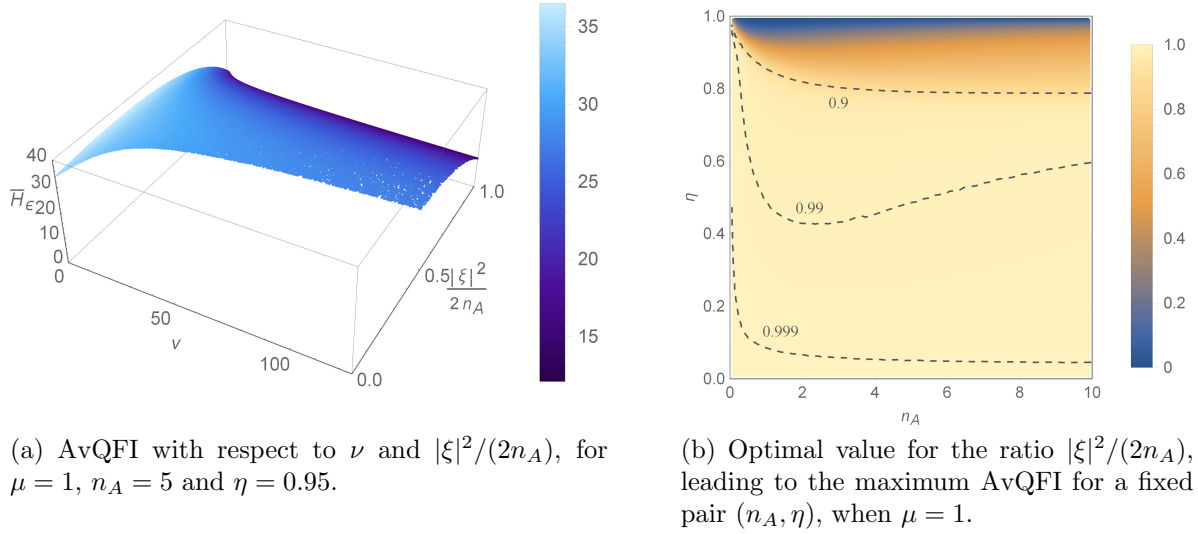


Figure 3.3: Numerical investigations leading to the identifications of the optimal single-mode probe in noisy condition. From analytical considerations, we already know that we can fix  $\phi = 0$  and  $\psi = \pi/2$ . Panel (a) then shows how the largest AvQFI values are reached for pure states (with minimal  $\nu$ ). In this particular example we considered  $\mu = 1$ ,  $n_A = 5$  and  $\eta = 0.95$ , and we uniformly sampled  $10^5$  random Gaussian states with  $\phi = 0$  and  $\psi = \pi/2$ , according to the method detailed in Section 3.2.4. The result of the final optimisation over the displacement ratio  $|\xi|^2/(2n_A)$  is shown in panel (b).

specific case of  $\mu = 1$ . Notice that we recover the fact that for vanishing losses ( $\eta = 1$ ) the best strategy consists in squeezing the input state as much as possible. However, it is also possible to see that for high losses the opposite choice leads to better average performances. In particular, for a wide range of values of  $\eta$ , below a certain  $n_A$ -dependent threshold, the optimal probe can be considered a coherent state for all practical purposes (although strictly speaking a vanishingly small squeezing component is always required). In the intermediate regime, the optimal probe has a non-zero amount of both squeezing and displacement.

As final remark, we should comment on the dependence of this result on the value of  $\mu = 1$ , which was arbitrarily chosen. In Figure 3.4 it is possible to see the plots associated with  $\mu = 0.5$  or  $0.1$ : the qualitative behaviour is similar, but with a much faster transition between the two extreme regimes where the optimal parameter  $|\xi|^2/(2n_A)$  is 0 or 1.

### Comparison with correlated probes

We now compare single-mode results with the paradigmatic example of two-mode squeezed vacuum probes. In particular, for any  $\mu$ ,  $n_A$  (or total photon-number  $N$ ), and  $\eta$ , we can evaluate the relative increase in precision that can be obtained by using this correlated

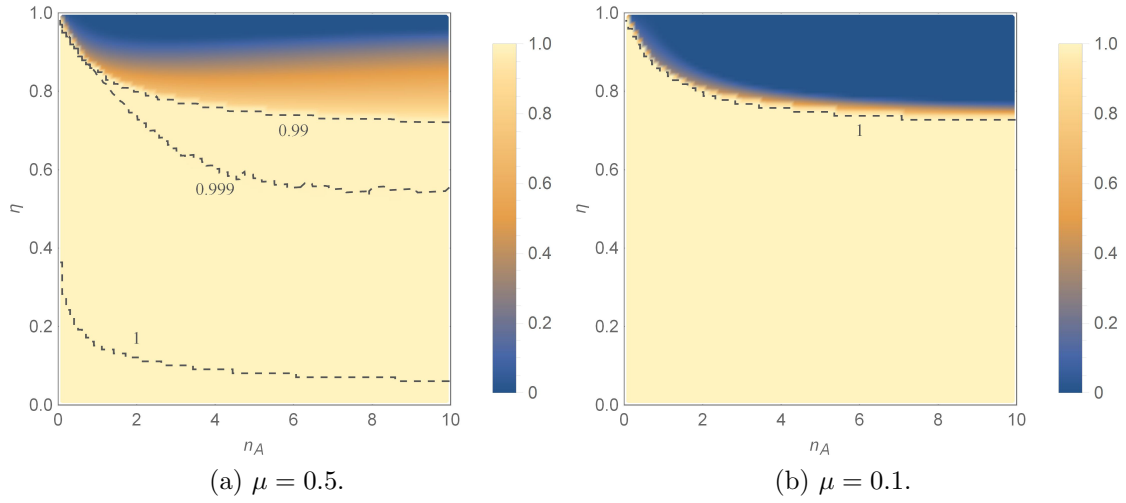


Figure 3.4: Optimal value for the ratio  $|\xi|^2/(2n_A)$ , which leads to the maximum AvQFI, for  $10^4$  pairs  $(n_A, \eta) \in [0, 10] \times [0, 1]$  and  $\mu = 0.5, 0.1$ . Only probes with the optimal parameters  $\phi = 0, \psi = \pi/2$  and  $\nu = 1$  were considered.

input state rather than the optimal single-mode probe, as quantified by

$$\mathcal{I} = \frac{\overline{H}_{\mu,\eta}(\rho_{AB}^{(\text{sq})}) - \max_{\rho_A} \overline{H}_{\mu,\eta}^{(1)}(\rho_A)}{\max_{\rho_A} \overline{H}_{\mu,\eta}^{(1)}(\rho_A)}. \quad (3.83)$$

The results are plotted in Figure 3.5 for different values of  $\mu$ . We can see how two-mode squeezed states always yield a better average precision than all single-mode probes with the same  $n_A$ . Remarkably, in certain conditions the same can be said even if we compare states with the same total number of photons  $N$ , thus keeping into account also the photons in the ancillary mode. Among the considered values of  $\mu$ , this last effect is larger for  $\mu = 1$ , and becomes less significant when  $\mu$  decreases.

Moreover, even in presence of losses the QFI associated with the correlated probe  $\rho_{AB}^{(\text{sq})}$  does not depend on the direction of squeezing  $\theta$  applied by the device under investigation, exactly as in the noiseless case. Hence, the stability of this particular class of entangled states against fluctuations in  $\theta$  is not affected by the introduction of photon losses.

## 3.5 Conclusions

In this chapter we considered a problem of quantum metrology in a Gaussian framework, where the goal was to estimate the squeezing strength of a certain device without knowing the direction of application. By assuming the knowledge of this phase at the measurement stage, we showed how the average quantum Fisher information associated with a probing state represents a significant figure of merit for its versatility.

For the specific example of squeezing estimation, we first studied the problem in ideal

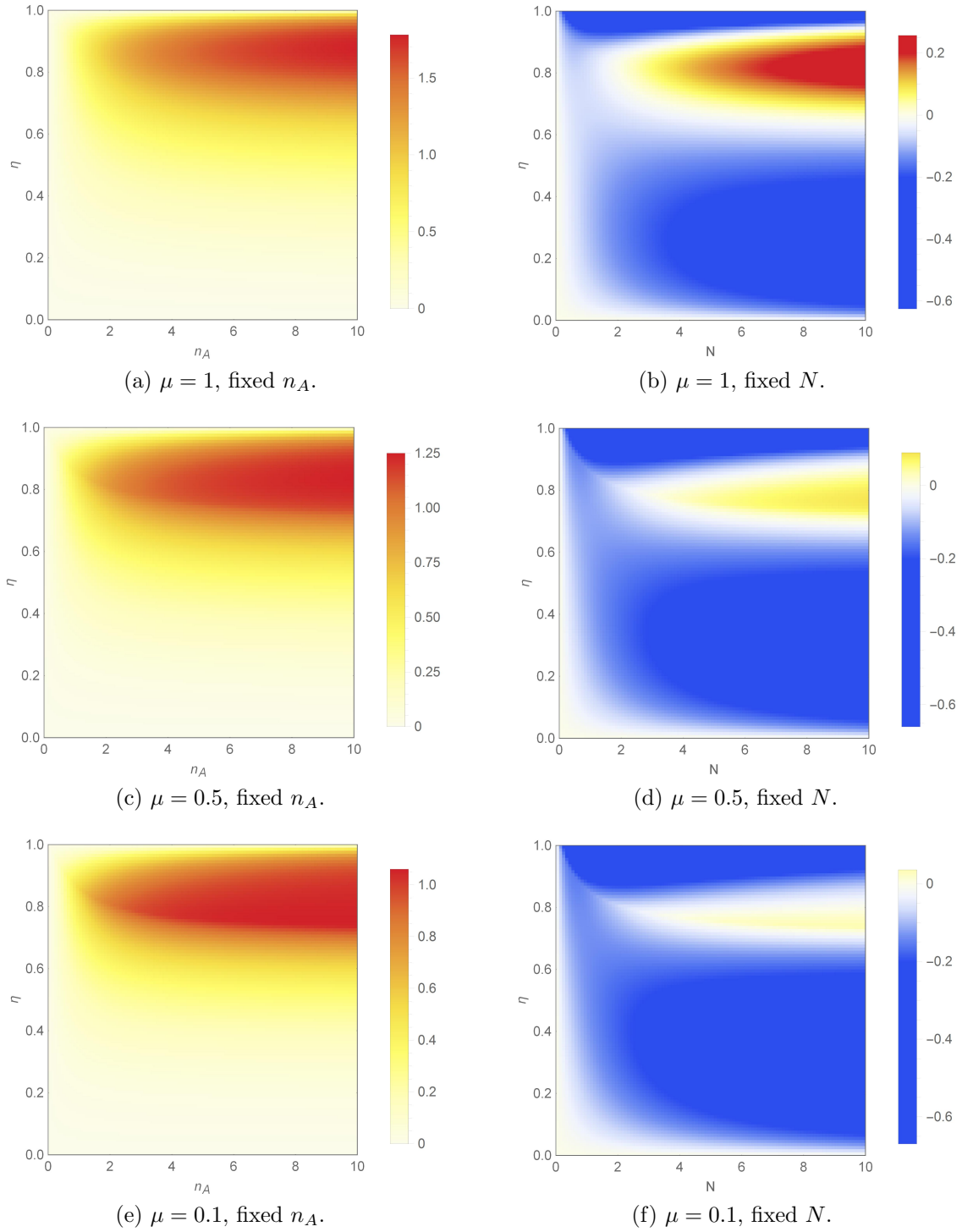


Figure 3.5: Relative increase  $\mathcal{I}$  in precision obtained by using two-mode squeezed vacuum probes rather than the optimal single-mode input state, numerically evaluated for  $10^4$  pairs  $(n_A, \eta)$  or  $(N, \eta)$  when  $\mu = 1, 0.5, 0.1$ .

noiseless conditions, and then we took into account the effects of photon losses. In absence of noise, the optimal single-mode probe is obtained by squeezing the vacuum state as much as possible. However, the variance of the resulting QFI increases with the energy of the state, and in the worst case the attained precision is no better than what could have been done by sending no photon at all. We also showed how the same average result could be obtained in a stable manner by employing a two-mode squeezed state, at the price of doubling the total number of photons composing the probe. When photon losses are taken into account, the situation changes. On the one hand, in the single-mode case using all available energy to squeeze the probe is not necessarily optimal, and introducing some displacement often helps. In particular, above an energy-dependent amount of loss, coherent states maximise the AvQFI. On the other hand, a numerical investigation revealed that in certain conditions two-mode squeezed states can yield an advantage over the best single-mode probe, even if the comparison is performed by fixing the total number of photons. This fact, together with the stability in precision which characterises this class of correlated probes, shows how two-mode squeezed states can be considered good versatile probes for squeezing estimation.

From a more general perspective, the quantum Fisher information is already a broadly applicable tool, which has been employed in many situations, not necessarily within the realm of estimation theory. For example, it has been successfully applied in the study of quantum speed limits, [156, 157], or in quantifying non-classicality [119, 129] or macroscopicity [158, 159]. Our study formalises the intuition behind the fact that its average value can be used as an indicator for the versatility of a probe employed to estimate a certain parameter across different settings. In particular, this can be done whenever the specific details characterising the evolution of each probe become available before the measurement stage. One of the advantages of this formulation is the possibility of taking into account fluctuating interactions, which randomly change across different probes. As this situation naturally arises whenever the probe evolution cannot be fully controlled experimentally, the AvQFI could find applications well beyond the particular example analysed in this chapter.



# Chapter 4

## Versatile communication bounds for quantum networks

### 4.1 Introduction

The possibility of communicating complex concepts has been a key aspect in the development of human society, and any improvement in our ability of effectively exchanging information always significantly altered our lifestyle. This can be repeatedly seen in modern history, with the invention of the telegraph, the first telephones and more recently with the diffusion of mobile phones. Alongside the development of these point-to-point communication devices, we also witnessed the overwhelming growth of the internet, which nowadays allows us to acquire and share information in a way not even imaginable before.

The field of quantum communication aims at assessing how direct exploitation of quantum mechanical laws could affect the communication among two or multiple parties. More specifically, it is mostly concerned with estimating the ultimate limits that quantum mechanics enforces on transmission rates, and with developing schemes with performances as close as possible to those thresholds.

In the remainder of this introduction, we first provide a brief overview of some aspects of quantum communication relevant to this chapter, and then we summarise our findings in context of previous literature. The layout of the rest of the chapter is postponed to Section [4.1.2](#).

#### 4.1.1 Quantum communication

In the most common scenario of point-to-point quantum communication two users, say Alice and Bob, are connected by a quantum channel. Loosely speaking this consists in a transmission line able to preserve, up to a certain amount of noise, the quantum features of the input state. For example, one can think of encoding information on the polarisation of a single photon, which can then be transmitted through an optical fibre

or a free-space link. Having a quantum channel at their disposal, Alice and Bob have now multiple options, depending on the sort of result they want to achieve. In a first case, they might enhance their ability to transmit classical information by encoding the letters composing their message into different quantum states. Alternatively, they might be interested in faithfully transmitting a quantum state itself, or their goal might be to share several maximally entangled qubits, known as ebits [160]. These are powerful resources, useful for many other tasks. In a final scenario, Alice and Bob might desire to share a string of secret and perfectly correlated random bits, also known as “private bits” or pbits [161, 162], which could then be used to encode highly confidential information, for example via the one-time pad encoding. This last situation is commonly known as “quantum key distribution”, or QKD, and it is currently one of the most studied and developed applications of quantum information [163]. The parties might also be allowed to use some additional free resources. Depending on their task, a free resource could be either some pre-shared entanglement (when the purpose is to transmit classical information), or the possibility of freely communicating over a public classical channel in one or both directions (for all other purposes). From this overview, it is clear how a single quantum channel could be characterised by a plethora of figures of merit, depending on the specific conditions in which it is used.

For any communication purpose, the typical approach used to assess the performance of a given quantum channel consists in quantifying the maximum rate at which it can faithfully transmit information in the asymptotic limit of many channel uses. This quantity, known as “capacity” of the channel, is obtained by taking the limit of the ratio between the number of relevant “bits” of information (e.g., classical bits, maximally entangled qubits [160], private bits [161, 162], etc.) shared at the end of the communication and the number of channel uses. The evaluation of this quantity is often far from trivial, because it requires to optimise over any possible choice of encoding and decoding operations performed by the parties, possibly applied globally on the whole set of input/output states [8]. Despite its abstract definition, the capacity of a quantum channel can also be practically useful in that it can be compared with the performance of known transmission schemes. As a result, this yields an indication on the extent of improvements that could be expected in the future. Similar conclusions could also be obtained by studying upper bounds on the channel capacity, which can be easier to compute than the capacity itself. For example, with this approach the authors of Ref. [164] were able to provide strong evidences for the need of quantum repeaters in long-distance QKD. In other words, rather than directly transmit the desired state to the receiver, above a certain distance it is beneficial to introduce multiple intermediate stations, with the purpose of establishing shorter entangled links before attempting to transmit the desired state (further details can be found in Ref. [7]). For the aforementioned practical reason, as well as for the appeal of characterising the ultimate transmission rate achievable by a quantum chan-



nel, recently several authors evaluated computable and simple upper bounds on channel capacities [164–172].

In the future, it is reasonable to expect that multiple parties might have simultaneously access to a network of several quantum channels. In this vision, each user will be located at a node of the network, and will have full quantum control over the systems under his/her influence. This scenario represents the evolution of today’s internet in a quantum regime, and is therefore known as “quantum internet” [2, 7, 173, 174]. Although this level of control is still quite far in the future, the first metropolitan networks of quantum channels are starting to appear, and experimental realisations of quantum key distribution schemes are currently under way [3–6].

### 4.1.2 Our contribution in light of recent results

In this chapter we will focus on two instances of channel capacities, which quantify the optimal rate at which ebits or pbits can be generated across a quantum channel  $\mathcal{N}$  when the two users are allowed to freely exchange classical public information. These are respectively known as two-way quantum capacity<sup>1</sup>  $Q(\mathcal{N})$  and private capacity  $K(\mathcal{N})$ , and satisfy the relation  $Q(\mathcal{N}) \leq K(\mathcal{N})$  because an ebit can be considered a special case of pbit [161, 162]. Indeed, by performing local measurements on a maximally entangled qubit, a pair of perfectly correlated bits could be obtained. In the following we follow a convention typically adopted in the literature, and we use the symbol  $C(\mathcal{N})$  to represent a generic one of these two capacities whenever the same property holds for both of them. Recently, upper bounds on  $K(\mathcal{N})$ , and therefore also on  $Q(\mathcal{N})$ , have been obtained by considering the maximum amount of entanglement that could be shared through a *single* use of the channel  $\mathcal{N}$ . This maximum entanglement is known as “entanglement of the channel” [165–167, 169], and can be written as

$$E(\mathcal{N}) \equiv \max_{\rho_{AA'}} E(\mathcal{N}_{A' \rightarrow B}[\rho_{AA'}]), \quad (4.1)$$

where  $E$  can be any measure quantifying the entanglement across the bipartition  $A : B$ , and the subscript  $A' \rightarrow B$  highlights the fact that  $\mathcal{N}$  takes as input a quantum state on Alice’s side and yields an output on Bob’s one. With this notation, the aforementioned bounds on channel capacities can be compactly written as

$$C(\mathcal{N}) \leq E(\mathcal{N}). \quad (4.2)$$

A result of this form has been proven for *any* quantum channel by considering a

---

<sup>1</sup>Technically, the quantum capacity of a channel quantifies the optimal rate at which quantum information could be transmitted, and it is generally smaller than the rate at which ebits can be distilled. However, because two-way classical communication is allowed in our case, any shared ebit leads to a faithfully transmitted qubit because of the teleportation protocol [175], and the two rates coincide.

particular entanglement measure, known as squashed entanglement  $E_{\text{sq}}$  [164, 165, 176]. However, because of the difficulty of computing  $E_{\text{sq}}(\mathcal{N})$  exactly, one often needs to further upper bound it, thus loosening the bound on the channel capacity. The relative entropy of entanglement  $E_{\text{R}}$  is also known to provide this sort of bound on the capacity, at least for those channels that are “Choi-stretchable” [166, 169]. Remarkably, this last upper bound often has no gap with respect to the best known lower bound on the capacity. When this happens, the capacity is exactly identified and the channel is dubbed as “distillable” [166]. However, a drawback of the approach based on the relative entropy of entanglement is that at the moment it is not known whether Equation (4.2), with  $E = E_{\text{R}}$ , is valid also when applied on a generic, non Choi-stretchable, quantum channel. Another option is to use in Equation (4.2) the max-relative entropy of entanglement  $E_{\text{max}}$  [167]. The resulting bound is formally proven only for quantum channels acting on finite-dimensional systems, but it is thought to hold in general (see Ref. [167] for a short discussion). Needless to say, if several entanglement measures yield an upper bound on  $C(\mathcal{N})$  via Equation (4.2), the optimal choice consists in retaining the tightest, i.e. smallest, among them. The set of pairs  $(E, \mathcal{N})$  for which Equation (4.2) is known to hold is object of ongoing research, and its extension represents an interesting and challenging problem.

When a quantum network is considered, one can seek for upper bounds on the number of ebits or pbits that can be shared by two parties after a communication protocol, possibly involving all the nodes of the network. The authors of Refs. [177] and [173] considered this situation, and obtained network versions of Equation (4.2) by using respectively  $E_{\text{R}}$  or  $E_{\text{sq}}$  as entanglement measures. In their approach, they split the whole network into two smaller components, and then used the maximum amount of entanglement generated by the channels connecting the two parts in order to bound the number of ebits (pbits) produced by a communication protocol. The result of Ref. [173], which is based on  $E_{\text{sq}}$ , holds for arbitrary quantum networks, because it inherits the broad applicability of the single-channel bound given in Equation (4.2) for  $E = E_{\text{sq}}$ . However, this single-channel bound is not typically tight, and therefore a non-vanishing gap can also be expected between the result of Ref. [173] and the optimal number of ebits or pbits that could be generated by the network. It is thus natural to wonder: (i) whether different entanglement measures could improve this sort of network bound, and (ii) to what extent the choice of entanglement measure could be tailored to the characteristics of the channels in the network.

In this chapter, we start by noticing how a common strategy has been adopted in most of the known proofs for the bounds with the same form as Equation (4.2). This allows us to formally identify two sufficient properties that, if satisfied by a given pair  $(E, \mathcal{N})$ , lead to a new instance of Equation (4.2). Then, we show how the same two properties are also sufficient to generalise the result of Ref. [173] on quantum networks to different entanglement measures:  $E_{\text{max}}$  or, when the channels in the network are Choi-stretchable,

$E_R$ . The possibility of using  $E_R$  is particularly interesting, because Equation (4.2) is known to be rather tight when stated in terms of  $E_R$ , rather than  $E_{sq}$ . Therefore, the same advantage is expected to arise for the corresponding upper bound on the performance of a quantum network. However, at this stage the presence of a *single* non-stretchable channel would prevent us from bounding the number of ebits (pbits) produced by the network by means of  $E_R$ , even though the relative entropy of entanglement would provide extremely tight bounds for all the remaining channels of the network. From this example, it should be clear how the possibility of changing entanglement measure on a channel-by-channel basis could greatly benefit the resulting bound on the performance of the network. Our second and most important result goes exactly in this direction.

By exploiting an intermediate step in the discussion of Ref. [167], we are able to obtain an upper bound on the performance of a quantum network where the maximum entanglement produced by a channel can be quantified by either  $E_R$  or  $E_{max}$ . In particular, we can use the relative entropy of entanglement on the Choi-stretchable part of the network, and the max-relative entropy of entanglement on the rest, where we could not use  $E_R$ . As  $E_R(\mathcal{N}) \leq E_{max}(\mathcal{N})$  for any quantum channel  $\mathcal{N}$ , this allows us to use the relative-entropy of entanglement as much as possible. As a consequence, we are able to maintain the precision guaranteed by the relative entropy of entanglement without the need to restrict the applicability of our method only to networks composed by Choi-stretchable channels. After having presented this general result, we will provide explicit examples of networks where our bound yields an advantage over its counterpart based on the squashed entanglement. In order to do this, we will also evaluate the max-relative entropy of entanglement for the most common qubit channels, by exploiting their symmetry under phase rotations.

The remainder of this chapter is organised as follows. In Section 4.2 we introduce our notation and some preliminary notions. In Section 4.3 we formally identify sufficient properties that, if satisfied by a pair  $(E, \mathcal{N})$ , lead to an upper bound on the capacity of a single-channel as in Equation (4.2). Furthermore, along the lines of Ref. [173], we show how the same properties are also sufficient to obtain an upper bound on the number of ebits (or pbits) generated through a quantum network. Our main result is presented in Section 4.4, where we derive a similar versatile upper bound, in which different entanglement measures are applied to the channels of the network depending on their Choi-stretchability. Analytical or numerical evaluations of the max-relative entropy of entanglement for the most common qubit channels can be found in Section 4.5, while two examples of networks where our bound performs better than the bound based on the squashed entanglement are presented in Section 4.6. A final discussion on our results can be found in Section 4.7, together with our conclusions.

## 4.2 Preliminary notions

In order to discuss the results of this chapter, we will need to extensively use the definitions and properties of a few entanglement measures, as well as the notions of private states and Choi-stretchable channels. In this section we introduce all these concepts, together with the formalism that will be used to describe a generic adaptive communication protocol over a quantum network. At the end of the section, we will also discuss the figure of merit that we will use in order to quantify the performance of a communication strategy, and we will comment on its relation to the usual single-channel capacity.

### 4.2.1 Entanglement measures

Given two quantum states  $\rho$  and  $\sigma$ , with supports satisfying  $\text{Supp}(\rho) \subseteq \text{Supp}(\sigma)$ , their relative entropy [178] and max-relative entropy [179] are respectively defined as

$$S(\rho||\sigma) = \text{Tr} [\rho(\log \rho - \log \sigma)], \quad (4.3)$$

$$D_{\max}(\rho||\sigma) = \inf\{x | 2^x \sigma - \rho \geq 0\}, \quad (4.4)$$

while their value is set to  $\infty$  if the condition on the supports is not satisfied. The relative and max-relative entropy of two states are related by

$$S(\rho||\sigma) \leq D_{\max}(\rho||\sigma), \quad (4.5)$$

are invariant under joint unitary operations, that is:

$$S(\hat{U}\rho\hat{U}^\dagger||\hat{U}\sigma\hat{U}^\dagger) = S(\rho||\sigma), \quad D_{\max}(\hat{U}\rho\hat{U}^\dagger||\hat{U}\sigma\hat{U}^\dagger) = D_{\max}(\rho||\sigma), \quad (4.6)$$

for any unitary operator  $\hat{U}$ , and they are also non-negative and equal to zero if and only if  $\rho = \sigma$ . For this reason, they can be intuitively considered as sort of “distances” between  $\rho$  and  $\sigma$  (however, one should keep in mind that the triangular inequality does not necessarily hold). Moreover, the relative entropy is jointly convex in its arguments, whereas the max-relative entropy is jointly quasi-convex. These properties can be formally written as

$$S\left(\sum_i p_i \rho_i \middle| \middle| \sum_i p_i \sigma_i\right) \leq \sum_i p_i S(\rho_i||\sigma_i), \quad (4.7)$$

$$D_{\max}\left(\sum_i p_i \rho_i \middle| \middle| \sum_i p_i \sigma_i\right) \leq \max_i D_{\max}(\rho_i||\sigma_i), \quad (4.8)$$

where  $\{\rho_i\}_i$  and  $\{\sigma_i\}_i$  are quantum states, and  $p_i \geq 0$  with  $\sum_i p_i = 1$ . The relative and max-relative entropies can be used to define entanglement measures respectively known

as relative entropy of entanglement [180] and max-relative entropy of entanglement [179]. For a given bipartite state  $\rho_{AB}$ , their values are obtained by optimising over all separable states:

$$E_{\text{R}}^{A:B}(\rho_{AB}) = \min_{\sigma_{AB} \in \text{SEP}} S(\rho_{AB} || \sigma_{AB}), \quad (4.9)$$

$$E_{\text{max}}^{A:B}(\rho_{AB}) = \min_{\sigma_{AB} \in \text{SEP}} D_{\text{max}}(\rho_{AB} || \sigma_{AB}). \quad (4.10)$$

The ordering relation in Equation (4.5) can also be straightforwardly extended to the entanglement measures  $E_{\text{R}}$  and  $E_{\text{max}}$ , as well as to the entanglement of a channel  $\mathcal{N}$  [see Equation (4.1)]:

$$E_{\text{R}}(\rho_{AB}) \leq E_{\text{max}}(\rho_{AB}), \quad E_{\text{R}}(\mathcal{N}) \leq E_{\text{max}}(\mathcal{N}). \quad (4.11)$$

Further details on  $E_{\text{max}}$  can be found in Ref. [181].

Another entanglement measure often mentioned in this chapter is the squashed entanglement  $E_{\text{sq}}$ . For a bipartite quantum state  $\rho_{AB}$ , this is defined by minimising the quantum conditional information  $I(A : B|C)$  over all its possible extensions  $\rho_{ABC}$ :

$$E_{\text{sq}}(\rho_{AB}) = \inf \left\{ \frac{1}{2} I(A : B|C) : \rho_{ABC} \text{ such that } \rho_{AB} = \text{Tr}_C [\rho_{ABC}] \right\}, \quad (4.12)$$

where

$$I(A : B|C) = S(\rho_{AC}) + S(\rho_{BC}) - S(\rho_{ABC}) - S(\rho_C), \quad (4.13)$$

and  $S(\rho) = -\text{Tr} [\rho \log \rho]$  is the von-Neumann entropy. Further details can be found in Refs. [164, 165, 176].

In the following, we will often omit to explicitly write the bipartition  $A : B$  in the symbols  $E_{\text{R}}$ ,  $E_{\text{max}}$  or  $E_{\text{sq}}$ , unless needed to avoid confusion. If the local quantum systems of Alice and Bob are divided into smaller subsystems, these will be labelled, for example, as  $A, A', A''$  and  $B, B', B''$ . When this is the case, the default evaluation of an entanglement measure has to be considered across the bipartition  $AA'A'' : BB'B''$ . As we have already mentioned, we consider a communication scenario in which the parties can locally alter the quantum states in their possession, and are allowed to exchange classical information to coordinate their strategy. This scenario is commonly summarised by saying that they are allowed to perform LOCC operations, which stands for “Local Operations and Classical Communication”. Any round of LOCC can be decomposed into several steps, in which Alice and Bob iteratively switch roles. At the beginning one of them, say Alice, performs a POVM measurement and communicates the result to Bob, who can rotate his subsystem according to the information received. Then Bob does the same and communicates his result to Alice, and so on. Importantly, any well-defined entanglement measure cannot, on average, increase under this sort of operations. This property can be explicitly written

for an entanglement measure  $E$  as

$$\sum_k p_k E\left(\rho_{AB}^{(k)}\right) \leq E(\rho_{AB}), \quad (4.14)$$

where  $\rho_{AB}$  is the input state,  $k$  represents a multi-index containing all the measurement outcomes obtained by Alice and Bob during the multiple measurements performed on the considered LOCC round,  $p_k$  is the probability of obtaining the outcomes in  $k$ , and  $\rho_{AB}^{(k)}$  is the output state of the system post-selected on such result. We stress that  $E_R$ ,  $E_{\max}$  and  $E_{\text{sq}}$ , as well as any other generic entanglement measure  $E$  considered in this chapter, satisfy Equation (4.14) and become zero when evaluated on separable states.

### 4.2.2 Target states: maximally entangled or private states

One of the most common goals of two parties in a quantum communication protocol is to share one or multiple copies of a  $d$ -dimensional maximally entangled state

$$\psi_{AB}(d) = \sum_{i,j=1}^d \frac{1}{d} |ii\rangle_{AB} \langle jj|, \quad (4.15)$$

where  $\{|i\rangle_{A(B)}\}_i$  is a local orthonormal basis. Any single copy of these states gives Alice and Bob access to  $\log_2 d$  ebits, which they can use to perform one of many possible tasks. For example, they can achieve perfect transmission of any  $d$ -dimensional state via the teleportation protocol [175], or they can perform a projective measurement on the computational basis in order to share a string of  $\log_2 d$  bits of private randomness. The maximally entangled state, however, is not the only quantum state from which a private key can be obtained by performing local measurements. The generation of a private string of correlated bits is the other main task being considered in this chapter, so it is worthwhile to expand a bit on this concept.

Any cryptographic setting, where Alice and Bob want to share some information without anyone else knowing about it, is naturally described in terms of three parties: Alice, Bob, and an eavesdropper, who for simplicity we can call Charlie<sup>2</sup>. Whenever Alice and Bob exchange information, whether over the public classical transmission line or by means of the quantum channel connecting them, Charlie can always be spying on them. With the typical caution of cryptography, we shall assume that the eavesdropper is all powerful: he can freely listen to the conversations taking place over the classical transmission line, has a fully functional quantum computer and possibly a perfect quantum memory, and has access to the environment of the quantum channel connecting Alice and Bob. In order to clarify this last point, we remind the reader that any quantum channel  $\mathcal{N}_{A \rightarrow B}$  can be

---

<sup>2</sup>We cannot follow the standard convention of calling the eavesdropper Eve, because the capital letter  $E$  has already been used to describe a generic entanglement measure.

written as an isometry from the Hilbert space of  $A$  to a larger system  $BC$ , followed by a partial trace over the environmental degrees of freedom  $C$  [182]:

$$\mathcal{N}_{A \rightarrow B}(\rho_A) = \text{Tr}_C \left[ \hat{U}_{A \rightarrow BC} \rho_A \hat{U}_{A \rightarrow BC}^\dagger \right], \quad (4.16)$$

where  $\hat{U}_{A \rightarrow BC}^\dagger \hat{U}_{A \rightarrow BC} = \hat{\mathbb{1}}_A$  and  $\hat{U}_{A \rightarrow BC} \hat{U}_{A \rightarrow BC}^\dagger$  projects on a subspace of  $BC$  isomorphic to  $A$ . The operator  $\hat{U}_{A \rightarrow BC}$  is not unambiguously characterised by the channel  $\mathcal{N}_{A \rightarrow B}$ , but it is defined up to a unitary rotation of system  $C$ . By saying that Charlie has access to the environment of the channel, we mean that he effectively receives the state  $\rho_C = \text{Tr}_B \left[ \hat{U}_{A \rightarrow BC} \rho_A \hat{U}_{A \rightarrow BC}^\dagger \right]$  when Alice sends  $\rho_A$  to Bob.

In this scenario, Alice and Bob can obtain a secret  $d$ -dimensional key if the communication protocol yields a final tripartite quantum state  $\gamma_{ABC}$  such that

$$\hat{\Pi}_A \otimes \hat{\Pi}_B(\gamma_{ABC}) = \frac{1}{d} \sum_{i=0}^{d-1} |i\rangle_A \langle i| \otimes |i\rangle_B \langle i| \otimes \sigma_C, \quad (4.17)$$

where  $\hat{\Pi}_A(\cdot) = \sum_i |i\rangle_A \langle i| (\cdot) |i\rangle_A \langle i|$  is a projective measurement channel in the basis  $\{|i\rangle_i\}$ , and similarly for  $\hat{\Pi}_B$ , while  $\sigma_C$  is a generic quantum state. Indeed, from Equation (4.17) we can see how Alice and Bob obtain  $\log_2 d$  bits of shared randomness, completely uncorrelated from the state which is left to Charlie. Although intuitive, this formulation has the drawback of involving tripartite states, and of explicitly taking projective measurements into account. A remarkable change in perspective came across with Refs. [161, 162], where the authors showed that a tripartite state  $\gamma_{ABC}$  leads to Equation (4.17) if and only if any one of its purifications  $\gamma_{ABA'B'C}$  is such that, when system  $C$  is traced away, one is left with

$$\gamma_{ABA'B'}(d) = U_{ABA'B'}^{(\text{twist})} (\psi_{AB}(d) \otimes \sigma_{A'B'}) U_{ABA'B'}^{(\text{twist})\dagger}, \quad (4.18)$$

where  $\psi_{AB}(d)$  is a maximally entangled state,  $\sigma_{A'B'}$  can be arbitrary, and the controlled unitary

$$\hat{U}_{ABA'B'}^{(\text{twist})} = \sum_{ij=1}^d |i\rangle_A \langle i| \otimes |j\rangle_B \langle j| \otimes \hat{U}_{A'B'}^{(ij)} \quad (4.19)$$

is known as “twisting unitary”, with each  $\hat{U}_{A'B'}^{(ij)}$  a unitary operator. In Equation (4.18) the systems  $A'$  and  $B'$  are considered to be respectively under the control of Alice and Bob, and are known as “shield systems”. Their role is to prevent an eavesdropper from getting access to the key component, and could have any dimension. Instead, subsystems  $A$  and  $B$  are known as “key systems”. A state of the form of Equation (4.18) is called “private state”, and is extremely useful because it allows us to transform a tripartite problem into a bipartite one. Note that a maximally entangled state is a particular example of private state, with a twisting unitary that is a multiple of the identity (the presence of  $\sigma_{A'B'}$  then becomes irrelevant). This shows that if two parties can distil entanglement they can also



share private random bits, but the converse is not true. A manifestation of this fact can be identified in the discovery of quantum channels with vanishing quantum capacity, but non-zero private capacity [183, 183].

### 4.2.3 Choi-stretchable channels

The idea of using quantum teleportation in order to simplify the structure of a computation or communication task has been used several times in the past [184–186]. More recently, this idea reacquired popularity after Refs. [166, 177], where it was used to define as “Choi-stretchable” any quantum channel  $\mathcal{N}$  such that its action on a quantum state  $\tilde{\rho}_{A'}$  can be written as

$$\mathcal{N}_{A' \rightarrow B'}(\tilde{\rho}_{A'}) = \Lambda_{A'A'' : B'}(\tilde{\rho}_{A'} \otimes \pi_{A''B'}(\mathcal{N})). \quad (4.20)$$

Here  $\Lambda_{A'A'' : B'}$  is a trace-preserving sequence of LOCC operations, and  $\pi_{A''B'}(\mathcal{N})$  represents the Choi-Jamiołkowski state associated with the quantum channel  $\mathcal{N}$ :

$$\pi_{A''B'}(\mathcal{N}) = \mathcal{N}_{\tilde{A} \rightarrow B'}(\psi_{A''\tilde{A}}), \quad (4.21)$$

where  $\psi_{A''\tilde{A}}$  is a maximally entangled state. This state contains all the information on the channel  $\mathcal{N}$ , and it is often used as a tool to characterise the properties of the channel itself. When Equation (4.20) holds, it is possible to equivalently describe the effect of a quantum channel as the application of some LOCC operations to the input state and a pre-shared Choi state [see Figure 4.1]. In this way, the description of a quantum communication protocol over the channel  $\mathcal{N}$  becomes much simpler, because  $\Lambda_{A'A'' : B'}$  can be joined to the other LOCCs performed by the parties. In the following, if a channel  $\mathcal{N}$  is Choi-stretchable we will write  $\mathcal{N} \in \mathcal{S}$ .

From the perspective of this chapter, the interest of Choi-stretchable channels lies in the fact that their relative entropy of entanglement, as defined in Equation (4.1), is an upper bound on their quantum and private capacities assisted by two-way classical communication [166, 169]. Moreover,  $E_R(\mathcal{N})$  exactly coincides with both capacities  $Q(\mathcal{N})$  and  $K(\mathcal{N})$  on particular examples of Choi-stretchable channels, known as “distillable” [166]. Among them, we can enumerate some well-known channels, as the erasure and dephasing channels in finite-dimensional systems, or the lossy channel in bosonic continuous variable systems. Interestingly, even for non-distillable but Choi-stretchable channels, often  $E_R(\mathcal{N})$  is closer to  $C(\mathcal{N})$  [166] than other known upper bounds based on the squashed entanglement of the channel [164, 165].

From Equation (4.20) we can now explicitly derive a property that the relative entropy of entanglement must satisfy when applied on the output of a Choi-stretchable channel. Although it is obvious from the discussion in Ref. [166], we will go through its proof



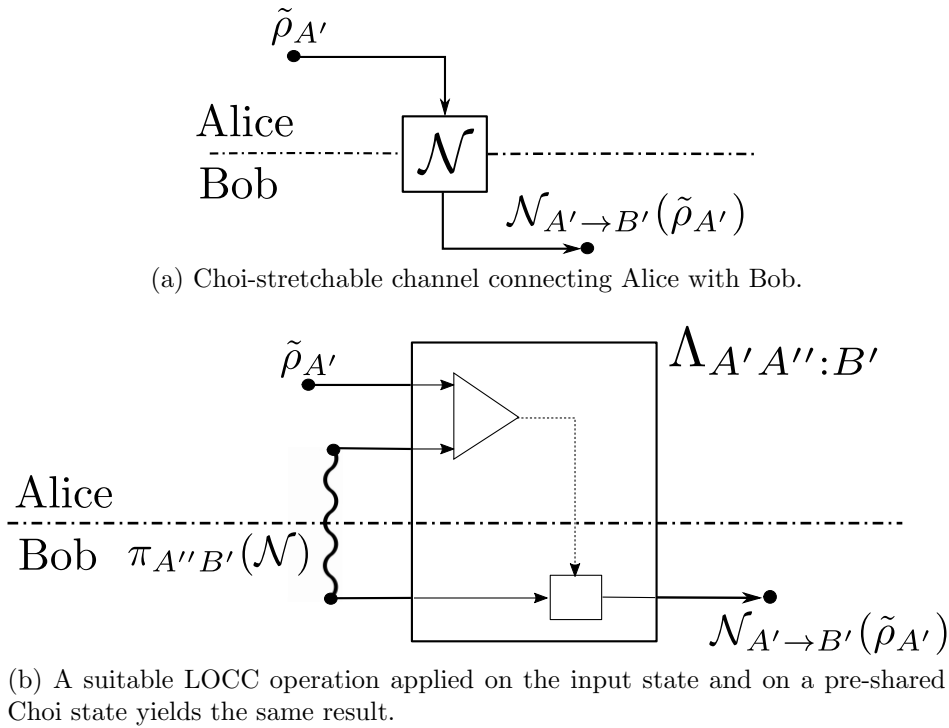


Figure 4.1: (a): Pictorial representation of a Choi-stretchable channel connecting Alice with Bob. (b): simulation of the same channel, obtained by applying LOCC operations on the input state and on a pre-shared Choi state. In this particular example, only a single measurement in  $A$  (triangular gate in the figure) and an outcome-dependent rotation in  $B$  (squared gate connected to the previous measurement by a dotted line, which corresponds to classical communication) are schematically shown inside the box representing the LOCC, but in general there could be more.

in detail because it will play a central role in the results presented in this chapter. In particular, when  $\mathcal{N} \in \mathcal{S}$  and  $\tilde{\rho}_{AB'B} = \mathcal{N}_{A' \rightarrow B'}(\rho_{AA'B})$  the following chain of inequalities holds:

$$\begin{aligned} E_R(\tilde{\rho}_{AB'B}) &\leq E_R(\rho_{AA'B} \otimes \pi_{A'B'}(\mathcal{N})) \\ &\leq E_R(\pi_{A'B'}(\mathcal{N})) + E_R(\rho_{AA'B}) \\ &= E_R(\mathcal{N}) + E_R(\rho_{AA'B}). \end{aligned} \quad (4.22)$$

The first inequality comes from Equation (4.20) and from the monotonicity of  $E_R$  under LOCC, while the second one follows from its sub-additivity under tensor products. The final equality can be proven by showing inequalities in both directions. The inequality “ $\leq$ ” is easily obtained by noticing that  $\pi_{A'B'}$  is a particular example of input state, whereas  $E_R(\mathcal{N})$  involves a maximisation over all possible inputs. The converse direction, instead, is a consequence of Equation (4.20) and of the monotonicity of  $E_R$  under LOCC [166]:

$$E_R(\mathcal{N}_{A' \rightarrow B'}[\rho_{AA'}]) = E_R(\Lambda_{A'A'' : B}[\rho_{AA'} \otimes \pi_{A''B'}(\mathcal{N})]) \leq E_R(\pi_{A''B}(\mathcal{N})), \quad (4.23)$$

which holds for any  $\rho_{AA'}$  and thus also for  $E_R(\mathcal{N})$ . The physical interpretation of Equation (4.22) is that the amount of entanglement found in output of a Choi-stretchable channel, as measured by  $E_R$ , can be upper bounded by the entanglement of the input state plus the maximum amount that could be possibly created by the channel itself. Perhaps surprisingly, to the best of the author’s knowledge, up to date whether the same conclusion could be obtained for any channel remains an open question.

#### 4.2.4 Adaptive communication protocols

Although the simplest setting that allows Alice and Bob to exchange quantum information consists in a single quantum channel connecting them, schematically shown in Figure 4.2a, we can more generally study a situation where they are two local users having access to a quantum network, as in Figure 4.2b. The structure of a network can be formally described by a directed graph  $\mathcal{G} = (V, L)$ , where  $V = \{V_0, \dots, V_{M+1}\}$  is the set of nodes and  $L$  is the set of directed edges, or links, between the nodes. For any edge  $l = (V_i, V_j) \in L$ , there is a quantum channel  $\mathcal{N}^{(l)}$  connecting node  $V_i$  to node  $V_j$ . Without loss of generality, we can assume that nodes  $A = V_0$  and  $B = V_{M+1}$  are respectively controlled by Alice and Bob, whereas the remaining nodes  $\{C_i\}_{i=1}^M$ , with  $C_i = V_i$ , are under the influence of other parties, that for our purposes are considered to be collaborative. Moreover, full quantum control over the local systems is provided to each node of the network. An important notion that will be useful in the following is that of “bipartition” of a quantum network, defined by dividing the nodes  $\{C_i\}_i$  into two disjoint sets:  $\mathcal{C}_A \subset \{C_i\}_i$  and  $\mathcal{C}_B = \{C_i\}_i \setminus \mathcal{C}_A$ . Once a bipartition has been chosen, the set of edges connecting the nodes in  $\{A\} \cup \mathcal{C}_A$

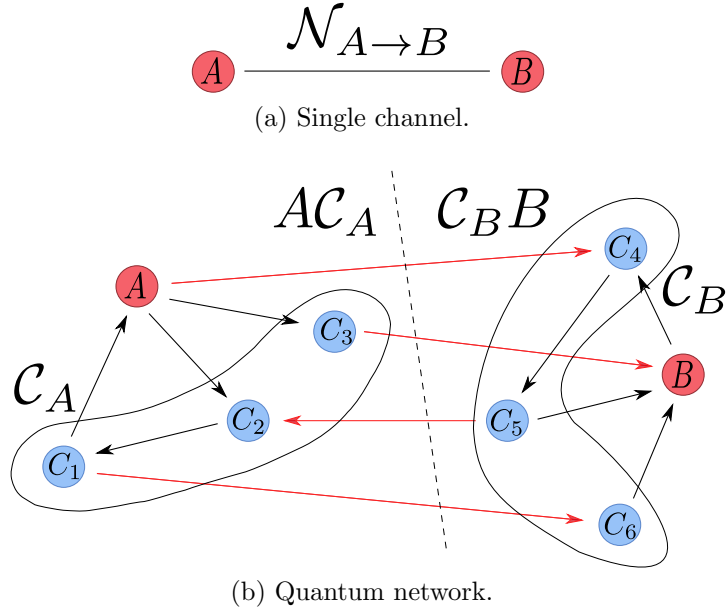


Figure 4.2: (a) Single-channel communication scenario, where  $A$  and  $B$  are connected through the channel  $\mathcal{N}_{A \rightarrow B}$ . (b) An example of quantum network, with  $M = 6$  additional nodes. Every arrow corresponds to a quantum channel. The bipartition  $\mathcal{C}_A = \{C_1, C_2, C_3\}$ ,  $\mathcal{C}_B = \{C_4, C_5, C_6\}$  is shown as an example, and the channels  $\mathcal{N}^{(l)}$  with  $l \in L_{\mathcal{C}_A}$ , which connect the two partitions, are coloured in red.

with those in  $\mathcal{C}_B \cup \{B\}$ , or vice-versa, will be labelled as  $L_{\mathcal{C}_A} \subset L$ . Moreover, in order to keep our notation simple, in the remainder of this chapter we will refer to the subsets of nodes  $\{A\} \cup \mathcal{C}_A$  and  $\mathcal{C}_B \cup \{B\}$  by writing respectively  $AC_A$  and  $C_B B$ .

We are now going to detail the steps of a generic communication protocol [173, 177, 187] that the parties can adopt when they are allowed to freely exchange public classical information in order coordinate their actions. At the beginning they initialise their local systems, so that they share a fully separable state  $\rho^{(1)}$  spread over the whole network. We should have written this as  $\rho_{ABC_1 \dots C_M}^{(1)}$ , but for the sake of simplicity from now on we drop the subscript whenever the state has components everywhere. Then, the parties can iteratively exchange (part of) their systems via the quantum channels, and perform LOCCs after any channel use. As these LOCC may involve measurements, every choice that the parties make at a certain stage of the protocol may depend on all previously obtained measurement outcomes. Note that, although several LOCC operations may be performed between any two channel uses, by grouping them together we can always consider them as a single round of LOCCs, yielding an overall outcome  $k$ . In this way, without loss of generality a single round of the protocol will be composed by the application of a channel followed by a round of LOCCs. Let us group within the vector  $\mathbf{k}_i = (k_0, k_1, \dots, k_{i-1}, k_i)$  the sequence of LOCC outcomes obtained in the first  $i$  rounds of the protocol, where  $k_0 \equiv 1$  is added for convenience. This choice allows us to write as  $\rho^{\mathbf{k}_{i-1}}$  the quantum state that the  $i$ -th round of the protocol receives as input. This is converted to  $\rho^{\mathbf{k}_i}$  by the

application of the following two steps, which compose the  $i$ th round of the protocol:

- Depending on previous LOCC outcomes, i.e. on  $\mathbf{k}_{i-1}$ , the parties may use the channel  $\mathcal{N}^{(l_{\mathbf{k}_{i-1}})}$  to transmit a quantum state along the edge  $l_{\mathbf{k}_{i-1}} \in L$  of the graph  $\mathcal{G}$  characterising the network. We label the global state at the end of this step as  $\tilde{\rho}^{\mathbf{k}_{i-1}}$ ;
- A round of LOCCs  $\Lambda^{(k_i)}$  is performed on  $\tilde{\rho}^{\mathbf{k}_{i-1}}$ , with output  $k_i$  obtained with probability  $p(k_i|\mathbf{k}_{i-1})$ . The output quantum state  $\rho^{\mathbf{k}_i}$  will be used as input for the following round of the protocol.

When the protocol stops, say after  $n$  rounds, the final state  $\rho_{AB}^{\mathbf{k}_n} = \text{Tr}_{C_1, \dots, C_M} [\rho^{\mathbf{k}_n}]$  shared by Alice and Bob has to be  $\epsilon$ -close in trace norm to an ideal target state  $\phi_{AB}(d_{\mathbf{k}_n})$ , i.e., such that for any sequence of outcomes  $\mathbf{k}_n$  one has

$$\|\rho_{AB}^{\mathbf{k}_n} - \phi_{AB}(d_{\mathbf{k}_n})\|_1 \leq \epsilon, \quad (4.24)$$

where  $\|O\|_1 \equiv \text{Tr} [\sqrt{O^\dagger O}]$ . The target state  $\phi_{AB}(d_{\mathbf{k}_n})$  can either be a maximally entangled state  $\psi_{AB}(d_{\mathbf{k}_n})$  or a private state  $\gamma_{AB}(d_{\mathbf{k}_n})$ , depending on the task of Alice and Bob. Differently from the definition of private state in Equation (4.18), the systems  $A$  and  $B$  appearing in  $\gamma_{AB}$  include both key and shield parts. To avoid confusion, we stress that  $d_{\mathbf{k}_n}$  corresponds to the dimension of a single key system.

All the details of the adaptive strategy leading to Equation (4.24) are determined by the specific protocol  $\mathcal{P}_{\epsilon, n}$  followed by the parties. Among these details we find the threshold  $\epsilon$ , the maximum number of rounds  $n$ , and the target states  $\phi_{AB}(d_{\mathbf{k}_n})$ . Moreover,  $\mathcal{P}_{\epsilon, n}$  also determines the set of rules which, at any round of the protocol, map the vectors of previous outcomes  $\{\mathbf{k}_i\}_{i=0}^{n-1}$  to the channel and LOCC operations applied in the following. In the remainder of this chapter, we will often have to average some function  $F(\mathbf{k}_n)$  over all possible LOCC outcomes  $\{\mathbf{k}_n\}$ . It is thus convenient to introduce the shorthand notation

$$\langle F \rangle_{\mathcal{P}_{\epsilon, n}} \equiv \sum_{\mathbf{k}_n} p(\mathbf{k}_n) F(\mathbf{k}_n), \quad (4.25)$$

where  $p(\mathbf{k}_n)$  is the probability of obtaining this particular sequence of LOCC outcomes according to the protocol  $\mathcal{P}_{\epsilon, n}$ .

We point out that the number of channels used when following the protocol  $\mathcal{P}_{\epsilon, n}$  will typically be smaller than  $n$ . This is because in any round the parties *may* decide to use a channel of the network, but are not forced to do so. However, without loss of generality, we can assume that a channel is used in any round of the protocol up to a certain outcomes-dependent point: after that stage, the parties can only perform LOCCs, and the communication protocol is effectively aborted. Indeed, if this was not the case, we could recover the aforementioned situation simply by merging all the LOCCs performed between two channel uses into a single round of LOCCs. In particular, for any edge  $l \in L$

and vector  $\mathbf{k}_n$ , we can define as  $m^{(l)}(\mathbf{k}_n)$  the total number of times channel  $\mathcal{N}^{(l)}$  has been used in that particular realisation of outcomes. Formally, this can be written as

$$m^{(l)}(\mathbf{k}_n) = \sum_{i=0}^{n-1} \delta_{l, l_{\mathbf{k}_i}}, \quad (4.26)$$

where the symbol  $\delta$  represents the Kronecker delta, while the total number of channels used when  $\mathbf{k}_n$  is obtained is

$$m(\mathbf{k}_n) = \sum_{l \in L} m^{(l)}(\mathbf{k}_n). \quad (4.27)$$

A value of  $m(\mathbf{k}_n)$  strictly smaller than  $n$  means that the protocol has been effectively interrupted after  $m(\mathbf{k}_n)$  rounds.

### 4.2.5 Quantifying the performance of a communication protocol

The quality of a point-to-point adaptive communication protocol  $\mathcal{P}_{\epsilon, n}$  depends on its ability to produce a large number of shared ebits (or pbits) between Alice and Bob. For any realisation  $\mathbf{k}_n$  of LOCC outcomes, this number can be quantified by the logarithm of the dimension  $d_{\mathbf{k}_n}$  that characterises the target state  $\phi_{AB}(d_{\mathbf{k}_n})$ , which is  $\epsilon$ -close to the final state  $\rho_{AB}^{\mathbf{k}_n}$  produced by the protocol. Therefore, a good figure of merit for  $\mathcal{P}_{\epsilon, n}$  can be obtained by averaging this quantity over all LOCC outcomes. In our notation, this can be written as  $\langle \log_2 d \rangle_{\mathcal{P}_{\epsilon, n}}$ .

This approach is particularly suitable to characterise the performance of protocols that use the channels of the network a finite number of times, because it directly provides the amount of ebits (pbits) that Alice and Bob can expect to share at the end of the communication. However, the quantity  $\langle \log_2 d \rangle_{\mathcal{P}_{\epsilon, n}}$  becomes unbounded when the asymptotic limit of infinitely many channel uses is considered. In a single-channel scenario, this issue is usually addressed by considering the communication rate, i.e., the number of produced bits per channel use. We point out that in this case the possibility of interrupting the communication depending on previous LOCC outcomes is typically not allowed. This is because otherwise with non-zero probability the asymptotic regime of infinitely many channel uses would not be reached. It is possible to enforce an interruption-free strategy by taking into account only protocols which make use of a quantum channel after *every* LOCC round. In this chapter a protocol of this kind will be labelled as  $\tilde{\mathcal{P}}_{\epsilon, N}$ , where  $\epsilon$  represents the error threshold and  $N$  is the fixed number of channel uses. With this notation, the quantum (or private) capacity of a quantum channel  $\mathcal{N}$  assisted by two-way classical communication can be obtained by first optimising the rate over all protocols  $\tilde{\mathcal{P}}_{\epsilon, N}$ , and then by taking the limits for  $N \rightarrow \infty$  and  $\epsilon \rightarrow 0$ :

$$C(\mathcal{N}) = \lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \sup_{\tilde{\mathcal{P}}_{\epsilon, N}} \frac{\langle \log_2 d \rangle_{\tilde{\mathcal{P}}_{\epsilon, N}}}{N}. \quad (4.28)$$

Note that the average over all LOCC outcomes still appears, because the results of the performed measurements can alter the number of ebits (pbits) shared at the end of the protocol.

For a generic quantum network the situation is more involved, and in the literature one can find multiple ways of assessing its communication performance in the asymptotic limit. For example, one can fix the frequency with which each channel is used, and divide  $\langle \log_2 d \rangle_{\mathcal{P}_{\epsilon,n}}$  by the total number of channel uses [188]. Other options, proposed in Ref. [177], consist in using each “path” connecting Alice and Bob with a certain probability, or in using each channel of the network exactly once. Then, the number of produced ebits (pbits) is respectively divided by the number of paths used, or by the total number of times the network has been accessed. Although the details characterising the considered figure of merit can change on a case-by-case basis, one typically has to optimise  $\langle \log_2 d \rangle_{\mathcal{P}_{\epsilon,n}}$  over a chosen class of protocols, and divide it by a quantity that counts how many times a basic operation has been repeated.

Similarly to Refs. [173, 188], in the following we are able to provide an upper bound on  $\langle \log d \rangle_{\mathcal{P}_{\epsilon,n}}$  for a generic adaptive protocol running on a quantum network with arbitrary topology. This bound only depends on the maximum amount of entanglement that could be generated by the quantum channels composing the network, and on the number of times each channel is used. From the previous discussion, it should be clear that our bound can be easily converted in a bound on a broad class of figures of merit which could be chosen to quantify the performance of the network. For example, in the case of a single channel, our bound can be reduced to an upper bound on its capacity by means of Equation (4.28).

### 4.3 Entanglement-based upper bounds

As we already briefly mentioned in the introduction, several authors have recently provided upper bounds on the number of ebits (pbits) shared by two parties at the end of a point-to-point quantum communication protocol assisted by two-way classical communication. Some studies deal with the capacity of a single quantum channel [164–167, 169], whereas others consider quantum networks with arbitrary topology [173, 177, 188]. Despite their differences and peculiarities, all these papers share some common features. Here we identify these, and show how they can lead to known and new communication bounds.

In this section, we start by considering a single channel  $\mathcal{N}$  and a generic entanglement measure  $E$ , and we summarise in Theorem 4.1 two important properties that have been used in the past in order to obtain upper bounds on channel capacities. One advantage of this abstract formulation is that of easing the identification of new entanglement measures which can be used to bound the capacity of a given channel. By comparing all these bounds, it would then be possible to select the one with the smallest value, which

represents the best known upper bound on the capacity  $C(\mathcal{N})$ . Moreover, the same two properties are also the main ingredients used in Ref. [173] to derive an upper bound for the number of shared ebits (pbits) produced by a quantum network. For this reason, a second advantage of our abstract approach lies in the possibility of easily extending previous results on quantum networks to other entanglement measures, not explicitly studied in the original papers. In particular, we show how the bound of Ref. [173], originally expressed in terms of the squashed entanglement, is also valid for other entanglement measures:  $E_{\max}$  and  $E_R$ , although the applicability of the latter is restricted to networks composed by Choi-stretchable channels. This is the result of Theorem 4.2, which can be considered the first important original contribution of this chapter.

As a consequence of the results presented in this section, multiple upper bounds on the communication performance of a quantum network become available, each based on a different entanglement measure. This opens the way for the possibility of combining them together, in order to obtain a bound as tight as possible. An obvious option consists in evaluating each upper bound separately, and then selecting the one yielding the smallest value. However, it is possible to do better than this, and in Section 4.4 we show how the bounds based on  $E_{\max}$  and  $E_R$  can be joined together in order to form a single tighter bound, which will be the main result of this chapter.

### 4.3.1 Key properties

To the best of the author's knowledge, any pair  $(E, \mathcal{N})$  currently leading to an upper bound on the channel capacity similar to Equation (4.2) satisfies two properties. The first resembles a continuity property for  $E$ , while the second concerns the relation between the entanglement content of the state used as input of the channel  $\mathcal{N}$  and its output. Here we discuss both these properties in detail, but one should keep in mind that a full characterisation of the pairs  $(E, \mathcal{N})$  satisfying these properties has not been completed yet.

The first property bounds the amount of entanglement present in a state that is  $\epsilon$ -close to a target state, i.e., either a maximally entangled or a private state.

- P1.** If a target state  $\phi_{AB}(d)$  is  $\epsilon$ -close to a quantum state  $\rho_{AB}$ , i.e., if  $\|\rho_{AB} - \phi_{AB}(d)\|_1 \leq \epsilon$ , then there exist two real functions  $f_E$  and  $g_E$ , with

$$\lim_{\epsilon \rightarrow 0} g_E(\epsilon) = 1, \quad \lim_{\epsilon \rightarrow 0} f_E(\epsilon) = 0, \quad (4.29)$$

such that

$$E(\rho_{AB}) \geq g_E(\epsilon) \log_2 d - f_E(\epsilon). \quad (4.30)$$

For a maximally entangled target state, this property can be easily proven for every *asymptotically continuous* [189] measure  $E$ . Let us show this explicitly. A function  $f$  is

asymptotic continuous if for any two  $D$ -dimensional states  $\rho$  and  $\sigma$ , such that  $\|\rho - \sigma\|_1 \leq \epsilon$ , one has

$$|f(\rho) - f(\sigma)| \leq k_1 \epsilon \log_2 D + O(\epsilon), \quad (4.31)$$

where  $k_1$  is a constant, while  $O(\epsilon) \rightarrow 0$  for  $\epsilon \rightarrow 0$  and is independent from  $D$ . If  $E$  is an asymptotically continuous entanglement measure, and the target state  $\phi_{AB}(d)$  is a maximally entangled state, by exploiting Equation (4.31) the hypothesis of property P1 yields

$$E(\rho_{AB})(d) \geq E(\phi_{AB})(d) - 2k_1 \epsilon \log_2 d - O(\epsilon) = (1 - 2k_1 \epsilon) \log_2 d - O(\epsilon), \quad (4.32)$$

because the total dimension of the system  $AB$  is  $D = d^2$ , and  $E(\phi_{AB})(d) = \log_2 d$ . This result coincides with Equation (4.30) by choosing  $g_E(\epsilon) = 1 - 2k_1 \epsilon$  and  $f_E(\epsilon) = O(\epsilon)$ . For the specific case of the relative entropy of entanglement, the asymptotic continuity is such that  $k_1 = 4$  and  $O(\epsilon) = -2[\epsilon \log_2 \epsilon + (1 - \epsilon) \log_2(1 - \epsilon)]$  [189], so

$$f_{E_R}(\epsilon) = -2[\epsilon \log_2 \epsilon + (1 - \epsilon) \log_2(1 - \epsilon)], \quad g_{E_R}(\epsilon) = 1 - 8\epsilon. \quad (4.33)$$

On the contrary, some more effort is usually needed to prove property P1 for private target states. This is because in this case the quantity  $d$  on the right-hand side of Equation (4.30) needs to be the dimension of the key systems, not of the whole key-shield local states. Nonetheless, this property has been proven for  $E_{\text{sq}}$  [190], and  $E_R$  [162, 166]. It can also be easily proven for  $E_{\text{max}}$ , by slightly varying the proof of Lemma IV.2 in Ref. [167] in order to obtain Equation (4.30) with

$$g_{E_{\text{max}}} = 1, \quad f_{E_{\text{max}}} = -2 \log_2(1 - \epsilon/2). \quad (4.34)$$

See Appendix B.1 for a short sketch of the proof. For the sake of completeness, at the end of this paragraph we explicitly prove property P1 for the relative entropy of entanglement, by following one of the proofs provided in Ref. [166].

The second important property that we discuss connects the entanglement of a state found as output of a channel  $\mathcal{N}$  with the entanglement already present in input.

**P2.** A pair  $(E, \mathcal{N})$  satisfies this property if for all states  $\rho_{AA'B}$  one has

$$E[\mathcal{N}_{A' \rightarrow B'}(\rho_{AA'B})] \leq E(\mathcal{N}) + E(\rho_{AA'B}), \quad (4.35)$$

where  $E(\mathcal{N})$  is the maximum entanglement shared through a single use of the channel, as defined in Equation (4.1).

This property is known to hold for any quantum channel when  $E = E_{\text{sq}}$  [164, 165], and for any Choi-stretchable channel when  $E = E_R$  [166, 169]. This last case has been already explicitly discussed when Choi-stretchable channels have been introduced [see



Equation (4.22)]. Moreover, the same result has been recently shown for the max-relative entropy of entanglement [167] when the channel acts on finite-dimensional systems, but it is conjectured to hold even without this last assumption.

*Proof sketch: property P1 for the relative entropy of entanglement [166].*

Say that the target state is the private state  $\gamma_{ABA'B'}(d)$  defined in Equation (4.18). We explicitly split the systems under control of Alice and Bob into a key-like and a shield-like part, so that the final state provided by the protocol can be written as  $\rho_{ABA'B'}$  rather than as  $\rho_{AB}$ . Then, let us defined the map  $\tau$ , which “untwist” any quantum state defined in  $ABA'B'$ , in order to retrieve its key component:

$$\tau(\rho_{ABA'B'}) = \text{Tr}_{A'B'} \left[ U_{ABA'B'}^{(\text{twist})\dagger} \rho_{ABA'B'} U_{ABA'B'}^{(\text{twist})} \right]. \quad (4.36)$$

In particular, because of the monotonicity of the trace distance under generic completely-positive and trace-preserving maps, one has

$$\|\tau(\rho_{ABA'B'}) - \psi_{AB}(d)\|_1 \leq \|\rho_{ABA'B'} - \gamma_{ABA'B'}\|_1 \leq \epsilon, \quad (4.37)$$

where  $\psi_{AB}(d)$  is the maximally entangled state obtained by untwisting the private state. It is also convenient to introduce the set  $\tau\text{SEP}$  obtained by untwisting any  $AA' : BB'$  separable state, which can be used to define a modified version of the relative entropy of entanglement:

$$E_{\text{R}}^{(\tau)}(\rho_{AB}) = \min_{\sigma_{AB} \in \tau\text{SEP}} S(\rho_{AB} || \sigma_{AB}) \geq E_{\text{R}}(\rho_{AB}), \quad (4.38)$$

where the inequality is due to the fact that  $\tau\text{SEP} \subset \text{SEP}$ . As  $E_{\text{R}}^{(\tau)}$  is asymptotic continuous (for the same reason of  $E_{\text{R}}$ , because the optimisation set is convex and it contains the maximally mixed state [189]), we can write

$$E_{\text{R}}^{(\tau)}[\psi_{AB}(d)] \leq E_{\text{R}}^{(\tau)}[\tau(\rho_{ABA'B'})] + 2k_1 \epsilon \log_2 d + O(\epsilon). \quad (4.39)$$

Now let  $\sigma_{ABA'B'}^{(0)}$  be the separable state such that  $E_{\text{R}}(\rho_{ABA'B'}) = S(\rho_{ABA'B'} || \sigma_{ABA'B'}^{(0)})$ , which implies

$$E_{\text{R}}^{(\tau)}[\tau(\rho_{ABA'B'})] \leq S(\tau(\rho_{ABA'B'}) || \tau(\sigma_{ABA'B'}^{(0)})) \leq E_{\text{R}}(\rho_{ABA'B'}). \quad (4.40)$$

The first inequality comes from selecting a particular element of  $\tau\text{SEP}$ , whereas the second one follows from the monotonicity of the relative entropy under the map  $\tau$ . The final observation is that because of the inequality in Equation (4.38) one has

$$E_{\text{R}}^{(\tau)}[\psi_{AB}(d)] \geq E_{\text{R}}[\psi_{AB}(d)] = \log_2 d, \quad (4.41)$$

which together with Equations (4.39) and (4.40) yields

$$E_R(\rho_{ABA'B'}) \geq (1 - 2k_1\epsilon) \log_2 d - O(\epsilon). \quad (4.42)$$

This concludes the proof of property P1, which can be written in terms of the same functions  $f_{E_R}(\epsilon)$  and  $g_{E_R}(\epsilon)$  of Equation (4.33).  $\square$

### 4.3.2 Upper bounds based on a single entanglement measure

Now we show how the properties P1 and P2 previously introduced can lead to upper bounds on the number of ebits (pbits) generated by a single channel or a quantum network. In order to ease the comparison between the two, we find beneficial to express both results as bounds on  $\langle \log_2 d \rangle_{\mathcal{P}_{\epsilon,n}}$ . The usual upper bounds on single-channel capacities [164–167, 169] can then be easily recovered as a corollary, by using the definition given in Equation (4.28).

**Theorem 4.1.** *If  $E$  and  $\mathcal{N}$  satisfy properties P1 and P2, the average number of ebits (pbits) generated by a single quantum channel via an adaptive protocol  $\mathcal{P}_{\epsilon,n}$ , assisted by two-way classical communication, can be upper bounded as*

$$\langle \log_2 d \rangle_{\mathcal{P}_{\epsilon,n}} \leq \frac{1}{g_E(\epsilon)} \left[ f_E(\epsilon) + \langle m \rangle_{\mathcal{P}_{\epsilon,n}} E(\mathcal{N}) \right], \quad (4.43)$$

where  $\langle m \rangle_{\mathcal{P}_{\epsilon,n}}$  is the average number of times the channel has been used.

**Corollary 4.1.** *If  $E$  and  $\mathcal{N}$  satisfy properties P1 and P2, the capacity of  $\mathcal{N}$  assisted by two-way classical communication can be upper bounded as*

$$C(\mathcal{N}) \leq E(\mathcal{N}). \quad (4.44)$$

*Proof of Corollary 4.1.* By definition of capacity [see Equation (4.28)], only protocols using the channel a fixed number of times should be considered. This corollary is a straightforward consequence of  $\langle m \rangle_{\tilde{\mathcal{P}}_{\epsilon,N}} = N$ ,  $\lim_{\epsilon \rightarrow 0} g_E(\epsilon) = 1$ , and  $\lim_{\epsilon \rightarrow 0} f_E(\epsilon) = 0$ .  $\square$

We should point out that  $E_R(\mathcal{N})$  and  $E_{\max}(\mathcal{N})$  have also been shown to be strong converse bounds on the capacity [167, 169]. This means not only that these quantities are larger than the maximum achievable rate of faithful transmission (i.e., in the limit of vanishing error), but that the error inevitably associated with any protocol with a larger rate grows exponentially fast with the number of channel uses.

Intuitively, a bipartite scenario  $A : B$  is much easier to study than a situation in which Alice and Bob need to cooperate with other nodes  $\{C_i\}_{i=1}^N$  of the network in order to achieve their communication goal. Building on this idea, the authors of Refs. [173, 177] derived upper bounds on the number of ebits (pbits) produced by a communication

protocol over a quantum network by considering a bipartition  $AC_A : C_BB$ . In particular, they imagined to extend the regions controlled by Alice and Bob so as to include in them also the remaining nodes on their same component of the bipartition. Intuitively, this construction can lead to an upper bound on the communication performance of the real Alice and Bob, because the communication is much easier for the “extended” parties, who can perform global operations on the nodes under their control. In this framework, only the channels associated with the edges in  $L_{C_A}$  can contribute to the communication bound, because they are the only ones connecting the two extended parties, and any given bipartition  $\{C_A, C_B\}$  of the network leads to a different upper bound. Although the proof that led to the result in Ref. [173] was based on a particular choice of entanglement measure, we can see how a similar reasoning applies to any entanglement measure which satisfies properties P1 and P2 for any channel associated with an edge in  $L_{C_A}$ . This is the result of the next theorem, which can be considered the first important original result of this chapter.

**Theorem 4.2.** *Consider a quantum network with associated directed graph  $\mathcal{G}$ . For a given bipartition  $\{C_A, C_B\}$  of the network nodes  $\{C_i\}_i$ , let  $L_{C_A} \subset L$  be the set of edges in  $\mathcal{G}$  that connect a node in  $AC_A$  with one in  $C_BB$ . The average number of ebits (or pbits) that Alice and Bob share at the end of a given adaptive protocol  $\mathcal{P}_{\epsilon,n}$ , assisted by two-way classical communication, can be upper bounded as:*

$$\langle \log_2 d \rangle_{\mathcal{P}_{\epsilon,n}} \leq \frac{1}{g_E(\epsilon)} [f_E(\epsilon) + \mathcal{B}_E(\mathcal{P}_{\epsilon,n}, C_A)], \quad (4.45)$$

where

$$\mathcal{B}_E(\mathcal{P}_{\epsilon,n}, C_A) \equiv \sum_{l \in L_{C_A}} \langle m^{(l)} \rangle_{\mathcal{P}_{\epsilon,n}} E(\mathcal{N}^{(l)}), \quad (4.46)$$

for any entanglement measure  $E$  satisfying properties P1 and P2 for any channel  $\mathcal{N}^{(l)}$  with  $l \in L_{C_A}$ .

At this point we can make a few comments on this bound. Because of Theorem 4.1, the entanglement  $E(\mathcal{N}^{(l)})$  of any channel  $\mathcal{N}^{(l)}$  has to be larger than its single-channel capacity  $C(\mathcal{N}^{(l)})$ . Therefore, the gap between the two sides of Equation (4.45) can be reduced by choosing an entanglement measure which leads to better approximations of the capacities of channels  $\mathcal{N}^{(l)}$  with  $l \in L_{C_A}$ . For any considered bipartition, the tightest available bound is the one based on the entanglement measure that yields the smallest possible  $\mathcal{B}_E(\mathcal{P}_{\epsilon,n}, C_A)$  in Equation (4.46), among those satisfying properties P1 and P2 for any channel  $\mathcal{N}^{(l)}$  with  $l \in L_{C_A}$ . By formally labelling a generic entanglement measure in this set as  $E|_{C_A}$ , the following bound can be obtained after an optimisation over all

possible bipartitions:

$$\langle \log_2 d \rangle_{\mathcal{P}_{\epsilon,n}} \leq \min_{\mathcal{C}_A} \min_{E|_{\mathcal{C}_A}} \frac{1}{g_E(\epsilon)} [f_E(\epsilon) + \mathcal{B}_E(\mathcal{P}_{\epsilon,n}, \mathcal{C}_A)]. \quad (4.47)$$

### 4.3.3 Proofs of Theorems 4.1 and 4.2

Any single channel can be interpreted as a simple quantum network, so we first show how Theorem 4.1 can be derived from Theorem 4.2, and then we prove the latter. The ideas that will be used for these proofs closely follow those of Refs. [164–167, 169, 173, 177, 188].

*Proof of Theorem 4.1.* For a single-channel scenario, the only possible bipartition  $AC_A : C_BB$  of the network is the trivial one  $A : B$ . Moreover, at every round of the adaptive strategy, the only channel Alice and Bob can use is  $\mathcal{N}_{A \rightarrow B}$ , which is associated with the only edge  $l_0$  of the graph. Therefore, for all  $\mathbf{k}_n$

$$m^{(l_0)}(\mathbf{k}_n) = m(\mathbf{k}_n), \quad (4.48)$$

and the thesis of Theorem 4.2 reduces to

$$\langle \log_2 d \rangle_{\mathcal{P}_{\epsilon,n}} \leq \frac{1}{g_E(\epsilon)} \left[ f_E(\epsilon) + \langle m \rangle_{\mathcal{P}_{\epsilon,n}} E(\mathcal{N}) \right]. \quad (4.49)$$

□

*Proof of Theorem 4.2.* In this proof we will make use of the notation introduced in Section 4.2.4 to describe a generic adaptive protocol  $\mathcal{P}_{\epsilon,n}$ . Property P1, together with Equation (4.24), implies:

$$\log_2 d_{\mathbf{k}_n} \leq \frac{1}{g_E(\epsilon)} (f_E(\epsilon) + E^{A:B}(\rho_{AB}^{\mathbf{k}_n})). \quad (4.50)$$

By exploiting the monotonicity of  $E$  under partial trace, and by averaging over all possible outcomes, we can write for any bipartition  $\{\mathcal{C}_A, \mathcal{C}_B\}$  of the set of nodes  $\{C_i\}_i$ :

$$\langle \log_2 d \rangle_{\mathcal{P}_{\epsilon,n}} = \sum_{\mathbf{k}_n} p(\mathbf{k}_n) \log_2 d_{\mathbf{k}_n} \leq \frac{1}{g_E(\epsilon)} \left[ f_E(\epsilon) + \sum_{\mathbf{k}_n} p(\mathbf{k}_n) E^{AC_A:CB}(\rho^{\mathbf{k}_n}) \right], \quad (4.51)$$

where  $\rho^{\mathbf{k}_n}$  is the final state of the protocol, spread across the whole network. The summation appearing in square brackets on the right-hand side can be expanded into two terms as

$$\sum_{\mathbf{k}_n} p(\mathbf{k}_n) E^{AC_A:CB}(\rho^{\mathbf{k}_n}) \leq \sum_{\mathbf{k}_{n-1}} p(\mathbf{k}_{n-1}) E^{AC_A:CB}(\rho^{\mathbf{k}_{n-1}}) + \sum_{\mathbf{k}_n} p(\mathbf{k}_n) \sum_{l \in L_{\mathcal{C}_A}} \delta_{l, l_{\mathbf{k}_{n-1}}} E[\mathcal{N}^{(l)}]. \quad (4.52)$$

The former has the same structure of the original expression, but is evaluated on the previous round of the protocol, while the latter characterises the ability of the last used channel to create entanglement across the chosen bipartition  $AC_A : C_BB$ . In particular, the second term does not always appear, because the channel  $\mathcal{N}^{(l_{k_{n-1}})}$  might not connect  $AC_A$  with  $C_BB$ , or the parties may have decided not to use a channel at all. This last case could be represented, for example, by any value of  $l_{k_{n-1}}$  not in the set  $L$  of the graph edges. In order to prove Equation (4.52), we can first expand the left-hand side as

$$\sum_{\mathbf{k}_n} p(\mathbf{k}_n) E^{AC_A:C_BB}(\rho^{\mathbf{k}_n}) = \sum_{\mathbf{k}_{n-1}} p(\mathbf{k}_{n-1}) \left[ \sum_{k_n} p(k_n|\mathbf{k}_{n-1}) E^{AC_A:C_BB}(\rho^{\mathbf{k}_n}) \right], \quad (4.53)$$

and then use the following chain of inequalities:

$$\begin{aligned} \sum_{k_n} p(k_n|\mathbf{k}_{n-1}) E^{AC_A:C_BB}(\rho^{\mathbf{k}_n}) &\stackrel{(i)}{\leq} E^{AC_A:C_BB}(\rho^{\mathbf{k}_{n-1}}) \\ &\stackrel{(ii)}{\leq} E^{AC_A:C_BB}(\rho^{\mathbf{k}_{n-1}}) + \sum_{l \in L_{C_A}} \delta_{l, l_{k_{n-1}}} E[\mathcal{N}^{(l)}], \end{aligned} \quad (4.54)$$

where (i) is due to the monotonicity of  $E$  under LOCC operations, while (ii) follows directly from property P2. After combining Equations (4.54) and (4.53), we can recover Equation (4.52) simply by noticing that the average over  $\mathbf{k}_{n-1}$  on the right-most term can be freely turned into an average over  $\mathbf{k}_n$ . This same procedure can be iteratively applied for every round of the protocol, so that in the end we are left with:

$$\begin{aligned} \sum_{\mathbf{k}_n} p(\mathbf{k}_n) E^{AC_A:C_BB}(\rho^{\mathbf{k}_n}) &\leq E^{AC_A:C_BB}(\rho^{(1)}) + \sum_{j=0}^{n-1} \sum_{\mathbf{k}_n} p(\mathbf{k}_n) \sum_{l \in L_{C_A}} \delta_{l, l_{k_j}} E[\mathcal{N}^{(l)}] \\ &= \sum_{l \in L_{C_A}} \langle m^{(l)} \rangle E[\mathcal{N}^{(l)}], \end{aligned} \quad (4.55)$$

where the last equality is due to the separability of the initial state  $\rho^{(1)}$ , and to the definition of  $\langle m^{(l)} \rangle$  given in Equation (4.26). At this point, the thesis of Theorem 4.2 follows directly from the inequality given in Equation (4.51).  $\square$

## 4.4 Versatile upper bound on quantum networks

As we have seen, an entanglement measure  $E$  can lead to an upper bound on the capacity of a quantum channel if it satisfies a continuity inequality (property P1), and a recursive relation (property P2), connecting the entanglement of the state before and after the channel application. In the previous section, we discussed how we can optimise the choice of the entanglement measure on each bipartition, as long as we make sure that the chosen

entanglement measure satisfies property P2 for *every* channel  $\mathcal{N}^{(l)}$  with  $l \in L_{\mathcal{C}_A}$ . This constraint leads to upper bounds that are less tight than what would be obtained if we could change entanglement measure on a channel-by-channel basis. For example, consider a situation where all the channels in a given bipartition are Choi-stretchable, with only one exception: the presence of this single unstretchable channel prevents us from using  $E_R$  in the bound of Theorem 4.2. Instead, we are forced to use some broadly applicable entanglement measure, as  $E_{\text{sq}}$  or  $E_{\text{max}}$ , on every channel of the bipartition, thus loosening the bound.

In this section we overcome this issue, by exploiting a recent result on sandwiched Rényi entropies [167]. In particular, we construct an upper bound on  $\langle \log_2 d \rangle_{\mathcal{P}_{\epsilon,n}}$  that allows us to apply either  $E_R$  or  $E_{\text{max}}$  to any channel  $\mathcal{N}^{(l)}$  with  $l \in L_{\mathcal{C}_A}$ , depending on its Choi-stretchability. To begin with, in the following we describe the recent result obtained in Ref. [167], which is the cornerstone of our method. Then, we prove our main result.

#### 4.4.1 Versatile Property P2

The authors of Ref. [167] recently showed that for any quantum channel  $\mathcal{N}_{A' \rightarrow B'}$ , and any real parameter  $1 \leq \alpha < \infty$ , if

$$\tilde{\rho}_{AB'B} \equiv \mathcal{N}_{A' \rightarrow B'}(\rho_{AA'B}), \quad (4.56)$$

one has:

$$E_\alpha(\tilde{\rho}_{AB'B}) \leq E_{\text{max}}(\mathcal{N}_{A' \rightarrow B'}) + E_\alpha(\rho_{AA'B}). \quad (4.57)$$

Here the tilde is simply used to label the output state of the quantum channel, and it is introduced with the only purpose of easing the application of this result to a scenario in which the parties communicate over a quantum network. Indeed, in this case with the notation of Sec. 4.2.4 the tilded states represent the output of quantum channels, and non-tilded states are the results of LOCC operations. The quantity  $E_\alpha$  is defined in terms of the sandwiched Rényi relative entropy  $\tilde{D}_\alpha$  [191, 192], which interpolates between the relative entropy  $S$  and the max-relative entropy  $D_{\text{max}}$ :

$$\begin{aligned} E_\alpha(\rho_{AB}) &= \min_{\sigma_{AB} \in \text{SEP}} \tilde{D}_\alpha(\rho_{AB} || \sigma_{AB}) \\ &\equiv \min_{\sigma_{AB} \in \text{SEP}} \begin{cases} \frac{1}{\alpha-1} \log_2 \text{Tr} \left[ \left( \sigma_{AB}^{\frac{1-\alpha}{2\alpha}} \rho_{AB} \sigma_{AB}^{\frac{1-\alpha}{2\alpha}} \right)^\alpha \right], & \text{if } \text{Supp}(\rho_{AB}) \subseteq \text{Supp}(\sigma_{AB}), \\ \infty, & \text{otherwise,} \end{cases} \end{aligned} \quad (4.58)$$

where  $\sigma_{AB}$  is optimised over all separable states. The quantity  $E_\alpha$  tends respectively to  $E_R$  and  $E_{\text{max}}$  in the limits of  $\alpha \rightarrow 1$  and  $\alpha \rightarrow \infty$ , so by setting  $\alpha = 1$  in Equation (4.59)

we obtain

$$E_R(\tilde{\rho}_{AB'B}) \leq E_{\max}(\mathcal{N}_{A' \rightarrow B'}) + E_R(\rho_{AA'B}). \quad (4.59)$$

This inequality closely resembles property P2 for  $E_R$ , which was obtained in Equation (4.22) for Choi-stretchable channels. However, thanks to the introduction of  $E_{\max}$  on the right hand side, Equation (4.59) now holds even for non-stretchable channels. If we combine Equation (4.22) with Equation (4.59), we can obtain a versatile property P2 for the relative entropy of entanglement, in which the right-hand side changes according to the Choi-stretchability of  $\mathcal{N}_{A' \rightarrow B'}$ :

$$E_R(\tilde{\rho}_{AB'B}) \leq E_R(\rho_{AA'B}) + \begin{cases} E_R(\mathcal{N}), & \text{if } \mathcal{N} \in \mathcal{S}, \\ E_{\max}(\mathcal{N}), & \text{otherwise,} \end{cases} \quad (4.60)$$

where  $\mathcal{S}$  is the set of Choi-stretchable channels. Note that this is the optimal choice, i.e., the choice that minimises the right-hand side, because  $E_R \leq E_{\max}$  for all channels [see Equation (4.11)].

#### 4.4.2 Upper bound combining multiple entanglement measures

With the tools developed so far, we can write a versatile upper bound on the average number of ebit (or pbits) shared by Alice and Bob at the end of a generic adaptive protocol  $\mathcal{P}_{\epsilon,n}$  over a quantum network, when they can freely communicate classically. This theorem is the main result of this chapter.

**Theorem 4.3.** *Consider a quantum network with an associated directed graph  $\mathcal{G}$ . For a given bipartition  $\{\mathcal{C}_A, \mathcal{C}_B\}$  of the network nodes  $\{C_i\}_i$ , let  $L_{\mathcal{C}_A} \subset L$  be the set of edges in  $\mathcal{G}$  that connect a node in  $\mathcal{C}_A$  with one in  $\mathcal{C}_B$ . The average number of ebits (or pbits) that Alice and Bob share at the end of a given adaptive communication protocol  $\mathcal{P}_{\epsilon,n}$  can be upper bounded as*

$$\langle \log_2 d \rangle_{\mathcal{P}_{\epsilon,n}} \leq \frac{1}{g_{E_R}(\epsilon)} [f_{E_R}(\epsilon) + \mathcal{B}'(\mathcal{P}_{\epsilon,n}, \mathcal{C}_A)], \quad (4.61)$$

where

$$\mathcal{B}'(\mathcal{P}_{\epsilon,n}, \mathcal{C}_A) \equiv \sum_{\substack{l \in L_{\mathcal{C}_A}: \\ \mathcal{N}^{(l)} \in \mathcal{S}}} \langle m^{(l)} \rangle_{\mathcal{P}_{\epsilon,n}} E_R(\mathcal{N}^{(l)}) + \sum_{\substack{l \in L_{\mathcal{C}_A}: \\ \mathcal{N}^{(l)} \notin \mathcal{S}}} \langle m^{(l)} \rangle_{\mathcal{P}_{\epsilon,n}} E_{\max}(\mathcal{N}^{(l)}), \quad (4.62)$$

with  $f_{E_R}(\epsilon) = -2[\epsilon \log_2 \epsilon + 2(1 - \epsilon) \log_2(1 - \epsilon)]$  and  $g_{E_R}(\epsilon) = 1 - 8\epsilon$ .

*Proof of Theorem 4.3.* The proof closely follows the one provided for Theorem 4.2, with  $E = E_R$ . The only difference lies in Equation (4.54), where we use the new inequality written in Equation (4.60) instead of the original property P2. Therefore, Equation (4.54)

has to be substituted with

$$E_{\mathbf{R}}^{AC_A:CB}(\tilde{\rho}^{k_{n-1}}) \leq E_{\mathbf{R}}^{AC_A:CB}(\rho^{k_{n-1}}) + \sum_{\substack{l \in L_{C_A}: \\ \mathcal{N}^{(l)} \in \mathcal{S}}} \delta_{l,k_{n-1}} E_{\mathbf{R}}(\mathcal{N}^{(l)}) + \sum_{\substack{l \in L_{C_A}: \\ \mathcal{N}^{(l)} \notin \mathcal{S}}} \delta_{l,k_{n-1}} E_{\max}(\mathcal{N}^{(l)}), \quad (4.63)$$

where we split the sum over the Choi-stretchable and non-Choi-stretchable channels connecting the two components of the considered bipartition. The remainder of the proof then follows the same steps used in the proof of Theorem 4.2. We also explicitly reported the expressions for the functions  $f_{E_{\mathbf{R}}}(\epsilon)$  and  $g_{E_{\mathbf{R}}}(\epsilon)$  [see Equation (4.33)].  $\square$

Thanks to this result, we have been able to merge the upper bounds based on the quantities  $\mathcal{B}_{E_{\mathbf{R}}}$  and  $\mathcal{B}_{E_{\max}}$  into a single bound, having the advantages of both entanglement measures: tightness and broad applicability. Therefore, in assessing the communication performance of an adaptive protocol  $\mathcal{P}_{\epsilon,n}$  over a quantum network, for any given bipartition  $AC_A : CB$  we just need to compare  $\mathcal{B}'$  with the quantity  $\mathcal{B}_{E_{\text{sq}}}$  based on the squashed entanglement [173]. This is because we can ignore the dependence of the upper bound on the quantities  $f_E$  and  $g_E$ , which become irrelevant for small  $\epsilon$ . In particular, the advantage of using  $\mathcal{B}'$  over  $\mathcal{B}_{E_{\text{sq}}}$  for the bipartition  $AC_A : CB$  can be quantified by the parameter

$$\mu_{C_A}(\mathcal{P}_{\epsilon,n}) = \frac{\mathcal{B}_{E_{\text{sq}}}(\mathcal{P}_{\epsilon,n}, C_A) - \mathcal{B}'(\mathcal{P}_{\epsilon,n}, C_A)}{\mathcal{B}_{E_{\text{sq}}}(\mathcal{P}_{\epsilon,n}, C_A) + \mathcal{B}'(\mathcal{P}_{\epsilon,n}, C_A)}, \quad (4.64)$$

which is defined in the range  $[-1, +1]$ , and becomes positive when the versatile bound  $\mathcal{B}'$  is tighter than  $\mathcal{B}_{E_{\text{sq}}}$ . Although the sign of  $\mu_{C_A}(\mathcal{P}_{\epsilon,n})$  will ultimately depend on the details of the bipartition and on the average number of times each channel has been used, we can expect  $\mathcal{B}'$  to be tighter than  $\mathcal{B}_{E_{\text{sq}}}$  on bipartitions mostly connected by Choi-stretchable channels. In general, the sign of  $\mu_{C_A}(\mathcal{P}_{\epsilon,n})$  will strongly depend on the sign of  $E_{\text{sq}}(\mathcal{N}) - E_{\max}(\mathcal{N})$  on the considered non-stretchable channels: every one of them characterised by a positive difference will enhance the usefulness of  $\mathcal{B}'$  over  $\mathcal{B}_{E_{\text{sq}}}$ .

We should stress that Equation (4.64) usually cannot be easily computed, because it depends on the squashed entanglement or max-relative entropy of entanglement of the considered channels, quantities which are not known at the moment for most quantum channels. Nonetheless, a similar approach can be used to check whether the best known upper bound  $\tilde{\mathcal{B}}'$  on  $\mathcal{B}'$ , based on our versatile approach, is better than the best known upper bound  $\tilde{\mathcal{B}}_{E_{\text{sq}}}$  on  $\mathcal{B}_{E_{\text{sq}}}$ . These approximations are obtained by using available upper bounds on  $E_{\max}(\mathcal{N})$  and  $E_{\text{sq}}(\mathcal{N})$ , respectively in Equation (4.62) and Equation (4.46). If we write as  $\tilde{\mu}_{C_A}(\mathcal{P}_{\epsilon,n})$  the modified parameter obtained by using  $\tilde{\mathcal{B}}'$  and  $\tilde{\mathcal{B}}_{E_{\text{sq}}}$  in Equation (4.64), we can say that for the time being the versatile upper bound yields a better result than the bound based on the squashed entanglement when  $\tilde{\mu}_{C_A}(\mathcal{P}_{\epsilon,n}) > 0$ .



## 4.5 Max-relative entropy of entanglement of some qubit channels

Before discussing examples of networks where the bound provided by Theorem 4.3 becomes tighter than its counterpart based on the squashed entanglement, we first need to evaluate  $E_{\max}(\mathcal{N})$  for some channels of interest. In particular, in this section we will consider typical examples of quantum channels acting on qubit systems. More precisely, we develop technical tools in order to analytically evaluate lower and upper bounds on the max-relative entropy of entanglement of the amplitude damping channel  $\mathcal{N}^{(\text{ad})}$ . Being non-stretchable, its capacity is still unknown, although it has been recently bounded by several authors [166, 168]. Our analysis shows that it is very unlikely that  $E_{\max}$  could lead in the future to a better upper bound on  $C(\mathcal{N}^{(\text{ad})})$  than the squashed entanglement. Our strategy to bound the max-relative entropy of entanglement is based on the invariance under phase rotations of the amplitude damping channel, hence it can be adopted also for other qubit channels with the same property. In particular, we numerically evaluate the max-relative entropy of entanglement of three Choi-stretchable qubit channels: dephasing, depolarising and erasure channel.

In order to calculate the max-relative entropy of entanglement of a channel, one should perform a max-min optimisation [see Equations (4.4) and (4.10)]:

$$E_{\max}(\mathcal{N}) = \max_{\rho_{AA'}} \min_{\sigma_{AB} \in \text{SEP}} \inf_x \{x | 2^x \sigma_{AB} - \mathcal{N}_{A' \rightarrow B}[\rho_{AA'}] \geq 0\}. \quad (4.65)$$

However,  $E_{\max}(\mathcal{N})$  can be bounded from both sides as stated in the following proposition. The upper bound is a re-elaborated version of the upper bound on the max-relative entropy of entanglement of a channel studied in Ref. [167]. After proving this result, in the next technical section we simplify the optimisations appearing in the upper and lower bound of Equation (4.66) by exploiting as much as possible the symmetry under phase rotations of most qubit channels. After that, we will bound, or evaluate, the max-relative entropy of entanglement of the aforementioned qubit channels.

**Proposition 4.1.** *Let  $\pi_{\mathcal{N}} = \mathbb{1}_A \otimes \mathcal{N}_{A' \rightarrow B}[\psi_{AA'}]$  be the Choi-Jamiołkowski state associated with the quantum channel  $\mathcal{N}$ , with  $\psi_{AA'}$  a maximally entangled state. If  $d$  is the input dimension of  $\mathcal{N}$ , then we have*

$$\min_{\sigma_{AB} \in \text{SEP}} D_{\max}(\pi_{\mathcal{N}} || \sigma_{AB}) \leq E_{\max}(\mathcal{N}) \leq \min_{\substack{\sigma_{AB} \in \text{SEP} \\ \text{Tr}_B[\sigma_{AB}] = \mathbb{1}_A/d}} D_{\max}(\pi_{\mathcal{N}} || \sigma_{AB}). \quad (4.66)$$

Moreover, if  $\mathcal{N}$  is Choi-stretchable, the lower bound is equal to  $E_{\max}(\mathcal{N})$  itself.

*Proof.* The lower bound can be easily obtained by using the maximally entangled state  $\psi_{AA'}$  as input in Equation (4.65), without performing the optimisation over all  $\rho_{AA'}$ .

Furthermore, we can prove its equality with  $E_{\max}(\mathcal{N})$  itself for Choi-stretchable channels exactly as in the last step of Equation (4.22).

The upper bound is a re-elaborated version of the upper bound on the max-relative entropy of a channel studied in Ref. [167]. There, the authors considered the set of entanglement breaking (EB) channels, and wrote the following chain of inequalities:

$$\begin{aligned} E_{\max}(\mathcal{N}) &\leq \max_{\rho_{AA'}} \min_{\mathcal{T} \in \text{EB}} \inf_x \{x | (2^x \mathcal{T}_{A' \rightarrow B} - \mathcal{N}_{A' \rightarrow B}) [\rho_{AA'}] \geq 0\} \\ &\leq \min_{\mathcal{T} \in \text{EB}} \inf_x \{x | 2^x \pi_{\mathcal{T}} - \pi_{\mathcal{N}} \geq 0\} = \min_{\mathcal{T} \in \text{EB}} D_{\max}(\pi_{\mathcal{N}} || \pi_{\mathcal{T}}). \end{aligned} \quad (4.67)$$

The first inequality is obtained by optimising over a smaller set of separable states, in which  $\sigma_{AB}$  can be written as output of an entanglement-breaking channel acting on the same input state  $\rho_{AA'}$ . The second inequality then follows by noticing that one surely has

$$(2^x \mathcal{T}_{A' \rightarrow B} - \mathcal{N}_{A' \rightarrow B}) [\rho_{AA'}] \geq 0, \quad (4.68)$$

for any input  $\rho_{AA'}$ , if the operator  $(2^x \mathcal{T}_{A' \rightarrow B} - \mathcal{N}_{A' \rightarrow B})$  is completely positive. This last condition can be verified just by looking at its Choi-Jamiołkowski state, and Equation (4.67) is thus proven.

In order to obtain the upper bound of Proposition 4.1, we need to show that the set of states  $\pi_{\mathcal{T}}$  appearing in Equation (4.67) corresponds to the set of separable density matrices  $\sigma_{AB}$  such that  $\text{Tr}_B [\sigma_{AB}] = \mathbb{1}_A/d$ . This follows from the known fact that a CPT map is EB if and only if its Choi state is separable [193]. Therefore, for any EB channel  $\mathcal{T}$  its Choi state  $\pi_{\mathcal{T}}$  is separable and such that  $\text{Tr}_B [\pi_{\mathcal{T}}] = \mathbb{1}_A/d$ . On the other hand, given a separable state  $\sigma_{AB}$  with  $\text{Tr}_B [\sigma_{AB}] = \mathbb{1}_A/d$ , we can find the corresponding CPT map  $\mathcal{T}_{A' \rightarrow B}^{(\sigma_{AB})}$  as

$$\mathcal{T}_{A' \rightarrow B}^{(\sigma_{AB})}(\tau_{A'}) = d^2 \text{Tr}_{A'A} [\psi_{A'A} (\tau_{A'} \otimes \sigma_{AB})], \quad (4.69)$$

where  $\psi_{A'A}$  is a maximally entangled state. Notice that this map is CPT because from Equation (4.69) it is possible to obtain a set of Kraus operators

$$N_{A' \rightarrow B}^{(h,k)} = d_{A'A} \langle \psi | \sqrt{\sigma_{AB}} | k \rangle_A | h \rangle_B, \quad (4.70)$$

with  $\sum_{h,k=1}^d \left( N_{A' \rightarrow B}^{(h,k)} \right)^\dagger N_{A' \rightarrow B}^{(h,k)} = \mathbb{1}_{A'}$ . A straightforward calculation shows that  $\pi_{\mathcal{T}_{A' \rightarrow B}^{(\sigma_{AB})}}$  is separable and hence that  $\mathcal{T}_{A' \rightarrow B}^{(\sigma_{AB})} \in \text{EB}$ .  $\square$

### 4.5.1 Tools for qubit channels invariant under phase rotations

Most of the qubit channels that are typically studied are invariant under rotations around the axis associated with the Pauli operator  $\hat{\sigma}_z = |0\rangle\langle 0| - |1\rangle\langle 1|$ , where  $|0\rangle, |1\rangle$  are the elements of the computational basis. It is thus interesting to study the consequences of this fact for the evaluation of the upper and lower bounds identified in Proposition 4.1.

Let  $\mathcal{N}$  be a quantum channel acting on a qubit, such that

$$\mathcal{N}(e^{i\theta\hat{\sigma}_z}\rho e^{-i\theta\hat{\sigma}_z}) = e^{i\theta\hat{\sigma}_z}\mathcal{N}(\rho)e^{-i\theta\hat{\sigma}_z}, \quad (4.71)$$

for all angles  $\theta$  and input states  $\rho$ . As the maximally entangled state  $\psi_{AA'}$  is left invariant by the unitary operation

$$\hat{J}_\theta = e^{+i\frac{\theta}{2}\hat{\sigma}_z^{(A)}} \otimes e^{-i\frac{\theta}{2}\hat{\sigma}_z^{(B)}}, \quad (4.72)$$

also its Choi state  $\pi_{\mathcal{N}}$  is left invariant by  $\hat{J}_\theta$ , for any  $\theta \in [0, 2\pi]$ . Therefore, one also has

$$\pi_{\mathcal{N}} = \int \frac{d\theta}{2\pi} \hat{J}_\theta \pi_{\mathcal{N}} \hat{J}_\theta^\dagger. \quad (4.73)$$

This allows us to prove the following lemma.

**Lemma 4.1.** *Let  $\pi_{\mathcal{N}}$  be a bipartite state invariant under the separable unitary evolution  $\hat{J}_\theta$  defined in Equation (4.72), and  $\sigma_{AB}^*$  be the state which minimises  $D_{\max}(\pi_{\mathcal{N}}||\sigma_{AB})$  among all separable states  $\sigma_{AB}$ . If  $\overline{\sigma}_{AB}^*$  is the averaged version of  $\sigma_{AB}^*$ , i.e.*

$$\overline{\sigma}_{AB}^* \equiv \int \frac{d\theta}{2\pi} \hat{J}_\theta \sigma_{AB}^* \hat{J}_\theta^\dagger, \quad (4.74)$$

then  $\overline{\sigma}_{AB}^*$  is separable and

$$D_{\max}(\pi_{\mathcal{N}}||\sigma_{AB}^*) = D_{\max}(\pi_{\mathcal{N}}||\overline{\sigma}_{AB}^*). \quad (4.75)$$

Similarly, if  $\sigma_{AB}^*$  is the state which minimises  $D_{\max}(\pi_{\mathcal{N}}||\sigma_{AB})$  over all separable states  $\sigma_{AB}$  with  $\text{Tr}_B[\sigma_{AB}] = \mathbb{1}_A/2$ , the same conclusion holds with  $\text{Tr}_B[\overline{\sigma}_{AB}^*] = \mathbb{1}_A/2$ .

As a corollary of Lemma 4.1, in Equation (4.66) instead of minimising over all separable states  $\sigma_{AB}$ , we can equivalently consider only the separable states which are left unaltered when averaged over all possible  $\theta$  rotations. The density matrices associated with these states in the basis  $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$  can be parametrised as

$$\sigma_{AB} = \frac{1}{2} \begin{pmatrix} \alpha & & \xi e^{i\phi} \\ & \gamma & \\ \xi e^{-i\phi} & & \beta \end{pmatrix}, \quad (4.76)$$

with  $\alpha, \beta, \gamma, \delta, \xi \geq 0$ ,  $\alpha + \beta + \gamma + \delta = 2$ ,  $\phi \in [0, 2\pi]$  and  $0 \leq \xi \leq \min\{\sqrt{\alpha\beta}, \sqrt{\gamma\delta}\}$ . Notice that the last inequality comes from the positive partial transpose (PPT) criterion, which for two-qubit states is a necessary and sufficient condition for separability [160]. For the optimisation leading to the upper bound in Proposition 4.1, we simply need to add the additional constraints  $\gamma = 1 - \alpha$  and  $\delta = 1 - \beta$ , in order to assure  $\text{Tr}_B[\sigma_{AB}] = \mathbb{1}_A/2$ .

*Proof of Lemma 4.1.* As mentioned in Section 4.2.1, the max-relative entropy  $D_{\max}(\rho||\sigma)$  is jointly quasi-convex and invariant under joint unitary operations. These facts, together with Equation (4.73), lead to the following inequality:

$$\begin{aligned} D_{\max}\left(\pi_{\mathcal{N}_\lambda^{(\text{ad})}}\left|\left|\overline{\sigma}_{AB}^*\right.\right.\right) &= D_{\max}\left(\int \frac{d\theta}{2\pi} \hat{J}_\theta \pi_{\mathcal{N}_\lambda^{(\text{ad})}} \hat{J}_\theta^\dagger \left|\left|\int \frac{d\theta}{2\pi} \hat{J}_\theta \sigma_{AB}^* \hat{J}_\theta^\dagger\right.\right.\right) \\ &\leq \max_{\theta} D_{\max}(\hat{J}_\theta \pi_{\mathcal{N}_\lambda^{(\text{ad})}} \hat{J}_\theta^\dagger || \hat{J}_\theta \sigma_{AB}^* \hat{J}_\theta^\dagger) = D_{\max}(\pi_{\mathcal{N}_\lambda^{(\text{ad})}} || \sigma_{AB}^*). \end{aligned} \quad (4.77)$$

The converse inequality follows because  $\overline{\sigma}_{AB}^*$  is separable, as can be easily seen from the structure of  $U_\theta$  in Equation (4.72), and because  $\sigma_{AB}^*$  minimises  $D_{\max}(\pi_{\mathcal{N}} || \sigma_{AB})$  over all separable states. The final remark can be easily proven by noticing that  $\text{Tr}_B [\overline{\sigma}_{AB}^*] = \mathbb{1}_A/2$  if  $\text{Tr}_B [\sigma_{AB}^*] = \mathbb{1}_A/2$ .  $\square$

## 4.5.2 Evaluation on qubit channels

We consider four examples of qubit channels invariant under phase rotations: amplitude damping, dephasing, depolarising, and erasure channel. The first one is not Choi-stretchable, so we can only compute the upper and lower bounds on its max-relative entropy of entanglement found in Proposition 4.1. The other three channels instead are Choi-stretchable, and their max-relative entropy of entanglement can be computed numerically via Proposition 4.1.

### Amplitude Damping Channel

We begin by studying the most important example among channels that are not Choi-stretchable: the qubit amplitude damping channel  $\mathcal{N}_\lambda^{(\text{ad})}$ . This can be written as

$$\mathcal{N}_\lambda^{(\text{ad})}(\rho) = \sum_{i=1}^2 M_i(\mathcal{N}_\lambda^{(\text{ad})}) \rho M_i^\dagger(\mathcal{N}_\lambda^{(\text{ad})}), \quad (4.78)$$

in terms of the Kraus operators

$$M_1(\mathcal{N}_\lambda^{(\text{ad})}) = |0\rangle\langle 0| + \sqrt{1-\lambda}|1\rangle\langle 1|, \quad M_2(\mathcal{N}_\lambda^{(\text{ad})}) = \sqrt{\lambda}|0\rangle\langle 1|. \quad (4.79)$$

Note that  $\mathcal{N}_\lambda^{(\text{ad})}$  reduces to the identity channel when  $\lambda = 0$ , whereas any input state is sent to  $|0\rangle\langle 0|$  when  $\lambda = 1$ . By exploiting the results of the previous section, we can analytically calculate the lower and upper bounds on  $E_{\max}(\mathcal{N}_\lambda^{(\text{ad})})$  found in Proposition 4.1. The proofs are quite long and technical, but the interested reader can find them respectively in Appendix B.2 and B.3.

**Proposition 4.2.** *The quantity  $E_{\max}(\mathcal{N}_\lambda^{(\text{ad})})$  can be bounded on both sides by:*

$$F(\lambda) \leq E_{\max}(\mathcal{N}_\lambda^{(\text{ad})}) \leq \tilde{E}_{\max}(\mathcal{N}_\lambda^{(\text{ad})}), \quad (4.80)$$

where

$$\tilde{E}_{\max}(\mathcal{N}_\lambda^{(\text{ad})}) \equiv \log_2(2 - \lambda), \quad F(\lambda) \equiv \begin{cases} \log_2 \left[ \frac{1}{2}(1 + \sqrt{1 - \lambda})^2 \right], & \text{if } \lambda \leq \frac{\sqrt{5}-1}{2}, \\ \log_2 \left( \frac{1+\lambda}{2\lambda} \right), & \text{if } \lambda \geq \frac{\sqrt{5}-1}{2}. \end{cases} \quad (4.81)$$

We stress that this upper bound  $\tilde{E}_{\max}(\mathcal{N}_\lambda^{(\text{ad})})$  is also an upper bound on the capacity  $C(\mathcal{N}_\lambda^{(\text{ad})})$ , whereas the lower bound does not have any known relation with the latter. However, in the region of parameters  $\lambda$  where this lower bound is larger than the best known upper bound on the squashed entanglement of the channel, we are guaranteed that  $E_{\text{sq}}(\mathcal{N}_\lambda^{(\text{ad})}) < E_{\max}(\mathcal{N}_\lambda^{(\text{ad})})$ . The plot in Figure 4.3 shows how these two bounds can be compared with other quantities known in the literature. The function  $\tilde{E}_{\max}(\mathcal{N}_\lambda^{(\text{ad})})$  is plotted as a black curve, and is an upper bound on  $E_{\max}(\mathcal{N}_\lambda^{(\text{ad})})$  and thus also on  $C(\mathcal{N}_\lambda^{(\text{ad})})$ . We can see that it is much smaller than the upper bound on the capacity obtained in Ref. [166], represented by the top blue solid curve in Figure 4.3. However, the upper bounds on the channel capacity based on the squashed entanglement [166, 168] are smaller than our result obtained through  $E_{\max}$ . In particular, the tightest among these has been plotted as a red solid curve in Figure 4.3, and remains below the lower bound on  $E_{\max}(\mathcal{N}_\lambda^{(\text{ad})})$  found in Proposition 4.2 (green dotted line in Figure 4.3) for almost any values of  $\lambda$ . By inspection, we see that the green line crosses the red line near  $\lambda \sim 0.88$ . Therefore, only when  $\lambda \gtrsim 0.88$  it might still be possible that  $C(\mathcal{N}^{(\text{ad})}) \leq E_{\max}(\mathcal{N}^{(\text{ad})}) < E_{\text{sq}}(\mathcal{N}^{(\text{ad})})$ , although this seems to be highly unlikely. For completeness, we also plotted the best known lower bound on  $C(\mathcal{N}_\lambda^{(\text{ad})})$ , which narrows the region where the capacity value could be [166, 168]. From this analysis, we can conclude that at the moment the best known upper bound on the capacity of the amplitude damping channel remains based on its squashed entanglement.

### Three Choi-stretchable channels

We numerically evaluate the value of the max-relative entropy of entanglement for three Choi-stretchable qubit channels, by computing the lower bound in Proposition 4.1 with techniques similar to those used for the amplitude damping channel. The dephasing channel  $\mathcal{N}_\lambda^{(\text{deph})}$  and depolarising channel  $\mathcal{N}_\lambda^{(\text{depo})}$  can be respectively written in terms of a set of two and five Kraus operators:

$$M_1(\mathcal{N}_\lambda^{(\text{deph})}) = \sqrt{1 - \frac{\lambda}{2}} \hat{1}, \quad M_2(\mathcal{N}_\lambda^{(\text{deph})}) = \sqrt{\frac{\lambda}{2}} \hat{\sigma}_z, \quad (4.82)$$

$$M_0(\mathcal{N}_\lambda^{(\text{depo})}) = \sqrt{1 - \lambda} \hat{1}, \quad M_{ij}(\mathcal{N}_\lambda^{(\text{depo})}) = \sqrt{\frac{\lambda}{2}} |i\rangle \langle j|, \quad (4.83)$$

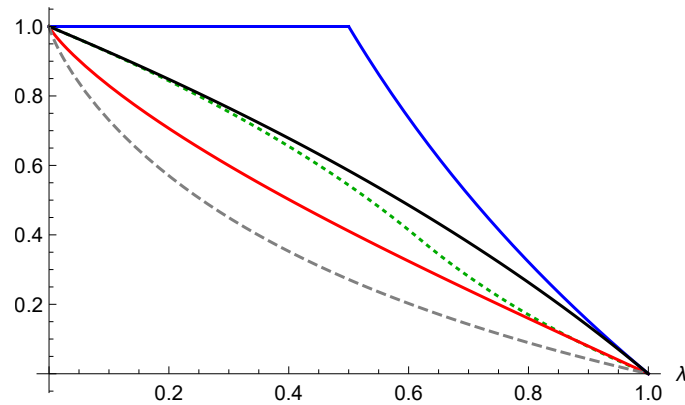


Figure 4.3: Grey dashed line: Best known lower bound on the capacity  $C(\mathcal{N}^{(\text{ad})})$ , corresponding to the reverse coherent information of the channel (see [166]). Green dotted line: lower bound on  $E_{\max}(\mathcal{N}^{(\text{ad})})$  found in Proposition 4.2. The remaining solid lines are all upper bounds on the capacity  $C(\mathcal{N}^{(\text{ad})})$ . In particular, the top blue line is the bound based on the relative entropy of entanglement discussed in Ref. [166], the bottom red line is the best known bound based on the squashed entanglement [166], and the black line in the middle is  $\tilde{E}_{\max}(\mathcal{N}^{(\text{ad})})$  found in Proposition 4.2.

with  $i, j = 0, 1$ . The erasure channel  $\mathcal{N}^{(\text{er})}$ , instead, is characterised by the Kraus operators

$$M_2(\mathcal{N}_\lambda^{(\text{er})}) = \sqrt{1-\lambda}\hat{1}, \quad M_i(\mathcal{N}_\lambda^{(\text{er})}) = \sqrt{\lambda}|e\rangle\langle i|, \quad (4.84)$$

where  $i = 0, 1$ , and  $|e\rangle$  is an error state orthogonal to both  $|0\rangle$  and  $|1\rangle$ . All these channels reduce to the identity channel when  $\lambda = 0$ . Moreover, the dephasing and erasure channels are distillable, and their relative entropy of entanglement coincides with their capacity.

We point out that exact values for the max-relative entropy of entanglement of these channels are not needed when applying the versatile network bound of Theorem 4.3. This is because all these channels are Choi-stretchable, and as such it is always convenient to use  $E_R$  in the bound rather than the larger measure  $E_{\max}$ . Nonetheless, we numerically evaluated the max-relative entropy of entanglement of these channels in order to see whether in some cases it could be smaller than its counterpart based on the squashed entanglement. The results can be seen in Figure 4.4: for all three channels  $E_R$  yields the tightest upper bound on the capacity, followed by the squashed entanglement, while  $E_{\max}$  provides the loosest bound.

## 4.6 Examples

By studying the sign of the parameter  $\mu_{c_A}$  introduced in Equation (4.64), or its approximation  $\tilde{\mu}_{c_A}$  written in terms of the tightest known upper bounds on the involved quantities, we can assess whether for a certain bipartition of the network Theorem 4.3 leads to a tighter bound than the version of Theorem 4.2 based on the squashed entan-

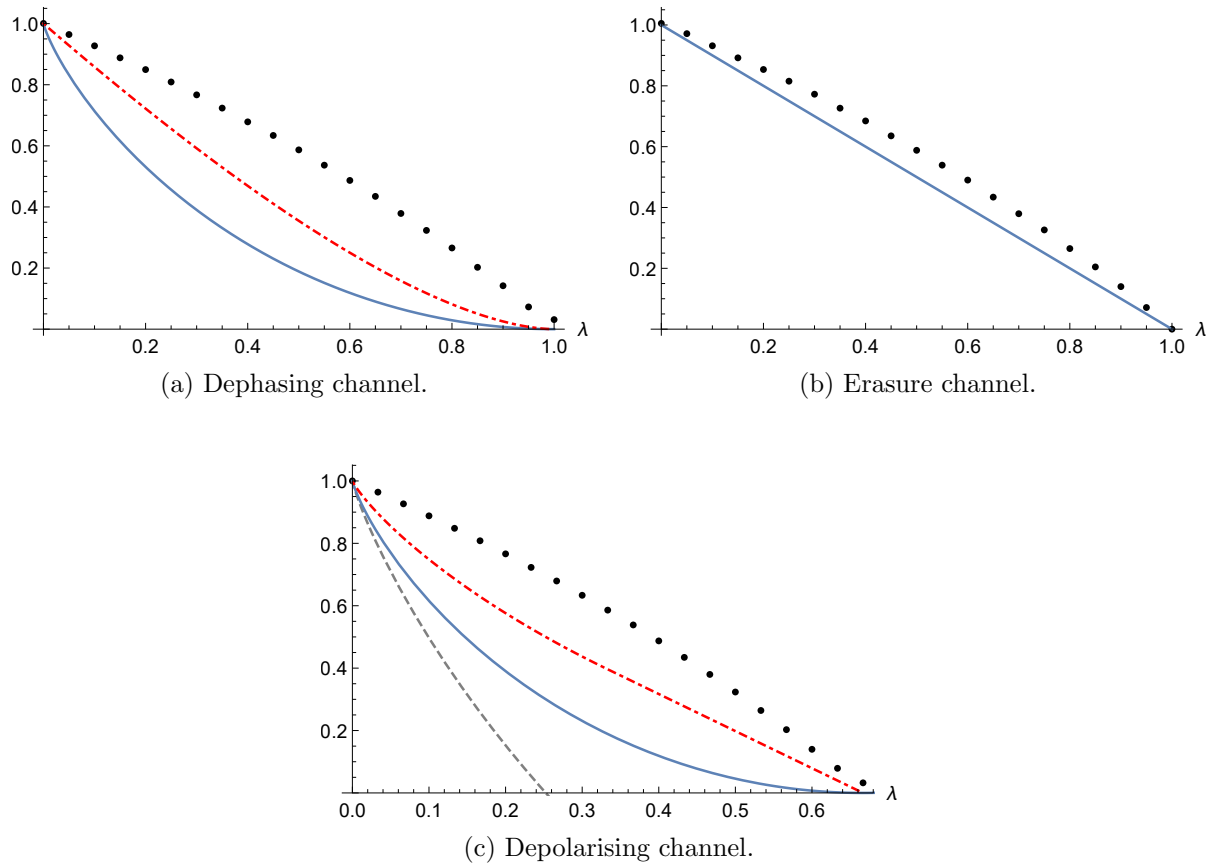


Figure 4.4: Bounds on the capacity of three Choi-stretchable qubit channels. In each plot, the solid blue line represents  $E_R(\mathcal{N}_\lambda)$ , and coincides with the capacity  $C(\mathcal{N}_\lambda)$  for the dephasing and erasure channel. For the depolarising channel, the capacity  $C(\mathcal{N}_\lambda^{(\text{depo})})$  lies between the blue solid line and the gray dashed line, which respectively represent its best known upper and lower bounds (see, e.g., Ref. [166]). The depolarising channel has zero capacity for  $\lambda > 2/3$ , where it becomes entanglement breaking, so that region has not been plotted. Red dot-dashed line: smaller known upper bound on the squashed entanglement of the channel [165, 168]. In the specific case of the erasure channel, one has  $E_{\text{sq}}(\mathcal{N}_\lambda^{(\text{er})}) = E_R(\mathcal{N}_\lambda^{(\text{er})})$  [166, 168]. Black dots: numerical evaluation of  $E_{\text{max}}(\mathcal{N}_\lambda)$ , obtained by performing the minimisations over the set of separable states given in Equation (4.76).

lement In what follows we provide two examples where these parameters are positive, and our versatile bound results advantageous.

### 4.6.1 Bipartition composed by flower and distillable channels

A first example is based on the so-called “flower channels” [194, 195], which are quantum channels cleverly designed so that the gap between  $E_{\text{sq}}(\mathcal{N})$  and  $E_{\text{max}}(\mathcal{N})$  can be made arbitrarily large, depending on the dimension of the input system [167]. As a consequence, our versatile upper bound  $\mathcal{B}'$  is much smaller, and thus much better, than  $\mathcal{B}_{E_{\text{sq}}}$  when evaluated on bipartitions composed by flower channels and other Choi-stretchable channels  $\mathcal{N}$  with  $E_{\text{sq}}(\mathcal{N}) \geq E_{\text{R}}(\mathcal{N})$  [e.g., distillable channels, for which  $E_{\text{R}}(\mathcal{N}) = C(\mathcal{N})$ ].

In what follows we discuss this example in more detail, by explicitly showing that the squashed entanglement of flower channels can be arbitrarily larger than their max-relative entropy of entanglement. We will follow Ref. [167].

For any  $d \in \mathbb{N}$ , we can consider the flower state

$$\rho_{AA'BB'}^{(f)} = \frac{1}{2d} \sum_{i,k=1}^d \sum_{j,l=1}^2 \langle k | \hat{U}_i^\dagger \hat{U}_j | i \rangle | ii \rangle_{AB} \langle kk | \otimes | jj \rangle_{A'B'} \langle ll |, \quad (4.85)$$

where  $\hat{U}_1 = \hat{1}_d$  and  $\hat{U}_2$  is the quantum Fourier transform operator

$$\hat{U}_2 = \sum_{j,k=1}^d \frac{1}{\sqrt{d}} e^{i \frac{2\pi jk}{d}} |j\rangle \langle k|. \quad (4.86)$$

The flower channel  $\mathcal{T}_{\tilde{A}\tilde{A}' \rightarrow BB'}^{(\rho^{(f)})}$  is then defined as the unital CPT map associated with the flower state in Equation (4.85) via the Choi-Jamiołkowski isomorphism (notice that both the reduced density matrices  $\rho_{AA'}^{(f)}$  and  $\rho_{BB'}^{(f)}$  are maximally mixed):

$$\mathcal{T}_{\tilde{A}\tilde{A}' \rightarrow BB'}^{(\rho^{(f)})}[\tau_{\tilde{A}\tilde{A}'}] = (d_A d_{A'})^2 \text{Tr}_{AA'\tilde{A}\tilde{A}'} \left[ \psi_{AA'\tilde{A}\tilde{A}'} \left( \tau_{\tilde{A}\tilde{A}'} \otimes \rho_{AA'BB'}^{(f)} \right) \right], \quad (4.87)$$

$$\rho_{AA'BB'}^{(f)} = \mathcal{T}_{\tilde{A}\tilde{A}' \rightarrow BB'}^{(\rho^{(f)})}[\psi_{AA'\tilde{A}\tilde{A}'}], \quad (4.88)$$

as in the proof of Proposition 4.1, where  $\psi_{AA'\tilde{A}\tilde{A}'}$  is a maximally entangled state across the bipartition  $AA' : \tilde{A}\tilde{A}'$ . The squashed entanglement of the flower state across the bipartition  $AA' : BB'$  has been computed exactly [194, 195], and it is given by

$$E_{\text{sq}}(\rho_{AA'BB'}^{(f)}) = 2 + \frac{1}{2} \log_2 d, \quad (4.89)$$

quantity which can be made arbitrarily large simply by increasing the dimension  $d$  of subsystems  $A$  and  $B$ . As a consequence, the squashed entanglement of the flower channel is larger than this quantity, because the maximally entangled state in Equation (4.88) is



not necessarily the state leading to the maximum amount of entanglement between  $AA'$  and  $BB'$ :

$$E_{\text{sq}} \left[ \mathcal{T}_{\tilde{A}\tilde{A}' \rightarrow BB'}^{(\rho^{(f)})} \right] = \max_{\rho_{AA'\tilde{A}\tilde{A}'}} E_{\text{sq}} \left[ \mathcal{T}_{\tilde{A}\tilde{A}' \rightarrow BB'}^{(\rho^{(f)})} [\rho_{AA'\tilde{A}\tilde{A}'}] \right] \geq E_{\text{sq}} \left[ \rho_{AA'BB'}^{(f)} \right] = 1 + \frac{1}{2} \log_2 d. \quad (4.90)$$

If we now move to the max-relative entropy of entanglement, we can exploit the fact that it is “non-lockable” [167], i.e., such that for any state  $\sigma_{AA'BB'}$  one has

$$E_{\text{max}}(\sigma_{AA'BB'}) \leq E_{\text{max}}(\sigma_{AA'B}) + 2 \log_2(d_{B'}), \quad (4.91)$$

where  $\sigma_{AA'B} = \text{Tr}_{B'} [\sigma_{AA'BB'}]$  and  $d_{B'}$  is the dimension of subsystem  $B'$ . In other words, partial trace can change the entanglement content of a state by a value which is at most twice the logarithm of the dimension of the system that has been traced away. As a corollary [167], it follows that the max-relative entropies of entanglement of two channels  $\mathcal{T}_{\tilde{A}\tilde{A}' \rightarrow BB'}$  and  $\mathcal{T}_{\tilde{A}\tilde{A}' \rightarrow B} \equiv \text{Tr}_{B'} [\mathcal{T}_{\tilde{A}\tilde{A}' \rightarrow BB'}]$  are related by

$$E_{\text{max}}(\mathcal{T}_{\tilde{A}\tilde{A}' \rightarrow BB'}) \leq E_{\text{max}}(\mathcal{T}_{\tilde{A}\tilde{A}' \rightarrow B}) + 2 \log_2(d_{B'}). \quad (4.92)$$

This can be easily shown by choosing as  $\sigma_{AA'BB'}$  in Equation (4.91) the output state of  $\mathcal{T}_{\tilde{A}\tilde{A}' \rightarrow BB'}$  with the maximum entanglement. We can now apply this result to the flower channel. Note that by tracing over subsystem  $B'$  in Equation (4.85) we are left with

$$\rho_{AA'B}^{(f)} = \text{Tr}_{B'} \left[ \rho_{AA'BB'}^{(f)} \right] = \frac{1}{2d} \sum_i^d \sum_j^2 |ii\rangle_{AB} \langle ii| \otimes |j\rangle_{A'} \langle j|, \quad (4.93)$$

which is separable. Therefore, the reduced channel  $\mathcal{T}_{\tilde{A}\tilde{A}' \rightarrow B}^{(\rho^{(f)})}$  is entanglement breaking, and by using  $d_{B'} = 2$  Equation (4.92) yields [167]

$$E_{\text{max}} \left[ \mathcal{T}_{AA' \rightarrow BB'}^{(\rho^{(f)})} \right] \leq E_{\text{max}} \left[ \mathcal{T}_{AA' \rightarrow B}^{(\rho^{(f)})} \right] + 2 \log_2(d'_B) = 2, \quad (4.94)$$

which does not depend on  $d$ .

Overall, by combining Equations (4.90) and (4.94) we obtain

$$E_{\text{sq}} \left[ \mathcal{T}_{AA' \rightarrow BB'}^{(\rho^{(f)})} \right] - E_{\text{max}} \left[ \mathcal{T}_{AA' \rightarrow BB'}^{(\rho^{(f)})} \right] \geq \frac{1}{2} \log_2 d - 1, \quad (4.95)$$

which is positive when  $d \geq 4$ , and can become arbitrarily large if  $d$  is increased. As our versatile bound uses  $E_{\text{max}}$  rather than  $E_{\text{sq}}$ , it performs much better than previously known bounds when the dimension  $d$  of the flower channels connecting the two sides of the network bipartition becomes large.

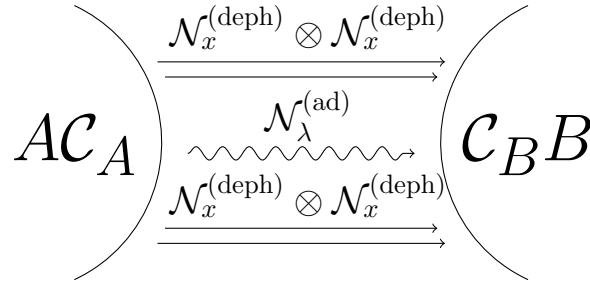


Figure 4.5: Example of bipartition  $AC_A : C_BB$  connected by  $k = 4$  dephasing channels (straight lines) and 1 amplitude damping channel (wiggling line). Once a bipartition of the network has been selected, it is not necessary to keep track of the precise nodes connected by the channels in order to apply Theorem 4.3.

#### 4.6.2 Bipartition composed by amplitude damping and dephasing channels

It could be argued that a flower channel is rather artificial, and it is not likely to appear in any realistic communication scenario. For this reason, we also consider a different example in which the two components of a bipartition  $AC_A : C_BB$  are connected by  $k$  dephasing channels  $\mathcal{N}_x^{(\text{deph})}$  and 1 amplitude damping channel  $\mathcal{N}_\lambda^{(\text{ad})}$ , as shown in Figure 4.5.

For simplicity we assume that each channel is used, on average, the same number of times. Moreover, we will compute  $\tilde{\mu}_{C_A}$  rather than  $\mu_{C_A}$  because the exact values of many quantities appearing in  $\mu_{C_A}$  are not known. In particular, we can write:

$$\tilde{\mu}_{C_A} = \frac{k[\tilde{E}_{\text{sq}}(\mathcal{N}_x^{(\text{deph})}) - E_{\text{R}}(\mathcal{N}_x^{(\text{deph})})] + [\tilde{E}_{\text{sq}}(\mathcal{N}_\lambda^{(\text{ad})}) - \tilde{E}_{\text{max}}(\mathcal{N}_\lambda^{(\text{ad})})]}{k[\tilde{E}_{\text{sq}}(\mathcal{N}_x^{(\text{deph})}) + E_{\text{R}}(\mathcal{N}_x^{(\text{deph})})] + [\tilde{E}_{\text{sq}}(\mathcal{N}_\lambda^{(\text{ad})}) + \tilde{E}_{\text{max}}(\mathcal{N}_\lambda^{(\text{ad})})]}, \quad (4.96)$$

where  $\tilde{E}_{\text{sq}}(\mathcal{N}_x^{(\text{deph})})$  and  $\tilde{E}_{\text{sq}}(\mathcal{N}_\lambda^{(\text{ad})})$  are the best known upper bounds respectively on  $E_{\text{sq}}(\mathcal{N}_x^{(\text{deph})})$  and  $E_{\text{sq}}(\mathcal{N}_\lambda^{(\text{ad})})$  [165, 166]. These upper bounds have been plotted as red dot-dashed curves in Figures 4.4a and 4.3, and can be explicitly written as

$$\tilde{E}_{\text{sq}}(\mathcal{N}_x^{(\text{deph})}) = h\left(\sqrt{\frac{x}{2}\left(1 - \frac{x}{2}\right)} + \frac{1}{2}\right), \quad (4.97)$$

$$\tilde{E}_{\text{sq}}(\mathcal{N}_\lambda^{(\text{ad})}) = h\left(\frac{1}{2} - \frac{\lambda}{4}\right) - h\left(1 - \frac{\lambda}{4}\right), \quad (4.98)$$

with  $h(y) \equiv -y \log y - (1 - y) \log(1 - y)$ . The quantity  $E_{\text{R}}(\mathcal{N}_x^{(\text{deph})})$  is known to be equal to  $1 - h(x/2)$  [166], and the upper bound  $\tilde{E}_{\text{max}}(\mathcal{N}_\lambda^{(\text{ad})})$  on  $E_{\text{max}}(\mathcal{N}_\lambda^{(\text{ad})})$  has been obtained in Proposition 4.2. In this way, we can write  $\tilde{\mu}_{C_A}$  as a function of the number of dephasing channels  $k$  and of the parameters  $x, \lambda \in [0, 1]$ .

In Figure 4.6 we plotted the results obtained for  $\tilde{\mu}_{C_A}$  when  $k = 1, 5, 10$ , as function of  $x$  and  $\lambda$ . As expected, we can see that the region in which our versatile bound is

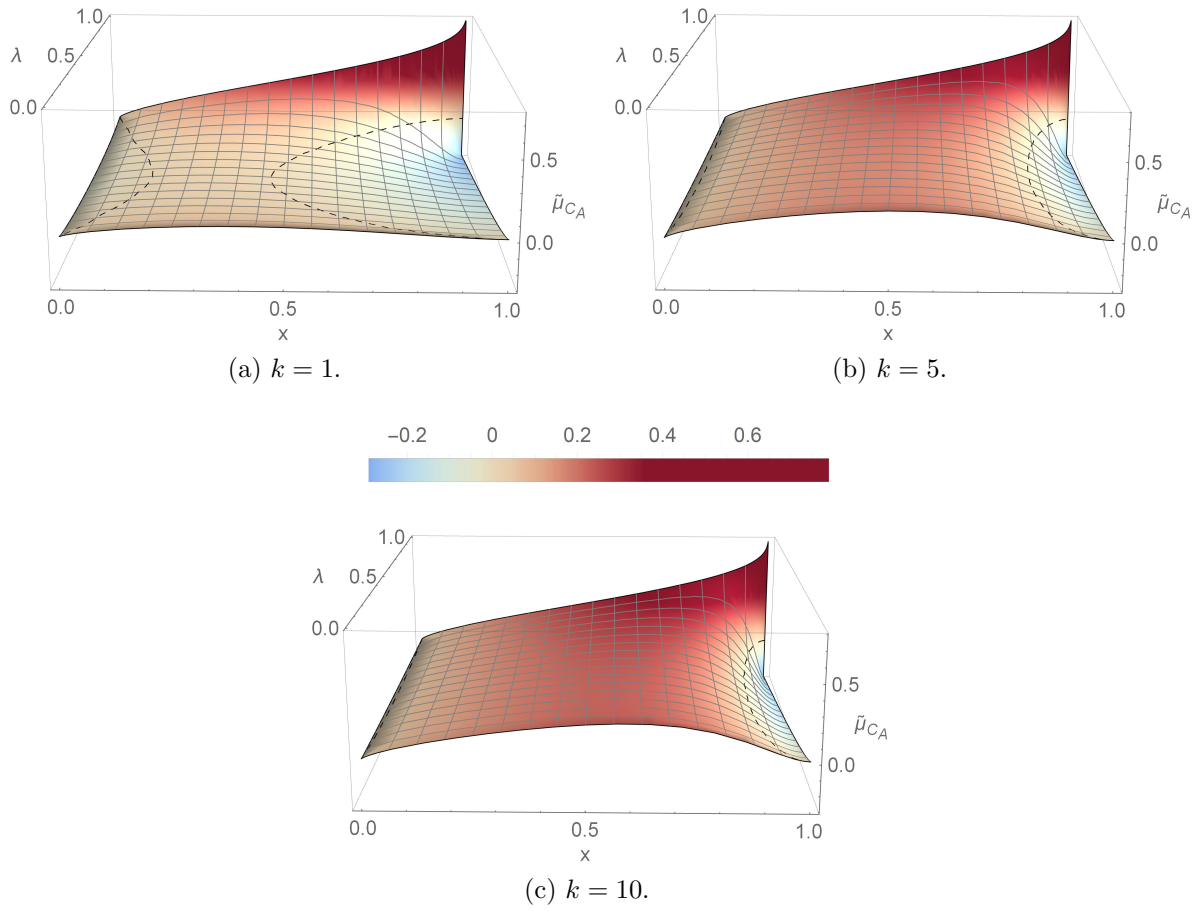


Figure 4.6: Relative advantage of the versatile upper bound  $\mathcal{B}'$  over the upper bound  $\mathcal{B}_{E_{\text{sq}}}$  based on the squashed entanglement, as measured by the parameter  $\tilde{\mu}_{C_A}$ , for a bipartition of the network whose components are connected by  $k$  dephasing channels  $\mathcal{N}_x^{(\text{deph})}$  and 1 amplitude damping channel  $\mathcal{N}_\lambda^{(\text{ad})}$ . The set of points characterised by  $\tilde{\mu}_{C_A} = 0$  is highlighted on the plots by dashed black curves.

advantageous (i.e. such that  $\tilde{\mu}_{C_A} > 0$ ) becomes larger when  $k$  increases. Nonetheless, even for  $k=1$ , for a broad set of parameters the versatile bound performs better than the bound based on the squashed entanglement. In particular, this happens for  $\lambda \simeq 1$ , where the negative contribution in  $\tilde{\mu}_{C_A}$  coming from the fact that  $\tilde{E}_{\text{max}}(\mathcal{N}_\lambda^{(\text{ad})}) \geq \tilde{E}_{\text{sq}}(\mathcal{N}_\lambda^{(\text{ad})})$  disappears, because both those quantities become equal to zero. On the contrary, the bound based on the squashed entanglement is preferable when  $x \simeq 1$ , because when this is the case there is no advantage in using  $E_R$  rather than  $E_{\text{sq}}$  to quantify the maximum amount of entanglement that could be produced by  $\mathcal{N}_x^{(\text{deph})}$ : in both cases this is almost zero. The peak in  $\tilde{\mu}_{C_A}$  observed for  $x, \lambda \rightarrow 1$ , instead, is due to the fact that in this regime the denominator of  $\tilde{\mu}_{C_A}$  goes to zero, and even small differences in performance are amplified.

## 4.7 Conclusions

In quantum communication, one of the most important research topics aims at quantifying the ultimate transmission rate achievable over a given set of quantum channels. When the possibility of exchanging public classical information is freely allowed to the parties, a successful approach to this problem is based on the quantification of the maximum amount of entanglement generated by the considered channels. Instead of focussing on a single entanglement measure, in this chapter we investigated the possibility of integrating multiple measures into a single approach, with the goal of writing a tight and versatile upper bound on the number of ebits (or pbits) shared by two parties at the end of a communication protocol over a generic quantum network. This has been possible thanks to the special relation between the relative entropy and the max-relative entropy of entanglement. With our result, it is possible to take advantage from the presence of Choi-stretchable channels in the network, without having to require this property beforehand.

We provided explicit examples where our versatile upper bound performs better than the previously known approach uniquely based on the squashed entanglement. Although in general one should check on a case-by-case basis which is the upper bound yielding the tightest result, we can expect the versatile bound introduced in Theorem 4.3 to be the best choice when the majority of the channels in the network are Choi-stretchable and such that  $E_R(\mathcal{N}) \leq E_{sq}(\mathcal{N})$ . This intuition was confirmed for a network composed by  $k$  dephasing channels and one amplitude damping channel, in which already for  $k = 5$  our versatile bound performed better on a broad range of parameters. We also reiterate that, according to the authors of Ref. [167], Equation (4.57) has been rigorously proven only for channels acting on finite-dimensional systems. As Theorem 4.3 heavily relies upon that inequality, special attention is required when that result is applied to channels receiving infinite-dimensional systems in input, at least as long as the proof of Equation (4.57) will not be suitably extended. However, we point out that at least some common bosonic channels (e.g., optical photon losses) are Choi-stretchable: in these cases we can safely utilise Equation (4.22) and bound their performances via  $E_R$  [166], so that Theorem 4.3 still holds.

The advantage of our approach would be further increased if more entanglement measures could be included within the same framework. For example, it would be interesting to include the squashed entanglement, because it typically provides tighter upper bounds on single-channel capacities than  $E_{max}$  for channels that are not Choi-stretchable. Furthermore, it would be interesting to check whether other entanglement measures could provide upper bounds on single-channel capacities. From this point of view, we feel that the schematic framework provided by Theorems 4.1 and 4.2 could act as a guideline for future investigations. Another idea worth considering concerns the possibility of extending this “versatile” approach to a multi-user scenario, in which the network is composed by broadcast quantum channels [187, 196–200].

# Chapter 5

## A global perspective

With this thesis we have provided original contributions to different aspects of quantum information theory, from the detection and quantification of nonclassicality to the investigation of the ultimate bounds that quantum mechanics enforces on communication rates. In this conclusive overview, we would like to comment on what we think are the main lessons and ideas that can be learned from each chapter. We will also try to highlight the limitations of our studies, with the hope of inspiring further research.

We feel that two main messages can be learned from Chapter 2. On the one hand, one should keep in mind that any quantum optical setting has a natural classical counterpart, which could be obtained by studying an analogous problem of classical electrodynamics. Often this aspect is overlooked in favour of another classical framework in which the photons are perfectly distinguishable, thus retaining the particle-like behaviour of light rather than its wave-like nature. The two pictures can be expected to provide different and complementary perspectives on light evolution, and can both be used in the attempt of interpreting or benchmarking the predictions of quantum mechanics. On the other hand, the technical proofs provided in the chapter highlight how the language of quantum information can be used to rewrite and solve linear algebraic problems originating from completely different contexts. It is possible that our results could be used in the future in order to better characterise experimental setups, or to certify a nonclassical behaviour in optical devices. However, at the moment our approach cannot be directly connected with other important research topics, as boson sampling or the theory of computational complexity, and it would be interesting to explore whether it is possible to use similar ideas to advance in those directions.

When studying or characterising a quantum system, it is often impossible to have perfect control over its evolution. However, it is reasonable to expect that some of the details affecting its dynamics could be deduced a-posteriori, without destroying the quantum state. Chapter 3 deals with this kind of framework, in a situation where the quantum state undergoing the evolution is used to deduce information on the evolution itself. The main contribution of the chapter consists in the introduction of a natural and computable

measure, which can be used to quantify the best average estimation precision obtainable when using a given probing system. On the other hand, the main limitation of this approach lies in the fact that it does not take into account whether the optimal read-out measurement could be actually implemented or not. This is a common problem in quantum metrology, where the possibility of reaching the ultimate precision allowed by quantum mechanics should be checked on a case-by-case basis. Unfortunately, the limitations on the experimentally available measurement schemes are of practical nature, and as such cannot be easily taken into account.

In Chapter 4 we considered a generic network of quantum channels, scenario that will become more and more important in the future. Despite the variety of situations that could possibly arise in this context, at the moment most of the theoretical efforts aiming at characterising the communication performance of this sort of networks have been largely inspired by previous strategies developed for single channels. It is important to go beyond this paradigm, and to draw expertise from classical graph theory in order to better address the challenges, and take advantage of the opportunities, that this developing field offers. With Chapter 4 we took a small step in this direction, by realising that diverse tools can be used to characterise different network components, and that combining together multiple approaches can often be beneficial.

If we consider the field of quantum information as a whole, some of the most mature and promising research topics are those studying quantum computation or simulation, quantum enhanced metrology, or quantum communication. The last two topics are probably those expected to yield technological applications in the near future (for example, think at the proposal of using squeezed light for gravitational wave detection, or at quantum key distribution). Instead, fully functional quantum computers are probably not going to appear within a few years, and quantum simulation is considered a temporary application for devices with limited capabilities. With this structure in mind, Chapters 3 and 4 naturally contribute, respectively, to quantum metrology and communication. The nonclassicality bound of Chapter 2, instead, could be used as a tool to certify “exotic” quantum behaviours in optical devices, and could play a role in the development of increasingly powerful quantum technologies. In this exciting time of change and transformation, quantum information has great impact potential, and it is certainly an area that still offers many research opportunities.

# Bibliography

- [1] C. E. Shannon, “A mathematical theory of communication”, *Bell Syst. Tech. J.*, 27, pp. 623–656, 1948.
- [2] H. J. Kimble, “The quantum internet”, *Nature*, 453, pp. 1023–1030, 2008.
- [3] N. Sangouard, C. Simon, H. de Riedmatten, and N. Gisin, “Quantum repeaters based on atomic ensembles and linear optics”, *Rev. Mod. Phys.*, 83, pp. 33–80, 2011.
- [4] M. Peev *et al.*, “The SECOQC quantum key distribution network in Vienna”, *New J. Phys.*, 11, p. 075001, 2009.
- [5] D. Stucki *et al.*, “Long-term performance of the swissquantum quantum key distribution network in a field environment”, *New J. Phys.*, 13, p. 123001, 2011.
- [6] M. Sasaki *et al.*, “Field test of quantum key distribution in the Tokyo QKD network”, *Opt. Express*, 19, pp. 10 387–10 409, 2011.
- [7] W. J. Munro, K. Azuma, K. Tamaki, and K. Nemoto, “Inside quantum repeaters”, *IEEE J. Sel. Top. Quantum Electron.*, 21, pp. 78–90, 2015.
- [8] M. M. Wilde, *Quantum Information Theory*, 1st ed. New York, NY, USA: Cambridge University Press, 2013.
- [9] L. Rigovacca, C. Di Franco, B. J. Metcalf, I. A. Walmsley, and M. S. Kim, “Non-classicality criteria in multiport interferometry”, *Phys. Rev. Lett.*, 117, p. 213602, 2016.
- [10] L. Rigovacca, A. Farace, L. A. M. Souza, A. De Pasquale, V. Giovannetti, and G. Adesso, “Versatile Gaussian probes for squeezing estimation”, *Phys. Rev. A*, 95, p. 052331, 2017.
- [11] L. Rigovacca, G. Kato, S. Bäuml, M. S. Kim, W. J. Munro, and K. Azuma, “Versatile relative entropy bounds for quantum networks”, 2017, arXiv:1707.05543 [quant-ph].
- [12] T. Young, *A course of lectures on natural philosophy and the mechanical arts*. London: Printed for J. Johnson, 1807, vol. 2.
- [13] A. A. Michelson and E. W. Morley, “On the relative motion of the earth and the luminiferous ether”, *Am. J. Sci.*, Series 3 Vol. 34, pp. 333–345, 1887.
- [14] C. K. Hong, Z. Y. Ou, and L. Mandel, “Measurement of subpicosecond time intervals between two photons by interference”, *Phys. Rev. Lett.*, 59, pp. 2044–2046, 1987.

- [15] F. W. Sun and C. W. Wong, “Indistinguishability of independent single photons”, *Phys. Rev. A*, 79, p. 013824, 2009.
- [16] E. Knill, R. Laflamme, and G. J. Milburn, “A scheme for efficient quantum computation with linear optics”, *Nature*, 409, pp. 46–52, 2001.
- [17] B. C. Sanders, “Quantum dynamics of the nonlinear rotator and the effects of continual spin measurement”, *Phys. Rev. A*, 40, pp. 2417–2427, 1989.
- [18] H. Lee, P. Kok, and J. P. Dowling, “A quantum rosetta stone for interferometry”, *J. Mod. Opt.*, 49, pp. 2325–2338, 2002.
- [19] V. Giovannetti, S. Lloyd, and L. Maccone, “Advances in quantum metrology”, *Nat. Photon.*, 5, pp. 222–229, 2011.
- [20] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński, “Chapter four - Quantum limits in optical interferometry”, *Progress in Optics*, E. Wolf, Ed. Elsevier, 2015, Vol. 60, pp. 345 – 435.
- [21] Y. H. Shih and C. O. Alley, “New type of Einstein-Podolsky-Rosen-Bohm experiment using pairs of light quanta produced by optical parametric down conversion”, *Phys. Rev. Lett.*, 61, pp. 2921–2924, 1988.
- [22] M. Halder, A. Beveratos, N. Gisin, V. Scarani, C. Simon, and H. Zbinden, “Entangling independent photons by time measurement”, *Nature Phys.*, 3, pp. 692–695, 2007.
- [23] M. C. Tichy, M. Tiersch, F. de Melo, F. Mintert, and A. Buchleitner, “Zero-transmission law for multiport beam splitters”, *Phys. Rev. Lett.*, 104, p. 220405, 2010.
- [24] A. Crespi, “Suppression laws for multiparticle interference in Sylvester interferometers”, *Phys. Rev. A*, 91, p. 013811, 2015.
- [25] Y. L. Lim and A. Beige, “Generalized Hong-Ou-Mandel experiments with bosons and fermions”, *New J. Phys.*, 7, p. 155, 2005.
- [26] R. Lopes, A. Imanaliev, A. Aspect, M. Cheneau, D. Boiron, and C. I. Westbrook, “Atomic Hong-Ou-Mandel experiment”, *Nature*, 520, pp. 66–68, 2015.
- [27] K. Toyoda, R. Hiji, A. Noguchi, and S. Urabe, “Hong-Ou-Mandel interference of two phonons in trapped ions”, *Nature*, 527, pp. 74–77, 2015.
- [28] M. Jachura and R. Chrapkiewicz, “Shot-by-shot imaging of Hong-Ou-Mandel interference with an intensified sCMOS camera”, *Opt. Lett.*, 40, pp. 1540–1543, 2015.
- [29] D. Sokolovski, J. Siewert, and L. M. Baskin, “Symmetry-assisted resonance transmission of noninteracting identical particles”, *Phys. Rev. A*, 93, p. 012705, 2016.
- [30] J. D. Urbina, J. Kuipers, S. Matsumoto, Q. Hummel, and K. Richter, “Multiparticle correlations in mesoscopic scattering: Boson sampling, birthday paradox, and Hong-Ou-Mandel profiles”, *Phys. Rev. Lett.*, 116, p. 100401, 2016.
- [31] J. B. Spring *et al.*, “Chip-based array of near-identical, pure, heralded single-photon sources”, *Optica*, 4, pp. 90–96, 2017.



- [32] M. Tillmann, S. H. Tan, S. E. Stoeckl, B. C. Sanders, H. de Guise, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, “Generalized multiphoton quantum interference”, *Phys. Rev. X*, 5, p. 041015, 2015.
- [33] N. Spagnolo, C. Vitelli, L. Aparo, P. Mataloni, F. Sciarrino, A. Crespi, R. Ramponi, and R. Osellame, “Three-photon bosonic coalescence in an integrated tritter”, *Nat. Commun.*, 4, 2013.
- [34] B. J. Metcalf *et al.*, “Multiphoton quantum interference in a multiport integrated photonic device”, *Nat. Commun.*, 4, 2013.
- [35] R. Crespi, A. Osellame, R. Ramponi, D. J. Brod, E. F. Galvao, N. Spagnolo, C. Vitelli, E. Maiorino, P. Mataloni, and F. Sciarrino, “Integrated multimode interferometers with arbitrary designs for photonic boson sampling”, *Nat. Photon.*, 7, pp. 545–549, 2013.
- [36] J. B. Spring *et al.*, “Boson sampling on a photonic chip”, *Science*, 339, pp. 798–801, 2013.
- [37] X. C. Yao, T. X. Wang, P. Xu, H. Lu, G. S. Pan, X. H. Bao, C. Z. Peng, C. Y. Lu, Y. A. Chen, and J. W. Pan, “Observation of eight-photon entanglement”, *Nat. Photon.*, 6, pp. 225–228, 2012.
- [38] J. C. F. Matthews, A. Politi, D. Bonneau, and J. L. O’Brien, “Heralding two-photon and four-photon path entanglement on a chip”, *Phys. Rev. Lett.*, 107, p. 163602, 2011.
- [39] N. Viggianiello *et al.*, “Experimental generalized quantum suppression law in Sylvester interferometers”, 2017, arXiv:1705.08650 [quant-ph].
- [40] V. S. Shchesnovich, “Universality of generalized bunching and efficient assessment of boson sampling”, *Phys. Rev. Lett.*, 116, p. 123601, 2016.
- [41] M. Bentivegna *et al.*, “Experimental scattershot boson sampling”, *Sci. Adv.*, 1, 2015.
- [42] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White, “Photonic boson sampling in a tunable circuit”, *Science*, 339, pp. 794–798, 2013.
- [43] N. Spagnolo *et al.*, “Experimental validation of photonic boson sampling”, *Nat. Photon.*, 8, pp. 615–620, 2014.
- [44] M. C. Tichy, K. Mayer, A. Buchleitner, and K. Mølmer, “Stringent and efficient assessment of boson-sampling devices”, *Phys. Rev. Lett.*, 113, p. 020502, 2014.
- [45] S. Aaronson and A. Arkhipov, “The computational complexity of linear optics”, in *Proceedings of the Forty-third Annual ACM Symposium on Theory of Computing*, STOC ’11. New York, NY, USA: ACM, 2011, pp. 333–342.
- [46] M. C. Tichy, M. Tiersch, F. Mintert, and A. Buchleitner, “Many-particle interference beyond many-boson and many-fermion statistics”, *New J. Phys.*, 14, p. 093015, 2012.

- [47] K. Mayer, M. C. Tichy, F. Mintert, T. Konrad, and A. Buchleitner, “Counting statistics of many-particle quantum walks”, *Phys. Rev. A*, 83, p. 062307, 2011.
- [48] S. Laibacher and V. Tamma, “From the physics to the computational complexity of multiboson correlation interference”, *Phys. Rev. Lett.*, 115, p. 243605, 2015.
- [49] M. C. Tichy, H. T. Lim, Y. S. Ra, F. Mintert, Y. H. Kim, and A. Buchleitner, “Four-photon indistinguishability transition”, *Phys. Rev. A*, 83, p. 062111, 2011.
- [50] Y.-S. Ra, M. C. Tichy, H.-T. Lim, O. Kwon, F. Mintert, A. Buchleitner, and Y.-H. Kim, “Nonmonotonic quantum-to-classical transition in multiparticle interference”, *Proceedings of the National Academy of Sciences*, 110, pp. 1227–1231, 2013.
- [51] Y. S. Ra, M. C. Tichy, H. T. Lim, O. Kwon, F. Mintert, A. Buchleitner, and Y. H. Kim, “Observation of detection-dependent multi-photon coherence times”, *Nat. Photon.*, 4, 2013.
- [52] L. Mandel, “Photon interference and correlation effects produced by independent quantum sources”, *Phys. Rev. A*, 28, pp. 929–943, 1983.
- [53] H. Paul, “Interference between independent photons”, *Rev. Mod. Phys.*, 58, pp. 209–231, 1986.
- [54] E. by E. R. Pike and H. Walter, *Photons and Quantum Fluctuations*, Malvern physics series. Adam Hilger, Bristol, 1988.
- [55] J. Bjorken and S. Drell, *Relativistic Quantum Fields*, Dover Books on Physics Series. Dover Publications, Incorporated, 2013.
- [56] K. J. Blow, R. Loudon, S. J. D. Phoenix, and T. J. Shepherd, “Continuum fields in quantum optics”, *Phys. Rev. A*, 42, pp. 4102–4114, 1990.
- [57] R. J. Glauber, “The quantum theory of optical coherence”, *Phys. Rev.*, 130, pp. 2529–2539, 1963.
- [58] R. J. Glauber, “Coherent and incoherent states of the radiation field”, *Phys. Rev.*, 131, pp. 2766–2788, 1963.
- [59] M. C. Tichy, “Interference of identical particles from entanglement to boson-sampling”, *J. Phys. B: At. Mol. Opt. Phys.*, 47, p. 103001, 2014.
- [60] V. S. Shchesnovich, “Partial indistinguishability theory for multiphoton experiments in multiport devices”, *Phys. Rev. A*, 91, p. 013844, 2015.
- [61] A. Zeilinger, “General properties of lossless beam splitters in interferometry”, *Am. J. Phys.*, 49, pp. 882–883, 1981.
- [62] Z. Y. Ou, “Quantum theory of fourth-order interference”, *Phys. Rev. A*, 37, pp. 1607–1619, 1988.
- [63] V. Tamma and S. Laibacher, “Multiboson correlation interferometry with arbitrary single-photon pure states”, *Phys. Rev. Lett.*, 114, p. 243601, 2015.
- [64] C. Thiel, T. Bastin, J. Martin, E. Solano, J. von Zanthier, and G. S. Agarwal, “Quantum imaging with incoherent photons”, *Phys. Rev. Lett.*, 99, p. 133603, 2007.

- [65] S. Oppel, T. Büttner, P. Kok, and J. von Zanthier, “Superresolving multiphoton interferences with independent light sources”, *Phys. Rev. Lett.*, 109, p. 233603, 2012.
- [66] D. Gatto Monticone, K. Katamadze, P. Traina, E. Moreva, J. Forneris, I. Ruobberchera, P. Olivero, I. P. Degiovanni, G. Brida, and M. Genovese, “Beating the Abbe diffraction limit in confocal microscopy via nonclassical photon statistics”, *Phys. Rev. Lett.*, 113, p. 143602, 2014.
- [67] M. E. Pearce, T. Mehringer, J. von Zanthier, and P. Kok, “Precision estimation of source dimensions from higher-order intensity correlations”, *Phys. Rev. A*, 92, p. 043831, 2015.
- [68] M. Walschaers, J. Kuipers, J. Urbina, K. Mayer, M. C. Tichy, K. Richter, and A. Buchleitner, “Statistical benchmark for bosonsampling”, *New J. Phys.*, 18, p. 032001, 2016.
- [69] M. Walschaers, J. Kuipers, and A. Buchleitner, “From many-particle interference to correlation spectroscopy”, *Phys. Rev. A*, 94, p. 020104, 2016.
- [70] A. Crespi, R. Osellame, R. Ramponi, M. Bentivegna, F. Flamini, N. Spagnolo, N. Viggianiello, L. Innocenti, P. Mataloni, and F. Sciarrino, “Suppression law of quantum states in a 3D photonic fast Fourier transform chip”, *Nat. Commun.*, 7, 2016.
- [71] A. J. Menssen, A. E. Jones, B. J. Metcalf, M. C. Tichy, S. Barz, W. S. Kolthammer, and I. A. Walmsley, “Distinguishability and many-particle interference”, *Phys. Rev. Lett.*, 118, p. 153603, 2017.
- [72] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, “Optimal design for universal multiport interferometers”, *Optica*, 3, pp. 1460–1465, 2016.
- [73] E. C. G. Sudarshan, “Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams”, *Phys. Rev. Lett.*, 10, pp. 277–279, 1963.
- [74] J. Sperling, “Characterizing maximally singular phase-space distributions”, *Phys. Rev. A*, 94, p. 013814, 2016.
- [75] E. Shchukin, T. Richter, and W. Vogel, “Nonclassicality criteria in terms of moments”, *Phys. Rev. A*, 71, 2005.
- [76] L. Mandel, “Sub-Poissonian photon statistics in resonance fluorescence”, *Opt. Lett.*, 4, pp. 205–207, 1979.
- [77] A. Miranowicz, M. Bartkowiak, X. Wang, Y.-x. Liu, and F. Nori, “Testing nonclassicality in multimode fields: A unified derivation of classical inequalities”, *Phys. Rev. A*, 82, p. 013824, 2010.
- [78] J. Sperling, W. Vogel, and G. S. Agarwal, “True photocounting statistics of multiple on-off detectors”, *Phys. Rev. A*, 85, p. 023820, 2012.
- [79] J. Sperling, W. Vogel, and G. S. Agarwal, “Sub-binomial light”, *Phys. Rev. Lett.*, 109, p. 093601, 2012.

- [80] J. Sperling, W. Vogel, and G. S. Agarwal, “Correlation measurements with on-off detectors”, *Phys. Rev. A*, 88, 2013.
- [81] J. Sperling, M. Bohmann, W. Vogel, G. Harder, B. Brecht, V. Ansari, and C. Silberhorn, “Uncovering quantum correlations with time-multiplexed click detection”, *Phys. Rev. Lett.*, 115, p. 023601, 2015.
- [82] J. Sperling *et al.*, “Identification of nonclassical properties of light with multiplexing layouts”, *Phys. Rev. A*, 96, p. 013804, 2017.
- [83] J. Sperling *et al.*, “Detector-independent verification of quantum light”, *Phys. Rev. Lett.*, 118, p. 163602, 2017.
- [84] Arvind and N. Mukunda, “Bell’s inequalities, multiphoton states and phase space distributions”, *Phys. Lett. A*, 259, pp. 421 – 426, 1999.
- [85] M. S. Kim, W. Son, V. Bužek, and P. L. Knight, “Entanglement by a beam splitter: Nonclassicality as a prerequisite for entanglement”, *Phys. Rev. A*, 65, p. 032323, 2002.
- [86] W. Xiang-bin, “Theorem for the beam-splitter entangler”, *Phys. Rev. A*, 66, p. 024303, 2002.
- [87] M. M. Wolf, J. Eisert, and M. B. Plenio, “Entangling power of passive optical elements”, *Phys. Rev. Lett.*, 90, p. 047904, 2003.
- [88] J. K. Asbóth, J. Calsamiglia, and H. Ritsch, “Computable measure of nonclassicality for light”, *Phys. Rev. Lett.*, 94, p. 173602, 2005.
- [89] J. Xiong, N. Zhou, and G. Zeng, “Second-order coherence of light fields with a beam splitter”, *J. Phys. B: At. Mol. Opt. Phys.*, 38, p. 4301, 2005.
- [90] S. C. Springer, J. Lee, M. Bellini, and M. S. Kim, “Conditions for factorizable output from a beam splitter”, *Phys. Rev. A*, 79, p. 062303, 2009.
- [91] J. S. Ivan, S. Chaturvedi, E. Ercolessi, G. Marmo, G. Morandi, N. Mukunda, and R. Simon, “Entanglement and nonclassicality for multimode radiation-field states”, *Phys. Rev. A*, 83, p. 032118, 2011.
- [92] Z. Jiang, M. D. Lang, and C. M. Caves, “Mixing nonclassical pure states in a linear-optical network almost always generates modal entanglement”, *Phys. Rev. A*, 88, p. 044301, 2013.
- [93] W. Vogel and J. Sperling, “Unified quantification of nonclassicality and entanglement”, *Phys. Rev. A*, 89, p. 052302, 2014.
- [94] C. Croal, C. Peuntinger, V. Chille, C. Marquardt, G. Leuchs, N. Korolkova, and L. Mišta, “Entangling the whole by beam splitting a part”, *Phys. Rev. Lett.*, 115, p. 190501, 2015.
- [95] M. Brunelli, C. Benedetti, S. Olivares, A. Ferraro, and M. G. A. Paris, “Single- and two-mode quantumness at a beam splitter”, *Phys. Rev. A*, 91, p. 062315, 2015.
- [96] A. Streltsov, U. Singh, H. S. Dhar, M. N. Bera, and G. Adesso, “Measuring quantum coherence with entanglement”, *Phys. Rev. Lett.*, 115, p. 020403, 2015.

- [97] C. M. Caves, “Quantum-mechanical noise in an interferometer”, *Phys. Rev. D*, 23, pp. 1693–1708, 1981.
- [98] Y. Shih, “Entangled biphoton source - property and preparation”, *Rep. Prog. Phys.*, 66, p. 1009, 2003.
- [99] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, “Gaussian quantum information”, *Rev. Mod. Phys.*, 84, pp. 621–669, 2012.
- [100] S. A. Fulling, “Nonuniqueness of canonical field quantization in Riemannian space-time”, *Phys. Rev. D*, 7, pp. 2850–2862, 1973.
- [101] P. C. W. Davies, “Scalar production in Schwarzschild and Rindler metrics”, *J. Phys. A: Math. Gen.*, 8, p. 609, 1975.
- [102] W. G. Unruh, “Notes on black-hole evaporation”, *Phys. Rev. D*, 14, pp. 870–892, 1976.
- [103] G. J. Milburn, W.-Y. Chen, and K. R. Jones, “Hyperbolic phase and squeeze-parameter estimation”, *Phys. Rev. A*, 50, pp. 801–804, 1994.
- [104] G. Chiribella, G. M. D’Ariano, and M. F. Sacchi, “Optimal estimation of squeezing”, *Phys. Rev. A*, 73, p. 062103, 2006.
- [105] R. Gaiba and M. G. Paris, “Squeezed vacuum as a universal quantum probe”, *Phys. Lett. A*, 373, pp. 934 – 939, 2009.
- [106] D. Šafránek, A. R. Lee, and I. Fuentes, “Quantum parameter estimation using multi-mode Gaussian states”, *New J. Phys.*, 17, p. 073016, 2015.
- [107] D. Šafránek and I. Fuentes, “Optimal probe states for the estimation of Gaussian unitary channels”, *Phys. Rev. A*, 94, p. 062313, 2016.
- [108] M. G. Genoni, C. Invernizzi, and M. G. A. Paris, “Enhancement of parameter estimation by Kerr interaction”, *Phys. Rev. A*, 80, p. 033842, 2009.
- [109] S. L. Braunstein and C. M. Caves, “Statistical distance and the geometry of quantum states”, *Phys. Rev. Lett.*, 72, pp. 3439–3443, 1994.
- [110] V. Giovannetti, S. Lloyd, and L. Maccone, “Quantum metrology”, *Phys. Rev. Lett.*, 96, p. 010401, 2006.
- [111] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guta, “The elusive Heisenberg limit in quantum-enhanced metrology”, *Nat. Commun.*, 3, 2012.
- [112] G. Toth and I. Apellaniz, “Quantum metrology from a quantum information science perspective”, *J. Phys. A: Math. Theor.*, 47, p. 424006, 2014.
- [113] R. Demkowicz-Dobrzański and L. Maccone, “Using entanglement against noise in quantum metrology”, *Phys. Rev. Lett.*, 113, p. 250801, 2014.
- [114] R. Schnabel, N. Mavalvala, D. E. McClelland, and P. K. Lam, “Quantum metrology for gravitational wave astronomy”, *Nat. Commun.*, 1, 2010.
- [115] R. Demkowicz-Dobrzański, K. Banaszek, and R. Schnabel, “Fundamental quantum interferometry bound for the squeezed-light-enhanced gravitational wave detector

- GEO 600”, *Phys. Rev. A*, 88, p. 041802, 2013.
- [116] M. A. Taylor and W. P. Bowen, “Quantum metrology and its application in biology”, *Phys. Rep.*, 615, pp. 1 – 59, 2016.
  - [117] J. C. Matthews, X.-Q. Zhou, H. Cable, P. J. Shadbolt, D. J. Saunders, G. A. Durkin, G. J. Pryde, and J. L. O’Brien, “Towards practical quantum metrology with photon counting”, *npjQI*, 2, p. 16023, 2016.
  - [118] M. Jachura, R. Chrapkiewicz, R. Demkowicz-Dobrzański, W. Wasilewski, and K. Banaszek, “Mode engineering for realistic quantum-enhanced interferometry”, *Nat. Commun.*, 7, p. 11411, 2016.
  - [119] D. Girolami, A. M. Souza, V. Giovannetti, T. Tufarelli, J. G. Filgueiras, R. S. Sarthour, D. O. Soares-Pinto, I. S. Oliveira, and G. Adesso, “Quantum discord determines the interferometric power of quantum states”, *Phys. Rev. Lett.*, 112, p. 210401, 2014.
  - [120] A. Farace, A. D. Pasquale, L. Rigovacca, and V. Giovannetti, “Discriminating Strength: a bona fide measure of non-classical correlations”, *New J. Phys.*, 16, p. 073010, 2014.
  - [121] D. Girolami, T. Tufarelli, and G. Adesso, “Characterizing nonclassical correlations via local quantum uncertainty”, *Phys. Rev. Lett.*, 110, p. 240402, 2013.
  - [122] W. Roga, S. M. Giampaolo, and F. Illuminati, “Discord of response”, *J. Phys. A: Math. Theor.*, 47, p. 365301, 2014.
  - [123] H. Ollivier and W. H. Zurek, “Quantum discord: A measure of the quantumness of correlations”, *Phys. Rev. Lett.*, 88, p. 017901, 2001.
  - [124] L. Henderson and V. Vedral, “Classical, quantum and total correlations”, *J. Phys. A: Math. Gen.*, 34, p. 6899, 2001.
  - [125] G. Adesso, T. R. Bromley, and M. Cianciaruso, “Measures and applications of quantum correlations”, *J. Phys. A: Math. Theor.*, 49, p. 473001, 2016.
  - [126] W. Roga, D. Spehner, and F. Illuminati, “Geometric measures of quantum correlations: characterization, quantification, and comparison by distances and operations”, *J. Phys. A: Math. Theor.*, 49, p. 235301, 2016.
  - [127] T. R. Bromley, I. A. Silva, C. O. Oncebay-Segura, D. O. Soares-Pinto, E. R. deAzevedo, T. Tufarelli, and G. Adesso, “There is more to quantum interferometry than entanglement”, *Phys. Rev. A*, 95, p. 052313, 2017.
  - [128] D. Braun, G. Adesso, F. Benatti, R. Floreanini, U. Marzolino, M. W. Mitchell, and S. Pirandola, “Quantum enhanced measurements without entanglement”, 2017, arXiv:1701.05152 [quant-ph].
  - [129] G. Adesso, “Gaussian Interferometric Power”, *Phys. Rev. A*, 90, p. 022321, 2014.
  - [130] L. Rigovacca, A. Farace, A. De Pasquale, and V. Giovannetti, “Gaussian Discriminating Strength”, *Phys. Rev. A*, 92, p. 042331, 2015.
  - [131] W. Roga, D. Buono, and F. Illuminati, “Device-independent quantum reading and

- noise-assisted quantum transmitters”, *New J. Phys.*, 17, p. 013031, 2015.
- [132] A. Farace, A. D. Pasquale, G. Adesso, and V. Giovannetti, “Building versatile bipartite probes for quantum metrology”, *New J. Phys.*, 18, p. 013049, 2016.
  - [133] R. Nichols, T. R. Bromley, L. A. Correa, and G. Adesso, “Practical quantum metrology in noisy environments”, *Phys. Rev. A*, 94, p. 042101, 2016.
  - [134] M. G. A. Paris, “Quantum estimation for quantum technology”, *Int. J. Quantum Inf.*, 07, pp. 125–137, 2009.
  - [135] A. Ferraro, S. Olivares, and M. G. A. Paris, *Gaussian states in quantum information*, Napoli Series on Physics and Astrophysics. Bibliopolis, 2005.
  - [136] G. Adesso, S. Ragy, and A. R. Lee, “Continuous variable quantum information: Gaussian states and beyond”, *OSID*, 21, p. 1440001, 2014.
  - [137] H. Cramér, *Mathematical Methods of Statistics. (PMS-9)*. Princeton University Press, 1999.
  - [138] A. Fujiwara and H. Imai, “A fibre bundle over manifolds of quantum channels and its application to quantum statistics”, *J. Phys. A: Math. Theor.*, 41, p. 255304, 2008.
  - [139] W. K. Wootters, “Statistical distance and Hilbert space”, *Phys. Rev. D*, 23, pp. 357–362, 1981.
  - [140] M. Hübner, “Explicit computation of the Bures distance for density matrices”, *Phys. Lett. A*, 163, pp. 239 – 242, 1992.
  - [141] D. Bures, “An extension of Kakutani’s theorem on infinite product measures to the tensor product of semifinite  $w^*$ -algebras”, *Trans. Amer. Math. Soc.*, 135, pp. 199–212, 1969.
  - [142] A. Uhlmann, “The transition probability in the state space of a  $*$ -algebra”, *Rep. Math. Phys.*, 9, pp. 273 – 279, 1976.
  - [143] D. Šafránek, “Discontinuities of the quantum Fisher information and the Bures metric”, *Phys. Rev. A*, 95, p. 052320, 2017.
  - [144] O. E. Barndorff-Nielsen and R. D. Gill, “Fisher information in quantum statistics”, *J. Phys. A: Math. Gen.*, 33, p. 4481, 2000.
  - [145] A. D. Pasquale, D. Rossini, P. Facchi, and V. Giovannetti, “Quantum parameter estimation affected by unitary disturbance”, *Phys. Rev. A*, 88, p. 052117, 2013.
  - [146] S. Olivares, “Quantum optics in the phase space”, *EPJ ST*, 203, pp. 3–24, 2012.
  - [147] J. Williamson, “On the algebraic problem concerning the normal forms of linear dynamical systems”, *Am. J. Math.*, 58, pp. 141–163, 1936.
  - [148] M. G. Genoni, L. Lami, and A. Serafini, “Conditional and unconditional Gaussian quantum dynamics”, *Contemp. Phys.*, 57, pp. 331–349, 2016.
  - [149] O. Pinel, P. Jian, N. Treps, C. Fabre, and D. Braun, “Quantum parameter estimation using general single-mode Gaussian states”, *Phys. Rev. A*, 88, p. 040102, 2013.

- [150] L. Banchi, S. L. Braunstein, and S. Pirandola, “Quantum fidelity for arbitrary Gaussian states”, *Phys. Rev. Lett.*, 115, p. 260501, 2015.
- [151] N. Friis, M. Skotiniotis, I. Fuentes, and W. Dür, “Heisenberg scaling in Gaussian quantum metrology”, *Phys. Rev. A*, 92, p. 022106, 2015.
- [152] C. Lupo, S. Mancini, A. D. Pasquale, P. Facchi, G. Florio, and S. Pascazio, “Invariant measures on multimode quantum Gaussian states”, *J. Math. Phys.*, 53, p. 122209, 2012.
- [153] K. Życzkowski, K. A. Penson, I. Nechita, and B. Collins, “Generating random density matrices”, *J. Math. Phys.*, 52, p. 062201, 2011.
- [154] E. P. Wigner and M. M. Yanase, “Information contents of distributions”, *Proceedings of the National Academy of Sciences*, 49, pp. 910–918, 1963.
- [155] S. Luo, “Wigner-Yanase skew information and uncertainty relations”, *Phys. Rev. Lett.*, 91, p. 180403, 2003.
- [156] D. P. Pires, M. Cianciaruso, L. C. Céleri, G. Adesso, and D. O. Soares-Pinto, “Generalized geometric quantum speed limits”, *Phys. Rev. X*, 6, p. 021031, 2016.
- [157] M. M. Taddei, B. M. Escher, L. Davidovich, and R. L. de Matos Filho, “Quantum speed limit for physical processes”, *Phys. Rev. Lett.*, 110, p. 050402, 2013.
- [158] F. Fröwis and W. Dür, “Measures of macroscopicity for quantum spin systems”, *New J. Phys.*, 14, p. 093039, 2012.
- [159] B. Yadin and V. Vedral, “General framework for quantum macroscopicity in terms of coherence”, *Phys. Rev. A*, 93, p. 022122, 2016.
- [160] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, “Quantum entanglement”, *Rev. Mod. Phys.*, 81, pp. 865–942, 2009.
- [161] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, “Secure key from bound entanglement”, *Phys. Rev. Lett.*, 94, p. 160502, 2005.
- [162] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, “General paradigm for distilling classical key from quantum states”, *IEEE Trans. Inf. Theory*, 55, pp. 1898–1929, 2009.
- [163] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, “The security of practical quantum key distribution”, *Rev. Mod. Phys.*, 81, pp. 1301–1350, 2009.
- [164] M. Takeoka, S. Guha, and M. M. Wilde, “Fundamental rate-loss tradeoff for optical quantum key distribution”, *Nat. Commun.*, 5, 2014.
- [165] M. Takeoka, S. Guha, and M. M. Wilde, “The squashed entanglement of a quantum channel”, *IEEE Trans. Inf. Theory*, 60, pp. 4987–4998, 2014.
- [166] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi, “Fundamental limits of repeaterless quantum communications”, *Nat. Commun.*, 8, 2017.
- [167] M. Christandl and A. Müller-Hermes, “Relative entropy bounds on quantum, private and repeater capacities”, *Commun. Math. Phys.*, 353, pp. 821–852, 2017.



- [168] K. Goodenough, D. Elkouss, and S. Wehner, “Assessing the performance of quantum repeaters for all phase-insensitive Gaussian bosonic channels”, *New J. Phys.*, 18, p. 063005, 2016.
- [169] M. M. Wilde, M. Tomamichel, and M. Berta, “Converse bounds for private communication over quantum channels”, *IEEE Trans. Inf. Theory*, 63, pp. 1792–1817, 2017.
- [170] E. Kaur and M. M. Wilde, “Upper bounds on secret key agreement over lossy thermal bosonic channels”, 2017, arXiv:1706.04590 [quant-ph].
- [171] T. Cope, L. Hetzel, L. Bianchi, and S. Pirandola, “Simulation of non-Pauli Channels”, 2017, arXiv:1706.05384 [quant-ph].
- [172] R. Laurenza, S. L. Braunstein, and S. Pirandola, “Finite-resource teleportation stretching for continuous-variable systems”, 2017, arXiv:1706.06065 [quant-ph].
- [173] K. Azuma, A. Mizutani, and H.-K. Lo, “Fundamental rate-loss tradeoff for the quantum internet”, *Nat. Commun.*, 7, 2016.
- [174] E. Schoute, L. Mancinska, T. Islam, I. Kerenidis, and S. Wehner, “Shortcuts to quantum network routing”, 2016, arXiv:1610.05238 [quant-ph].
- [175] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, “Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels”, *Phys. Rev. Lett.*, 70, pp. 1895–1899, 1993.
- [176] M. Christandl and A. Winter, ““Squashed entanglement”: An additive entanglement measure”, *J. Math. Phys.*, 45, pp. 829–840, 2004.
- [177] S. Pirandola, “Capacities of repeater-assisted quantum communications”, 2016, arXiv:1601.00966 [quant-ph].
- [178] H. Umegaki, “Conditional expectation in an operator algebra. IV. Entropy and information”, *Kodai Math. Sem. Rep.*, 14, pp. 59–85, 1962.
- [179] N. Datta, “Min- and max-relative entropies and a new entanglement monotone”, *IEEE Trans. Inf. Theory*, 55, pp. 2816–2826, 2009.
- [180] V. Vedral and M. B. Plenio, “Entanglement measures and purification procedures”, *Phys. Rev. A*, 57, pp. 1619–1633, 1998.
- [181] N. Datta, “Max-relative entropy of entanglement, alias log robustness”, *Int. J. Quantum Inf.*, 07, pp. 475–491, 2009.
- [182] W. F. Stinespring, “Positive functions on C-algebras”, *Proc. Amer. Math. Soc.*, 6, p. 211216, 1955.
- [183] K. Horodecki, M. Horodecki, P. Horodecki, D. Leung, and J. Oppenheim, “Quantum key distribution based on private states: Unconditional security over untrusted channels with zero quantum capacity”, *IEEE Trans. Inf. Theory*, 54, pp. 2604–2620, 2008.
- [184] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, “Mixed-state entanglement and quantum error correction”, *Phys. Rev. A*, 54, pp. 3824–3851,

- 1996.
- [185] D. Gottesman and I. L. Chuang, “Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations”, *Nature*, 402, pp. 390–393, 1999.
  - [186] E. Knill, R. Laflamme, and G. J. Milburn, “A scheme for efficient quantum computation with linear optics”, *Nature*, 409, pp. 46–52, 2001.
  - [187] S. Bäuml and K. Azuma, “Fundamental limitation on quantum broadcast networks”, *Quantum Sci. Technol.*, 2, p. 024004, 2017.
  - [188] K. Azuma and G. Kato, “Aggregating quantum repeaters for the quantum internet”, 2016, arXiv:1606.00135 [quant-ph].
  - [189] B. Synak-Radtke and M. Horodecki, “On asymptotic continuity of functions of quantum states”, *J. Phys. A: Math. Gen.*, 39, p. L423, 2006.
  - [190] M. M. Wilde, “Squashed entanglement and approximate private states”, *Quantum Inf. Process.*, 15, pp. 4563–4580, 2016.
  - [191] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, “On quantum Rényi entropies: A new generalization and some properties”, *J. Math. Phys.*, 54, 2013.
  - [192] M. M. Wilde, A. Winter, and D. Yang, “Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy”, *Commun. Math. Phys.*, 331, pp. 593–622, 2014.
  - [193] M. Horodecki, P. W. Shor, and M. B. Ruskai, “Entanglement breaking channels”, *Rev. Math. Phys.*, 15, pp. 629–641, 2003.
  - [194] M. Christandl and A. Winter, “Uncertainty, monogamy, and locking of quantum correlations”, *IEEE Trans. Inf. Theory*, 51, pp. 3159–3165, 2005.
  - [195] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, “Locking entanglement with a single qubit”, *Phys. Rev. Lett.*, 94, p. 200501, 2005.
  - [196] J. Yard, P. Hayden, and I. Devetak, “Quantum broadcast channels”, *IEEE Trans. Inf. Theory*, 57, pp. 7147–7162, 2011.
  - [197] K. P. Seshadreesan, M. Takeoka, and M. M. Wilde, “Bounds on entanglement distillation and secret key agreement for quantum broadcast channels”, *IEEE Trans. Inf. Theory*, 62, pp. 2849–2866, 2016.
  - [198] R. Laurenza and S. Pirandola, “General bounds for sender-receiver capacities in multipoint quantum communications”, 2016, arXiv:1603.07262 [quant-ph].
  - [199] M. Takeoka, K. P. Seshadreesan, and M. M. Wilde, “Unconstrained distillation capacities of a pure-loss bosonic broadcast channel”, in *2016 IEEE International Symposium on Information Theory*, 2016, pp. 2484–2488.
  - [200] M. Takeoka, K. P. Seshadreesan, and M. M. Wilde, “Unconstrained capacities of quantum key distribution and entanglement distillation for pure-loss bosonic broadcast channels”, 2017, arXiv:1706.06746 [quant-ph].

- 
- [201] C. A. Fuchs and J. van de Graaf, “Cryptographic distinguishability measures for quantum-mechanical states”, *IEEE Trans. Inf. Theory*, 45, pp. 1216–1227, 1999.



# Appendix A

## Technicalities on squeezing estimation

### A.1 Two-mode AvQFI in noiseless conditions

The general expression for the noiseless two-mode AvQFI for squeezing estimation, as function of the parameters appearing in the standard form of Equation (3.61), can be written as

$$\overline{H}[\rho_{AB}] = \frac{N_1}{D_1} + \frac{N_2}{D_2} + \overline{H}_{\text{disp}}[\rho_{AB}], \quad (\text{A.1})$$

where  $\overline{H}_{\text{disp}}[\rho_{AB}]$  is the contribution coming from the displacement vector of the probe, with  $|\xi| = \sqrt{\xi_x^2 + \xi_p^2}$ :

$$\overline{H}_{\text{disp}}[\rho_{AB}] = \frac{b|\xi|^2}{\det \Gamma} (b(a_x + a_p) - c^2 - d^2), \quad (\text{A.2})$$

and the terms  $N_1, D_1, N_2$  and  $D_2$  can be explicitly written as

$$N_1 = c^2 [4d^2 - b(a_x + 5a_p)] + b [b(a_x^2 + 6a_x a_p + a_p^2 - 4a_{xp}^2) - d^2(5a_x + a_p)], \quad (\text{A.3})$$

$$D_1 = 2 [-b^2 a_{xp}^2 + (c^2 - a_x b) (d^2 - b a_p) - 1], \quad (\text{A.4})$$

$$N_2 = -4 (b^2 + cd + 1) [- (b^2 + 1) a_{xp}^2 + a_x (a_p b^2 - b d^2 + a_p) + c^2 (d^2 - b a_p) + cd] \\ - [a_x + b^2(a_x + a_p) - b(c^2 + d^2) + a_p]^2, \quad (\text{A.5})$$

$$D_2 = 2 [-b^2 a_{xp}^2 + (c^2 - a_x b) (d^2 - b a_p) - 1] \\ \cdot [a_x a_p + b^2 (a_x a_p - a_{xp}^2 + 1) - b (a_p c^2 + a_x d^2) - a_{xp}^2 + (cd + 1)^2]. \quad (\text{A.6})$$

## A.2 Single-mode probes with extreme AvQFI values

Here we formally prove that pure squeezed states maximise the AvQFI of Equation (3.69), while thermal states minimise it, among all single-mode probes with the same number of photons  $n_A$ . We parametrise the probes as in Equation (3.62), so that the AvQFI becomes a function of the absolute value of the input displacement  $|\xi|$  and of the quantities  $\text{Tr}[\Gamma_A] = 2\nu_A \cosh(2\alpha)$  and  $\det \Gamma_A = \nu_A^2$ .

At first, we prove that single-mode squeezed states have the maximum AvQFI. By exploiting the expression for  $n_A$  given in Equation (3.47), we can substitute the parameter  $\alpha$  in  $\overline{H}$  and obtain

$$\overline{H}|_{n_A} = 2 \frac{(2n_A + 1 - |\xi|^2)^2 + \nu_A^2}{1 + \nu_A^2} + 2|\xi|^2 \frac{2n_A + 1 - |\xi|^2}{\nu_A^2}. \quad (\text{A.7})$$

Notice how the displacement has an overall negative contribution in this expression: if the photon-number is fixed, reducing the displacement will always improves the performance of the probe. We now show that, for  $|\xi| = 0$ ,  $\overline{H}|_{n_A}$  is maximised by  $\nu_A = 1$ . Note that the first derivative of the function

$$f_{a,b}(x) = \frac{a+x}{b+x}, \quad (\text{A.8})$$

has the same sign of  $b - a$ . As the average QFI can be written as

$$\overline{H}|_{n_A, |\xi|=0} = 2f_{(2n_A+1)^2, 1}(\nu_A), \quad (\text{A.9})$$

we can conclude that the AvQFI is maximised by choosing  $\nu_A = 1$ , which is the minimum possible symplectic eigenvalue.

Let us turn our attention to the problem of finding the single-mode probe with minimum  $\overline{H}$  for fixed  $n_A$ . By using Equation (3.47) to substitute the value of  $\nu_A$  in Equation (3.69), we find

$$\begin{aligned} \overline{H}|_{n_A} &= 2(2n_A + 1 - |\xi|^2)^2 \frac{1 + \cosh^2(2\alpha)}{(2n_A + 1 - |\xi|^2)^2 + \cosh^2(2\alpha)} \\ &\quad + 2|\xi|^2 \cosh^2(2\alpha) \frac{1}{(2n_A + 1 - |\xi|)}. \end{aligned} \quad (\text{A.10})$$

The first fraction appearing on the right-hand side can be rewritten as  $f_{a,b}(\cosh^2[2\alpha])$ , with  $a = 1$  and  $b = (2n_A + 1 - |\xi|^2)^2 \geq 1$ . Therefore, with the same reasoning used above, we can conclude that the AvQFI is increasing with  $\cosh(2\alpha)$ , and is minimised by  $\alpha = 0$ . At last, we have to show that the minimum of  $\overline{H}|_{n_A, \alpha=0}$  is reached for  $|\xi| = 0$ . We can prove this by calculating the first derivative of Equation (A.10) on  $|\xi|^2$  for  $\alpha = 0$ , and by showing that it is always positive. After some straightforward manipulations, the

positivity of the aforementioned derivative can be written as

$$\sqrt{2}(Y - 2X)(Y + 2X) \geq 0, \quad (\text{A.11})$$

where we introduced the auxiliary positive quantities

$$X = (2n_A + 1 - |\xi|^2), \quad Y = \sqrt{2n_A + 1} [1 + X^2]. \quad (\text{A.12})$$

The proof is now concluded because

$$Y - 2X = (\sqrt{2n_A + 1} - 1)(1 + X^2) + (1 - X)^2 \geq 0. \quad (\text{A.13})$$

### A.3 Single-mode QFI under noisy conditions

By setting without loss of generality  $\phi = 0$  in Equation (3.62), the single-mode  $\theta$ -dependent QFI for the noisy evolution given in Equation (3.82) can be written as:

$$H_{\epsilon, \theta}^{(1)} = \frac{\tilde{N}_1}{\tilde{D}_1} + \frac{\tilde{N}_2}{\tilde{D}_2} + \frac{\tilde{N}_3}{\tilde{D}_3}, \quad (\text{A.14})$$

where the coefficients  $\tilde{N}_1, \tilde{N}_2, \tilde{N}_3, \tilde{D}_1, \tilde{D}_2, \tilde{D}_3$  are defined as:

$$\begin{aligned} \tilde{N}_1 &= [(\eta - 1)^2 \eta^2 e^{-4(\alpha + \mu)}] \\ &\cdot [(e^{4\alpha} - 1) \eta \nu (e^{4\mu} + 1) \cos(2\theta) + 4e^{2(\alpha + \mu)} \sinh(2\mu)(\eta \nu \cosh(2\alpha) - \eta + 1)]^2, \end{aligned} \quad (\text{A.15})$$

$$\begin{aligned} \sqrt{\frac{\tilde{D}_1}{2} + 1} &= \frac{1}{2} e^{-2(\alpha + \mu)} (e^{4\alpha} - 1) (1 - \eta) \eta^2 \nu (e^{4\mu} - 1) \cos(2\theta) + \eta^4 \nu^2 + (1 - \eta)^2 (\eta^2 + 1) \\ &+ 2(1 - \eta) \eta [\eta \nu \cosh(2\alpha)(\eta + \cosh(2\mu)) + (1 - \eta) \cosh(2\mu)], \end{aligned} \quad (\text{A.16})$$

$$\begin{aligned} \frac{\tilde{N}_2}{e^{-2\alpha} \eta^2} &= -2e^{4\alpha} \{ \eta \nu [\eta \nu \cosh(4\alpha) - 8(\eta - 1) \cosh(2\alpha)] + \eta [\eta (3\nu^2 + 4) - 8] + 4 \} \\ &+ (e^{4\alpha} - 1)^2 \eta^2 \nu^2 \cos(4\theta), \end{aligned} \quad (\text{A.17})$$

$$\begin{aligned} \frac{\tilde{D}_2}{2} &= e^{4\alpha} (\eta - 1) \eta^2 \nu [\eta + \cos(2\theta) \sinh(2\mu) + \cosh(2\mu)] \\ &- e^{2\alpha} \{ \eta [\eta (\eta \nu^2 + \eta - 2) + 2] + 2\eta (\eta - 1)^2 \cosh(2\mu) + 2 \} \\ &+ (\eta - 1) \eta^2 \nu [\eta - \cos(2\theta) \sinh(2\mu) + \cosh(2\mu)], \end{aligned} \quad (\text{A.18})$$

$$\begin{aligned}
\frac{\tilde{N}_3}{\eta^3} = & + 2(e^{4\alpha} - 1) \eta \nu \sin(4\theta) \xi_p \xi_x - (e^{4\alpha} - 1) \eta \nu \cos(4\theta) (\xi_p - \xi_x) (\xi_p + \xi_x) \\
& + 2e^{2\alpha}(\eta - 1) \sinh(2\mu) [-\cos(2\theta) \xi_p^2 + 2 \sin(2\theta) \xi_p \xi_x + \cos(2\theta) \xi_x^2] \\
& + 2e^{2\alpha}(\eta - 1) \cosh(2\mu) (\xi_p^2 + \xi_x^2) - 2e^{2\alpha} (\xi_p^2 + \xi_x^2) [\eta \nu \cosh(2\alpha) - \eta + 1], \quad (\text{A.19})
\end{aligned}$$

$$\begin{aligned}
\frac{\tilde{D}_3}{e^{2\alpha}} = & 2(\eta - 1) \eta [\eta \nu \sinh(2\alpha) \cos(2\theta) \sinh(2\mu) + \eta \nu \cosh(2\alpha) (\eta + \cosh(2\mu))] \\
& - 2\eta(\eta - 1)^2 \cosh(2\mu) - \nu^2 \eta^4 - (\eta - 1)^2 (\eta^2 + 1). \quad (\text{A.20})
\end{aligned}$$

## A.4 Optimal displacement direction for single-mode probes

We consider the noisy encoding evolution of Equation (3.82), in which photon losses are taken into account. In what follows we formally show that the optimal displacement direction for a single-mode probe characterised by  $\phi = 0$  in Equation (3.62), is  $\psi = \pi/2$ . With respect to the expression reported in Section A.3, note that only the term  $\tilde{N}_3/\tilde{D}_3$  depends on the displacement vector  $(\xi_x, \xi_p)$ , where  $\xi_x = |\xi| \cos \psi$  and  $\xi_p = |\xi| \sin \psi$ .

We start by highlighting the dependence of  $\tilde{N}_3/\tilde{D}_3$  on the angles  $\theta$  and  $\psi$ , by rewriting it as

$$\frac{\tilde{N}_3}{\tilde{D}_3} = \frac{x_0 + x_1 \cos[2(\theta + \psi)] + x_2 \cos[2(2\theta + \psi)]}{x_3 + x_4 \cos(2\theta)}, \quad (\text{A.21})$$

where the coefficients  $\{x_i\}_{i=0}^4$  depend on  $\mu, \eta, \alpha$  and  $\nu$ . In particular

$$x_4 = \nu_A \eta^2 (1 - \eta) (e^{4\alpha} - 1) \sinh(2\mu) \geq 0. \quad (\text{A.22})$$

When  $x_4 = 0$ , the average over  $\theta$  removes the dependence upon  $\psi$  and the thesis becomes trivially true. On the other hand, if  $x_4 \neq 0$ , we can assume without loss of generality that  $\mu, \alpha \geq 0$ , and a comparison with Eqs. (A.19) and (A.20) yields the inequalities

$$x_1 \geq 0, \quad x_2 \leq 0, \quad x_3/x_4 > 1, \quad (\text{A.23})$$

which will shortly become important. By explicitly performing the integration of Equation (A.21) over  $\theta \in [0, 2\pi]$ , and by dividing the result by  $2\pi$ , we obtain the following contribution to the AvQFI:

$$\int_0^{2\pi} \frac{\tilde{N}_3}{\tilde{D}_3} \frac{d\theta}{2\pi} = \frac{x_0}{2\sqrt{x_3^2 - x_4^2}} + \frac{\cos 2\psi}{2x_4} \left[ x_1 \int_0^{2\pi} \frac{d\theta}{2\pi} \frac{\cos \theta}{\frac{x_3}{x_4} + \cos \theta} + x_2 \int_0^{2\pi} \frac{d\theta}{2\pi} \frac{\cos 2\theta}{\frac{x_3}{x_4} + \cos \theta} \right]. \quad (\text{A.24})$$



For  $x_3/x_4 > 1$ , the integrals appearing in Equation (A.24) have a well defined sign:

$$\int_0^{2\pi} \frac{d\theta}{2\pi} \frac{\cos \theta}{\frac{x_3}{x_4} + \cos \theta} \leq 0, \quad \int_0^{2\pi} \frac{d\theta}{2\pi} \frac{\cos 2\theta}{\frac{x_3}{x_4} + \cos \theta} \geq 0, \quad (\text{A.25})$$

so that the term between square brackets cannot be positive. As a consequence, the maximum value of the average ratio  $\tilde{N}_3/\tilde{D}_3$  is reached when  $\psi = \pm\pi/2$ , and the same can be said for the maximum value of the AvQFI.



# Appendix B

## Technicalities on communication bounds

### B.1 Proof of property P1 for the max-relative entropy of entanglement

This is a straightforward adaptation of the reasoning used to prove Lemma IV.2 in Ref. [167], but we report it here for completeness. We want to show that if  $\rho_{ABA'B'}$  is  $\epsilon$ -close in trace norm to a private state  $\gamma_{ABA'B'}(d)$  [see Equation (4.18)] with  $d$ -dimensional key components, then

$$E_{\max}(\rho_{ABA'B'}) \geq \log d + 2 \log(1 - \epsilon/2). \quad (\text{B.1})$$

Notice that a maximally entangled target state is a particular example of private state, so we do not need to study it separately.

The proof is based on the notion of privacy test [169] associated with a private state, that is a projective measurement with projectors  $\{\hat{\Pi}_{ABA'B'}, \hat{\mathbb{1}}_{ABA'B'} - \hat{\Pi}_{ABA'B'}\}$ , where

$$\hat{\Pi}_{ABA'B'} = \hat{U}_{ABA'B'}^{(\text{twist})} (\psi_{AB}(d) \otimes \hat{\mathbb{1}}_{A'B'}) \hat{U}_{ABA'B'}^{(\text{twist})\dagger}. \quad (\text{B.2})$$

Here  $\psi_{AB}$  is a maximally entangled state and the twisting unitary operators are those appearing in the definition of  $\gamma_{ABA'B'}(d)$ . The probability  $p_{\Pi}$  of detecting the outcome associated with the projector  $\hat{\Pi}_{ABA'B'}$  can be bounded as [167, 169, 192]

$$p_{\Pi} = \text{Tr} \left[ \rho_{ABA'B'} \hat{\Pi}_{ABA'B'} \right] \geq \mathcal{F}(\rho_{ABA'B'}, \gamma_{ABA'B'}(d)), \quad (\text{B.3})$$

where  $\mathcal{F}$  represents the Uhlmann fidelity [142]. This probability can be used to bound the sandwiched Rényi divergence  $\tilde{D}_{\alpha}$  [see Equation (4.58)] between  $\rho_{ABA'B'}$  and a generic

separable state  $\sigma_{ABA'B'}$  as

$$\tilde{D}_\alpha(\rho_{ABA'B'}, \sigma_{ABA'B'}) \geq \frac{\alpha}{\alpha - 1} \log_2 p_\Pi + \log_2 d. \quad (\text{B.4})$$

This inequality was obtained in Ref. [167] by exploiting the properties of  $\tilde{D}_\alpha$  and the data-processing inequality. The max-relative entropy of entanglement of  $\rho_{ABA'B'}$  is obtained by optimising the sandwiched Rényi divergence over all separable states in the limit  $\alpha \rightarrow \infty$ , so

$$E_{\max}(\rho_{ABA'B'}) \geq \log_2 p_\Pi + \log_2 d \geq \log_2 \mathcal{F}(\rho_{ABA'B'}, \gamma_{ABA'B'}(d)) + \log_2 d, \quad (\text{B.5})$$

where we also used Equation (B.3). The thesis follows from the Fuchs-van-de-Graaf inequality [201]:

$$\sqrt{\mathcal{F}(\rho_{ABA'B'}, \gamma_{ABA'B'}(d))} \geq 1 - \frac{1}{2} \|\rho_{ABA'B'} - \gamma_{ABA'B'}(d)\|_1, \quad (\text{B.6})$$

which yields Equation (B.1) when used on the right-hand side of Equation (B.5), because  $\|\rho_{ABA'B'} - \gamma_{ABA'B'}(d)\|_1 \leq \epsilon$  by hypothesis.

## B.2 Proof of the upper bound in Proposition 4.2

In order to prove the upper bound appearing in Proposition 4.2, we need to explicitly perform the optimisation appearing in the definition of  $\tilde{E}_{\max}(\mathcal{N}_\lambda^{(\text{ad})})$ , i.e.

$$\tilde{E}_{\max}(\mathcal{N}_\lambda^{(\text{ad})}) \equiv \min_{\substack{\sigma_{AB} \in \text{SEP} \\ \text{Tr}_B[\sigma_{AB}] = \mathbb{1}_A/2}} D_{\max}(\pi_{\mathcal{N}_\lambda^{(\text{ad})}} || \sigma_{AB}) = \min_{\sigma_{AB}} \inf \{x | 2^x \sigma_{AB} - \pi_{\mathcal{N}_\lambda^{(\text{ad})}} \geq 0\}, \quad (\text{B.7})$$

where because of Lemma 4.1, on the rightmost minimisation  $\sigma_{AB}$  has the structure given in Equation (4.76), with  $\gamma = 1 - \alpha$  and  $\delta = 1 - \beta$ . With these assumptions, the condition  $y\sigma_{AB} - \pi_{\mathcal{N}_\lambda^{(\text{ad})}} \geq 0$  can be rewritten as the system of inequalities:

$$\begin{cases} y(1 - \beta) \geq \lambda, \\ y\tilde{\sigma} - \tilde{\pi}_\lambda \geq 0, \end{cases} \quad (\text{B.8})$$

where  $\tilde{\sigma}$  and  $\tilde{\pi}_\lambda$  are  $2 \times 2$  matrices

$$\tilde{\sigma} = \begin{pmatrix} \alpha & \xi e^{i\phi} \\ \xi e^{-i\phi} & \beta \end{pmatrix}, \quad \tilde{\pi}_\lambda = \begin{pmatrix} 1 & \sqrt{1-\lambda} \\ \sqrt{1-\lambda} & 1-\lambda \end{pmatrix}, \quad (\text{B.9})$$

because the Choi state  $\pi_{\mathcal{N}_\lambda^{(\text{ad})}}$  can be expanded in basis  $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$  as

$$\pi_{\mathcal{N}_\lambda^{(\text{ad})}} = \frac{1}{2} \left( \begin{array}{c|c|c} 1 & & \sqrt{1-\lambda} \\ \hline & 0 & \\ \hline & \lambda & \\ \hline \sqrt{1-\lambda} & & 1-\lambda \end{array} \right). \quad (\text{B.10})$$

We now define  $y_1(\lambda, \sigma_{AB})$  and  $y_2(\lambda, \sigma_{AB})$  as the smallest values of  $y$  that satisfy respectively the first and the second inequality appearing in Equation (B.8), and we rewrite the minimisation leading to the upper bound on  $E_{\max}(\mathcal{N}_\lambda^{(\text{ad})})$  as

$$\tilde{E}_{\max}(\mathcal{N}_\lambda^{(\text{ad})}) = \log_2 \min_{\sigma_{AB}} \max\{y_1(\lambda, \sigma_{AB}), y_2(\lambda, \sigma_{AB})\}. \quad (\text{B.11})$$

We can easily show that this quantity is smaller than or equal to  $\log_2(2-\lambda)$  by providing a matrix  $\sigma_{AB}$  of the desired form such that  $\max\{y_1(\lambda, \sigma_{AB}), y_2(\lambda, \sigma_{AB})\} = 2 - \lambda$ . This can be achieved with the choices:

$$\alpha = \frac{1}{2-\lambda}, \quad \beta = 1 - \alpha, \quad \xi = \sqrt{\alpha\beta}, \quad \phi = 0, \quad (\text{B.12})$$

which yield  $y_1 = \lambda(2 - \lambda)$  and  $y_2 = 2 - \lambda$ , as can be verified by directly substituting the values in Equation (B.12) into Equation (B.8). The converse inequality, i.e.

$\tilde{E}_{\max}(\mathcal{N}_\lambda^{(\text{ad})}) \geq \log_2(2 - \lambda)$ , requires some more work. Thanks to the monotonicity of the logarithm and the trivial relation  $\max\{y_1, y_2\} \geq y_2$ , we can bound  $\tilde{E}_{\max}(\mathcal{N}_\lambda^{(\text{ad})})$  from below as

$$\tilde{E}_{\max}(\mathcal{N}_\lambda^{(\text{ad})}) \geq \log_2 \min_{\sigma_{AB}} y_2(\lambda, \sigma_{AB}). \quad (\text{B.13})$$

Hence, we are left with the task of showing that  $\min_{\sigma_{AB}} y_2(\lambda, \sigma_{AB}) \geq 2 - \lambda$ , where the optimisation has to be effectively performed over the parameters  $\alpha, \beta, \xi, \phi$  satisfying the conditions detailed after Equation (4.76), with  $\gamma = 1 - \alpha$  and  $\delta = 1 - \beta$ .

The condition  $y\tilde{\sigma} - \tilde{\pi}\lambda \geq 0$  involves  $2 \times 2$  matrices, and can be rewritten in terms of Pauli matrices  $\boldsymbol{\sigma} = \{\sigma_x, \sigma_y, \sigma_z\}$  as

$$y(\alpha + \beta)(\mathbb{1} + \mathbf{v} \cdot \boldsymbol{\sigma}) - (2 - \lambda)(\mathbb{1} + \mathbf{n} \cdot \boldsymbol{\sigma}) \geq 0, \quad (\text{B.14})$$

where

$$\mathbf{v} = \frac{1}{\alpha + \beta} \begin{pmatrix} 2\xi \cos \phi \\ -2\xi \sin \phi \\ \alpha - \beta \end{pmatrix}, \quad \mathbf{n} = \frac{1}{2 - \lambda} \begin{pmatrix} 2\sqrt{1 - \lambda} \\ 0 \\ \lambda \end{pmatrix}. \quad (\text{B.15})$$

This in turn reduces to

$$y \geq \frac{2 - \lambda}{\alpha + \beta} \frac{2(1 - v \cos \psi)}{1 - v^2} \equiv y_2(\lambda, \sigma_{AB}), \quad (\text{B.16})$$

where  $v = |\mathbf{v}| \leq 1$  and  $\psi$  is the angle between  $\mathbf{v}$  and  $\mathbf{n}$ . Note that the second fraction appearing in Equation (B.16) is always larger than 1, therefore, when  $\alpha + \beta \leq 1$  the condition  $y_2(\lambda, \sigma_{AB}) \geq 2 - \lambda$  holds. On the other hand, if  $1 \leq \alpha + \beta \leq 2$ , we can use the parametrisation:

$$2\xi = \eta(2 - \alpha - \beta) \sin \zeta, \quad \alpha - \beta = \eta(2 - \alpha - \beta) \cos \zeta, \quad (\text{B.17})$$

with  $\eta \in [0, 1]$  and  $\zeta \in [0, \pi]$ . This allows us to conclude because of the following chain of inequalities:

$$\begin{aligned} y_2(\lambda, \sigma_{AB}) &= 2(2 - \lambda) \frac{(\alpha + \beta) - \eta(2 - \alpha - \beta) [\cos(\theta - \zeta) - \sin \theta \sin \zeta (1 - \cos \phi)]}{(\alpha + \beta)^2 - \eta^2(2 - \alpha - \beta)^2} \\ &\geq 2(2 - \lambda) \frac{(\alpha + \beta) - \eta(2 - \alpha - \beta)}{(\alpha + \beta)^2 - \eta^2(2 - \alpha - \beta)^2} \\ &= (2 - \lambda) \frac{2}{2\eta + (1 - \eta)(\alpha + \beta)} \geq (2 - \lambda), \end{aligned} \quad (\text{B.18})$$

where  $\theta = \arctan(2\sqrt{1 - \lambda}/\lambda)$  is the angle describing the direction of  $\mathbf{n}$ .

## B.3 Proof of the lower bound in Proposition 4.2

Here we prove the following lower bound on  $E_{\max}(\mathcal{N}_\lambda^{(\text{ad})})$ :

$$E_{\max}(\mathcal{N}_\lambda^{(\text{ad})}) \geq \min_{\sigma_{AB} \in \text{SEP}} \tilde{D}_{\max}(\pi_{\mathcal{N}_\lambda^{(\text{ad})}} || \sigma_{AB}) = \begin{cases} \log_2 \left( \frac{1}{2}(1 + \sqrt{1 - \lambda})^2 \right), & \text{if } \lambda \leq \frac{\sqrt{5}-1}{2}, \\ \log_2 \left( \frac{1+\lambda}{2\lambda} \right), & \text{if } \lambda \geq \frac{\sqrt{5}-1}{2}. \end{cases} \quad (\text{B.19})$$

Thanks to Lemma 4.1, we can reduce the optimisation over all separable states  $\sigma_{AB}$  which are left unaltered under all possible  $\theta$  rotations, which can be parametrised as in Equation (4.76). The condition  $y\sigma_{AB} - \pi_{\mathcal{N}_\lambda^{(\text{ad})}} \geq 0$  can be explicitly rewritten as

$$\begin{cases} y \geq \frac{\lambda}{\delta}, \\ y \geq \frac{\alpha(1-\lambda) + \beta - 2\sqrt{1-\lambda}\xi \cos \phi}{\alpha\beta - \xi^2}, \end{cases} \quad (\text{B.20})$$

so that

$$\tilde{D}_{\max}(\pi_{\mathcal{N}_\lambda^{(\text{ad})}} || \sigma_{AB}) = \log_2 \max \left\{ \frac{\alpha(1-\lambda) + \beta - 2\xi \cos \phi \sqrt{1-\lambda}}{\alpha\beta - \xi^2}, \frac{\lambda}{\delta} \right\}. \quad (\text{B.21})$$

In what follows, for any fixed  $\lambda$  we will minimise this quantity over the parameters  $\alpha, \beta, \gamma, \delta, \xi, \phi$ , satisfying the constraints detailed after Equation (4.76).

The minimisation in  $\phi$  can be easily performed, with the optimal choice being  $\phi = 0$ . Moreover, for any fixed  $\alpha, \beta, \xi$ , the maximum  $\delta$  (and thus the minimum  $\lambda/\delta$ ) is given by

$$\delta_{\max} = \frac{1}{2} \left( 2 - \alpha - \beta + \sqrt{(2 - \alpha - \beta)^2 - 4\xi^2} \right), \quad (\text{B.22})$$

that is, when  $\gamma\delta$  equals the smallest allowed value  $\xi^2$  and  $\delta > \gamma$ . Notice that this choice implies

$$2\xi = 2\sqrt{\gamma\delta} \leq \gamma + \delta = 2 - \alpha - \beta. \quad (\text{B.23})$$

At this stage, the optimisation problem (without the logarithm) has been reduced to

$$\min_{\alpha, \beta, \xi} \left\{ \max \left[ \frac{\alpha(1-\lambda) + \beta - 2\xi\sqrt{1-\lambda}}{\alpha\beta - \xi^2}, \frac{2\lambda}{2 - \alpha - \beta + \sqrt{(2 - \alpha - \beta)^2 - 4\xi^2}} \right] \mid \begin{array}{l} \alpha, \beta, \xi \geq 0 \wedge \xi^2 \leq \alpha\beta \\ \wedge \alpha + \beta + 2\xi \leq 2 \end{array} \right\}. \quad (\text{B.24})$$

Now we introduce the parameters  $\nu = (\alpha + \beta)/2$  and  $\mu = (\alpha - \beta)/2$ . As  $\alpha > \delta$  always yields a smaller value than the converse choice, we can limit our study to  $\mu \geq 0$  and rewrite the problem in the new parameters:

$$\min_{\nu, \mu, \xi} \left\{ \max \left[ \frac{\nu(2-\lambda) - \lambda\mu - 2\xi\sqrt{1-\lambda}}{\nu^2 - \xi^2 - \mu^2}, \frac{\lambda}{1 - \nu + \sqrt{(1-\nu)^2 - \xi^2}} \right] \mid \begin{array}{l} 0 \leq \mu \leq \sqrt{\nu^2 - \xi^2} \\ 0 \leq \xi \leq \nu \wedge \nu + \xi \leq 1 \end{array} \right\}. \quad (\text{B.25})$$

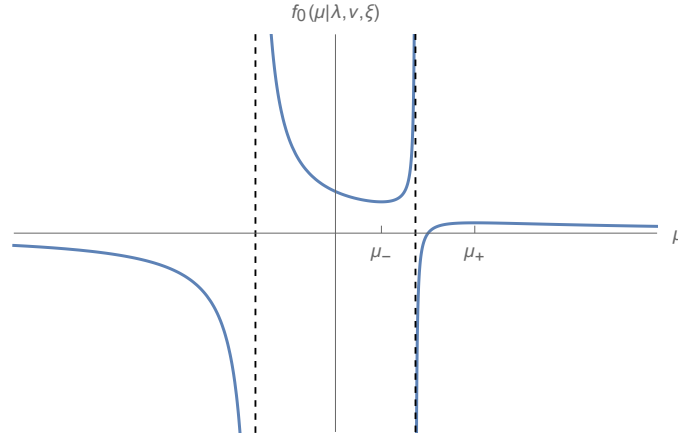


Figure B.1: Typical plot of the function  $f_0(\mu|\lambda, \nu, \xi)$  defined in Equation (B.26).

We can now minimise the first term over  $\mu$ . The value  $\mu_0$  for which the function

$$f_0(\mu|\lambda, \nu, \xi) = \frac{\nu(2 - \lambda) - \lambda\mu - 2\xi\sqrt{1 - \lambda}}{\nu^2 - \xi^2 - \mu^2} \quad (\text{B.26})$$

becomes zero is always larger than  $\sqrt{\nu^2 - \xi^2}$  in the considered region. Together with the asymptotic scaling  $f_0(\mu|\lambda, \nu, \xi) \sim \lambda/\mu$  for  $|\mu| \gg 1$ , this can be used to deduce the qualitative behaviour of  $f_0(\mu|\lambda, \nu, \xi)$ , which is shown in Figure B.1. Let  $\mu_{\pm}(\lambda, \nu, \xi)$  be the zeros of  $\partial_{\mu}f_0(\mu|\lambda, \nu, \xi)$ , with  $\mu_- \leq \mu_+$ :

$$\mu_{\pm}(\lambda, \nu, \xi) = \frac{1}{\lambda} \left[ (2 - \lambda)\nu - 2\sqrt{1 - \lambda}\xi \right] \pm \frac{1}{\lambda} |2\sqrt{1 - \lambda}\nu - (2 - \lambda)\xi|. \quad (\text{B.27})$$

As  $f_0(\mu_-|\lambda, \nu, \xi) \geq f_0(\mu_+|\lambda, \nu, \xi)$ , we can find the desired minimum of  $f_0(\mu|\lambda, \nu, \xi)$  in  $\mu \in [0, \sqrt{\nu^2 - \xi^2}]$  as

$$\min_{\mu} f_0(\mu|\lambda, \nu, \xi) = \max\{f_0(\mu_-), f_0(\mu_+)\} = \max \left\{ \frac{(1 - \sqrt{1 - \lambda})^2}{2(\nu - \xi)}, \frac{(1 + \sqrt{1 - \lambda})^2}{2(\nu + \xi)} \right\}. \quad (\text{B.28})$$

It is worth substituting  $\nu \rightarrow x(1 + y)/2$  and  $\xi \rightarrow x(1 - y)/2$ . In terms of the new variables, the problem after the optimisation in  $\mu$  becomes

$$\min_{x,y} \left\{ \max \left[ \frac{(1 - \sqrt{1 - \lambda})^2}{2xy}, \frac{(1 + \sqrt{1 - \lambda})^2}{2x}, \frac{\lambda}{1 - x\frac{(1+y)}{2} + \sqrt{(1-x)(1-xy)}} \right] \mid \begin{array}{l} 0 \leq x \leq 1 \\ 0 \leq y \leq 1 \end{array} \right\}, \quad (\text{B.29})$$

whose form is suitable to perform the minimisation in  $y$ . Let us label the three functions appearing between square brackets as  $f_1, f_2$  and  $f_3$ . Note that  $f_1$  and  $f_3$  are respectively monotonically decreasing and increasing with  $y$ , with only the first one diverging to infinity for  $y \rightarrow 0$ . If the two functions do not cross each other, i.e., if  $x \leq x_{\text{th}} \equiv (1 - \sqrt{1 - \lambda})/2$ , the minimum over  $y$  is thus obtained by evaluating  $f_1$  in  $y = 1$ , otherwise we need to pick



their intersection point. Explicitly, this can be written as

$$\min_y \{f_1, f_3\} = \begin{cases} \frac{(1-\sqrt{1-\lambda})^2}{2x}, & \text{if } 0 \leq x \leq x_{\text{th}}, \\ f_4(x, \lambda), & \text{if } x_{\text{th}} \leq x \leq 1, \end{cases} \quad (\text{B.30})$$

where

$$f_4(x, \lambda) = \frac{1}{2x^2} \left\{ 8 + x \left[ \left( \frac{1 - \sqrt{1-\lambda}}{\sqrt{\lambda}} \right)^2 - 4 \right] - 4 \sqrt{(1-x) \left[ 4 + x \left( \frac{1 - \sqrt{1-\lambda}}{\sqrt{\lambda}} \right)^2 \right]} \right\}. \quad (\text{B.31})$$

Finally, we can optimise over  $x$ . If  $x \leq x_{\text{th}}$ , we are left with:

$$\min_{x \leq x_{\text{th}}} \max \left\{ \frac{(1 - \sqrt{1-\lambda})^2}{2x}, \frac{(1 + \sqrt{1-\lambda})^2}{2x} \right\} = \frac{(1 + \sqrt{1-\lambda})^2}{2x_{\text{th}}} = f_2(x_{\text{th}}, \lambda). \quad (\text{B.32})$$

On the other hand, when  $x \geq x_{\text{th}}$ , we can apply the same reasoning used for the minimisation over  $y$ . In particular,  $f_2$  and  $f_4$  are respectively monotonically decreasing and increasing with  $x$ , with  $f_2(x_{\text{th}}) \geq f_4(x_{\text{th}})$ , and they have a crossing point only when  $\lambda \geq (\sqrt{5} - 1)/2$ . If there is no crossing, the minimum over  $x$  is given by  $f_2(x = 1, \lambda)$ , which is less than or equal to  $f_2(x_{\text{th}}, \lambda)$  of Equation (B.32). If there is a crossing, instead, the minimum corresponds to the value of the functions at the intersection, which is  $\frac{\lambda+1}{2\lambda}$ . This concludes the proof.