Bosonic systems in quantum information theory

Gaussian-dilatable channels, passive states, and beyond

une dissertation présentée
par
Michael G. Jabbour
à
l'École polytechnique de Bruxelles

en vue de l'obtention du grade de
Docteur
en
Sciences de l'Ingénieur et Technologie

Superviseur :
Nicolas J. Cerf

Université libre de Bruxelles
Bruxelles
Juin 2018

Michael G. Jabbour
Université libre de Bruxelles
École polytechnique de Bruxelles
50 av. F.D. Roosevelt - CP165/59
1050 Bruxelles
Belgique
Email: mjabbour@ulb.ac.be

Thèse de doctorat présentée en séance publique le 18 juin 2018 à l'Université libre de Bruxelles.

Jury : **Gerardo Adesso**, University of Nottingham, Angleterre

Raúl García-Patrón, ULB

Ognyan Oreshkov, Secrétaire, ULB

Jean-Marc Sparenberg, Président du jury, ULB

Andreas Winter, Universitat Autònoma de Barcelona, Espagne

Nicolas J. Cerf, Superviseur, ULB

À F. L.

ستبقى بسمة في قلوبنا ، يا جوجو .

Acknowledgments

Je voudrais avant tout remercier mon promoteur Nicolas Cerf. Merci Nicolas de m'avoir fait confiance en me proposant de rejoindre le QuIC en tant que doctorant. Tu as fait bien plus que me conseiller et me guider dans ma recherche. C'est ton cours de BA3 que j'ai suivi en Polytech qui m'a donné goût à la mécanique quantique, et le fait que tu sois très bon pédagogue y est pour beaucoup. Tu m'as aussi accepté comme mémorant tard durant ma MA2 alors que j'étais un peu perdu et ne savais pas vers quoi me diriger. Je ne regretterai jamais de t'avoir contacté ce jour-là, lorsque tu m'as finalement proposé de m'attaquer à l'étude de la *majorization* en information quantique. J'ai aussi apprécié le fait que tu me laisses tant de liberté dans le choix des sujets de recherche que je voulais explorer.

C'est aussi un plaisir pour moi de remercier les personnes dont j'ai énormément appris au QuIC. Je pense notamment à Evgueni, Ognyan et Raúl. Merci Evgueni d'avoir répondu à toutes mes questions, surtout celles dont le sujet s'écartait un peu de l'information quantique. Thank you Ognyan for all the discussions we had, especially about open quantum systems. I learnt a lot from our exchanges. Merci Raúl de m'avoir appris autant concernant l'optique quantique. J'apprécie le fait que tu aies pris le temps de répondre à mes questions alors que tu étais assez occupé ces derniers temps.

Ces années passées au QuIC n'auraient jamais été aussi agréable sans mes amis Anaelle, Matthieu, Mathieu et Levon. Je dois ajouter que sans toi, Anaelle, j'aurais oublié (ma spécialité) 90% des tâches que je devais effectuer. Par contre, tu n'auras malheureusement pas réussi à me stresser durant l'écriture de ma thèse. En passant, Matthieu, j'ai encore du mal à croire que mon poignet ne se soit jamais cassé. Je suis d'ailleurs ravi (?) d'avoir contribué à ton entraînement d'Aikido.

To the members of FC QuIC, Leo, Shan, Zacharie, Matthieu and Levon, our football matches were pretty fun! Also, Leo, Shan and Zacharie, I got to meet you later, but I really enjoyed your company during my last year at QuIC. Leo and Shan, I look forward to having more debates with you (you know what I am talking about).

Zacharie, ce fut un plaisir de pouvoir superviser ton mémoire. J'espère avoir bien rempli mon rôle et t'avoir suffisamment bien guidé. De mon côté, j'ai été ravi de travailler avec toi. Je trouve que tu as fait un excellent mémoire, et je suis sûr qu'il en sera de même pour ton doctorat. J'espère qu'on continuera à travailler ensemble dans le futur.

Pascale, merci pour ton aide précieuse durant ces années au QuIC. Comme je te l'ai souvent répété, qu'est-ce que j'aurais fait sans toi !

I would also like to thank everybody else at the QuIC. Atul, Jérémie, Luc, Stephan, Uttam, Zoé. Uttam, it was a pleasure working with you, I hope we continue doing research together in the future.

Je tiens à remercier particulièrement John. Sans toi, John, je ne sais pas si je serais actuellement en train de finaliser cette thèse. Il m'est impossible de te dire en quelques lignes à quel point je te suis reconnaissant. Mes années d'études en Polytech ont été beaucoup plus agréable en ta compagnie. Après la MA2, tu as été un des premiers à me motiver à me lancer dans cette aventure que fut le doctorat. Je suis heureux que nous soyons amis après tout ce temps, et après avoir partagé tant d'expériences, tous ces voyages, perdus au fin fond de Viñales à Cuba ou au sommet de Doi Inthanon en Thaïlande, en passant par Lofoten en Norvège… En ce qui concerne ce manuscrit, je te remercie d'être repassé dessus avec tant de soin et d'avoir été si rigoureux lors de tes corrections.

Ma famille sait à quel point je lui suis reconnaissant pour tout, Maman, Papa, Christophe, Lara. Christophe, on l'ouvrira peut-être un jour, ce bar à Cuba...

Fanny, j'ai toujours apprécié toutes nos conversations, de la science à la sociologie, des questions sur la nature des trous noirs à celle de la conscience, je ne me lasserai jamais d'en parler avec toi. Ne t'inquiète pas, je reviendrai te voir souvent après avoir quitté Bruxelles. Mano, à tous les matchs de foot qui s'annoncent, en espérant que ça dure encore des années !

I would also like to thank the members of the jury of my thesis for going through the present manuscript.

Thesis advisor: Nicolas J. Cerf Author: Michael G. Jabbour

Bosonic systems in quantum information theory

Gaussian-dilatable channels, passive states, and beyond

Abstract

The symplectic formalism applied to the phase-space representation of bosonic quantum systems provides us with a powerful mathematical tool for the characterisation of Gaussian states and transformations. As a consequence, quantum information protocols involving the latter are very well understood from a theoretical point of view. Nevertheless, it has become clear in recent years that the use of non-Gaussian resources is necessary in order to perform various crucial information-processing tasks. An illustration of this fact can for instance be found in situations where a Gaussian no-go theorem precludes the use of Gaussian transformations in order to achieve a task involving Gaussian states, such as quantum entanglement distillation, quantum error correction, or universal quantum computation. In the first part of this thesis, we develop a new method based on the generating function of a sequence, which gives rise to an elegant description of intrinsically non-Gaussian objects. Building on the generating function of the matrix elements of Gaussian unitaries in Fock basis, our approach gives access to the multi-photon transition probabilities via unexpectedly simple recurrence equations. The method is developed for Gaussian unitaries effecting both passive and active linear coupling between two bosonic modes. It predicts an interferometric suppression term which generalises the Hong-Ou-Mandel effect for more than two indistinguishable photons impinging on a balanced beam splitter. Furthermore, it exhibits an unsuspected 2-photon suppression effect in optical parametric amplification of gain 2, which originates from the indistinguishability between the input and output photon pairs. Finally, we extend our method to Bogoliubov transformations acting on an arbitrary number of modes. In the second part of this thesis, we introduce a class of Gaussian-dilatable bosonic quantum channels (characterised by a Gaussian unitary in their Stinespring dilation) called passive-environment channels. These channels are interesting from a quantum thermodynamical viewpoint because they correspond to the coupling of a bosonic system with a bosonic environment that is passive in the Fock-basis (that is, no energy can be extracted from it by using unitary transformations) followed by discarding the environment. Making use of the generating function, we provide a description of these channels in terms of Gaussian bosonic channels. We then introduce a new preorder relation called Fock-majorization, which coincides with regular majorization for passive states but also induces another relation in terms of mean boson number, thereby connecting the concepts of energy and disorder of a quantum state. We prove various properties of Fock-majorization, showing in particular that the latter can be interpreted as a relation indicating the existence of a heating or amplifying map between two quantum states. This new preorder relation happens to be relevant

in the context of passive-environment bosonic channels. Indeed, we show that these channels are Fock-majorization-preserving, so that any two input states that obey a Fock-majorization relation are transformed into output states respecting a similar relation. As a consequence, it also implies that passive-environment channels are majorization-preserving over the set of passive states of the harmonic oscillator. The consequences of majorization preservation are discussed in the context of the so-called entropy photon-number inequality. Most of our results being independent of the specific nature of the system under investigation, they could be generalised to other quantum systems and Hamiltonians, providing new tools that may prove useful in quantum information theory. In the last part of our thesis, we lay out a resource theory of local activity for bosonic systems. We introduce a notion of local-activity distance, and compare it with the work that can be extracted from a quantum state under local unitaries assisted by passive global unitaries. With this framework, we hope to connect the area of continuous-variable bosonic channels together with quantum thermodynamics.

Systèmes bosoniques en théorie de l'information quantique

Canaux gaussiens-dilatables, états passifs, et au-delà

Résumé

Le formalisme symplectique appliqué à la représentation des systèmes bosoniques dans l'espace des phases donne accès à un outil mathématique puissant pour la caractérisation des états gaussiens et transformations gaussiennes. Les protocoles d'information quantique impliquant ces derniers sont d'ailleurs très bien compris d'un point de vue théorique. Toutefois, il s'est avéré clair durant ces dernières années que l'utilisation de ressources non-gaussiennes est nécessaire afin d'effectuer des tâches cruciales de traitement de l'information. En effet, certaines tâches — telles que la distillation d'intrication quantique, le codage quantique ou encore le calcul quantique — impliquant des états gaussiens ne peuvent être effectuées avec des transformations gaussiennes. Dans la première partie de cette thèse, nous développons une nouvelle méthode basée sur la fonction génératrice d'une suite qui donne lieu à une description élégante d'objets intrinsèquement non-gaussiens. Se basant sur la fonction génératrice des éléments de matrice d'unitaires gaussiens dans la base de Fock, notre approche donne accès aux probabilités de transition multi-photon via des équations de récurrence étonnamment simples. La méthode est développée pour des unitaires gaussiens produisant des couplages linéaires passifs et actifs entres deux modes bosoniques. Elle prédit un terme d'interférence destructive qui généralise l'effet Hong-Ou-Mandel pour plus de deux photons indistinguables pénétrant dans un diviseur de faisceau équilibré. De plus, elle met en évidence un effet inattendu de suppression de deux photons dans un amplificateur paramétrique optique de gain 2. Cette suppression résulte de l'indistinguabilité entre les paires de photons d'entrée et de sortie. Finalement, nous étendons notre méthode à des transformations de Bogoliubov agissant sur un nombre de modes arbitraire. Dans la seconde partie de cette thèse, nous introduisons une classe de canaux quantiques bosoniques gaussiens-dilatables (caractérisés par un unitaire gaussien dans leur *Stinespring dilation*) appelés canaux à environnement passif. Ces canaux sont intéressants du point de vue de la thermodynamique quantique puisqu'ils correspondent au couplage d'un système bosonique avec un environnement bosonique qui est passif dans la base de Fock (en d'autres termes, il est impossible d'en extraire de l'énergie avec des transformations unitaires), suivi du rejet de l'environnement. Grâce à la fonction génératrice, nous fournissons une description de ces transformations en termes de canaux quantiques bosoniques gaussiens limités par le bruit du vide. Nous introduisons ensuite une nouvelle relation de pré-ordre appelé *majorization* de Fock, qui coïncide avec la *majorization* usuelle pour les états passifs mais induit une autre relation en terme du nombre moyen de bosons, connectant ainsi les concepts d'énergie et de désordre d'un état quantique. Dans ce contexte, nous prouvons des propriétés variées de la *majorization* de Fock et

montrons en particulier que cette dernière peut être interprétée comme une relation indiquant l'existence d'une transformation d'amplification entre deux états quantiques. Cette nouvelle relation de pré-ordre s'avère appropriée dans le contexte des canaux bosonique à environnement passif. En effet, nous montrons que ces canaux conservent la *majorization* de Fock, de sorte que n'importe quels deux états d'entrée obéissant une relation de *majorization* de Fock sont transformés en états de sortie vérifiant une relation similaire. En particulier, cela implique que les canaux à environnement passif préservent la *majorization* pour l'ensemble des états passifs de l'oscillateur harmonique. Les conséquences de la préservation de la *majorization* sont examinées dans le contexte de la *entropy photon-number inequality*. Étant indépendants de la nature spécifique du système étudié, la plupart de nos résultats peuvent être généralisés à d'autres systèmes et hamiltoniens quantiques, donnant lieu à de nouveaux outils qui pourraient s'avérer utiles en théorie de l'information quantique. Dans la dernière partie de notre thèse, nous mettons en place une théorie de l'activité locale pour les système bosoniques. Nous introduisons une notion de distance en terme d'activité locale et la comparons avec le travail qui peut être extrait d'un état quantique avec des unitaires locaux assistés par des unitaires globaux passifs. Le but à long terme est de se baser sur cette théorie afin de connecter les domaines des canaux bosoniques à variables continues et de la thermodynamique quantique.

List of Publications

Work related to the present thesis

[a] Majorization preservation of Gaussian bosonic channels, Michael G. Jabbour, Raúl García-Patrón and Nicolas J. Cerf, New J. Phys. **18**, 073047 (2016).
DOI:10.1088/1367-2630/18/7/073047, arXiv:1512.08225.

[b] Multiphoton interference effects in passive and active Gaussian transformations, Michael G. Jabbour and Nicolas J. Cerf (2018).
arXiv:1803.10734.

[c] Fock majorization in bosonic quantum channels with a passive environment, Michael G. Jabbour and Nicolas J. Cerf (2018).
arXiv:1806.06044.

[d] A resource theory of local activity for bosonic quantum systems, Michael G. Jabbour, Uttam Singh, Evgueni Karpov and Nicolas J. Cerf, in preparation.

Work not related to the present thesis

[e] Interconversion of pure Gaussian states requiring non-Gaussian operations, Michael G. Jabbour, Raúl García-Patrón and Nicolas J. Cerf, Phys. Rev. A **91**, 012316 (2015).
DOI:10.1103/PhysRevA.91.012316, arXiv:1409.8217.

[f] Entropy-power uncertainty relations: towards a tight inequality for all Gaussian pure states, Anaelle Hertz, Michael G. Jabbour and Nicolas J. Cerf, J. Phys. A: Math. Theor. **50**, 385301 (2017).
DOI:10.1088/1751-8121/aa852f, arXiv:1702.07286.

Contents

List of Symbols

Miscellaneous

N	Number of modes
$\mathbb{N}_0$	Set of all natural numbers (including zero)
$\mathbb{R}_{\geq 0}$	Set of all non-negative real numbers
$\mathbb{1}_N$	$N \times N$ identity matrix
$\mathbb{I}$	Identity operator
$\mathbf{\Omega}$	N-mode symplectic form
$\hat{n}$	Number operator
κ	Transmittance of pure-loss channel or gain of quantum-limited amplifier
$\mathbf{p}^{\downarrow}$	Vector containing the elements of $\mathbf{p}$ sorted in non-increasing order
$A^{\downarrow}$	Spherically decreasing symmetric rearrangement of a Borel set A
$f^{\downarrow}$	Spherically decreasing symmetric rearrangement of a measurable non-negative function f
$\succ$	Majorization symbol
$\equiv$	Equivalence in terms of majorization
$\succ_{\text{F}}$	Fock-majorization symbol
$\succ_{\text{T}}$	Catalytic majorization symbol
$f_X \star f_Y$	Convolution of the two functions f_X and f_Y
$P(X)$	Entropy power of the random variable X
$\mathfrak{N}(\rho)$	Entropy photon-number of the quantum state ρ
$\mathcal{T}$	Generating function
g^{BS}	Generating function of amplitudes for the beam splitter
g^{TMS}	Generating function of amplitudes for the two-mode squeezer
f^{BS}	Generating function of probabilities for the beam splitter
f^{TMS}	Generating function of probabilities for the two-mode squeezer
f^{PI}	Generating function of probabilities for the interferometer

States

$\lvert n\rangle$	Number (Fock) state if $n \in \mathbb{N}_0$
$\lvert \alpha\rangle$	Gaussian coherent state if $\alpha \in \mathbb{C}$
$\zeta_{\bar{n}}$	Gaussian thermal state of mean number of photon $\bar{n}$
τ_x	Gaussian thermal state of inverse temperature $-\log x$
$\Gamma_{\mathbf{x}}$	Tensor product of Gaussian thermal states τ_{x_i}, $\mathbf{x} = (x_1, x_2, \ldots)$
$\mathbf{V}_{\mathbf{x}}$	Covariance matrix of $\Gamma_{\mathbf{x}}$ if $\mathbf{x} \in \mathbb{R}^N$
$\mathbf{V}_\rho$	Covariance matrix of a quantum state ρ if ρ is a density matrix
$\lvert \psi_r^{\mathrm{EPR}}\rangle$	Two-mode squeezed vacuum state with squeezing parameter r
$\lvert \varphi_\lambda^{\mathrm{EPR}}\rangle$	Two-mode squeezed vacuum state with $\lambda = \tanh^2 r$
$\lvert \psi_{i,k}^{\mathrm{BS}}\rangle$	Double Fock state $\lvert i, k\rangle$ evolved through a beam splitter
$\lvert \psi_{i,k}^{\mathrm{TMS}}\rangle$	Double Fock state $\lvert i, k\rangle$ evolved through a two-mode squeezer
$\rho^{\downarrow}$	Fock-passive state with the same spectrum as ρ
$P_k^{\downarrow}$	Extremal-Fock-passive state
$\mathbf{\Sigma}(\rho)$	Vector whose elements constitute the spectrum of ρ

Channels

U_η^{BS}	Beam-splitter unitary of transmittance η acting on density matrices
U^{PI}	Passive interferometer unitary acting on density matrices
$\mathcal{U}^{\mathrm{PI}}$	Passive interferometer unitary characterising the evolution of the bosonic field operators
U_λ^{TMS}	Two-mode squeezing unitary of parameter λ acting on density matrices
$\mathcal{B}_\eta^{(\varepsilon)}$	Lossy channel
$\mathcal{B}_\eta^{[k]}$	Extremal passive beam-splitter channel
$\mathcal{B}_\eta$	Pure-loss channel $\mathcal{B}_\eta = \mathcal{B}_\eta^{(0)} = \mathcal{B}_\eta^{[0]}$
$\mathcal{A}_G^{(\varepsilon)}$	Amplifier channel
$\mathcal{A}_G^{[k]}$	Extremal passive two-mode squeezer channel
$\mathcal{A}_G$	Quantum-limited amplifier $\mathcal{A}_G = \mathcal{A}_G^{(0)} = \mathcal{A}_G^{[0]}$
$\tilde{\mathcal{A}}_G^{(\varepsilon)}$	Phase-conjugate channel
$\tilde{\mathcal{A}}_G$	Quantum-limited phase-conjugate channel $\tilde{\mathcal{A}}_G = \tilde{\mathcal{A}}_G^{(0)}$

Amplitudes and probabilities

$b_n^{(i,k)}$	Transition amplitudes $\langle n, m \vert\, U_\eta^{\mathrm{BS}} \,\vert i, k\rangle$ of a beam splitter U_η^{BS}
$a_n^{(i,k)}$	Transition amplitudes $\langle n, m \vert\, U_\lambda^{\mathrm{TMS}} \,\vert i, k\rangle$ of a two-mode squeezer U_λ^{TMS}
$B_n^{(i,k)}$	Transition probabilities $\vert \langle n, m \vert\, U_\eta^{\mathrm{BS}} \,\vert i, k\rangle \vert^2$ of a beam splitter U_η^{BS}
$A_n^{(i,k)}$	Transition probabilities $\vert \langle n, m \vert\, U_\lambda^{\mathrm{TMS}} \,\vert i, k\rangle \vert^2$ of a two-mode squeezer U_λ^{TMS}

List of Figures

List of Tables

1 Introduction

Quantum mechanics provides us with an elegant mathematical framework for the description of the Standard Model. As such, it remains one of the most successful theories ever built and can arguably be labelled as one of the most precisely tested in the history of Science. From solving mysteries such as the one shrouding the phenomenon of black-body radiation to being at the very origin of numerous theories, ranging from quantum computing to superconductivity, the assertion that quantum theory enjoys a plethora of applications is far from being an overstatement. One of the most fruitful theories originating from quantum mechanics can be found in quantum information theory, which can actually be seen as a refinement of classical information theory. In 1948, Claude Shannon published a seminal article in which he established a mathematical theory of information [1], a concept which was not precisely measurable until that time. He brought forward some definitions, as well as two revolutionary theorems characterising the information transmission through any classical communication system. The two key pillars of Shannon's theory are the entropy, a measure of uncertainty or information content, and the channel capacity, which sets a limit on the transmission rate achievable over a noisy communication channel. At that time, a consistent theory of quantum mechanics had already been established, and John von Neumann had long given a definition of the quantum entropy [2]. Still, one had to wait more than 40 years until the combination of the classical theory of information and quantum mechanics resulted in a truly new field focusing on the manipulation of information quanta (qubits). In the mid-'90s, Holevo, Schumacher and Westmoreland [3, 4] further generalised Shannon's communication theory to the quantum case, redefining quanti-

ties such as the classical capacity of quantum channels in the process. They brought forward some entropic quantities, making it possible to compute the capacity of a quantum channel when maximised over all sources.

One could say that quantum information theory comes in two flavours, depending on the way information is chosen to be encoded in a physical system. Indeed, one can adopt a procedure in which information is encoded using either discrete or continuous variables. The state of a quantum mechanical system can always be mathematically described by a density matrix characterised by a discrete eigenspectrum. As such, its eigenvalues can be used in order to store information in some way. This is the paradigm of a system in the framework of discrete quantum information. An example of such a system can readily be found in spin $1/2$ particles, such as an atom, which can be used to register a qubit. On the other hand, the archetype of a continuous-variable quantum system involves bosons, which are described in an infinite-dimensional Hilbert space. Its relevant degrees of freedom are given by the so-called bosonic field operators and corresponding quadrature observables, which admit a continuous eigenspectrum and can be employed to encode continuous information. It is important to understand that, in this case, the spectrum of the density matrix describing the state of the system is still discrete, even though it is infinite dimensional. An example of such a bosonic system is provided by a quantised mode of the electromagnetic field. In the framework of bosonic quantum information, the so-called Gaussian quantum states and transformations play a major role, as they closely model a great amount of systems handled in experimental conditions, especially in quantum optics and atomic physics. Furthermore, powerful analytical tools are available for treating Gaussian quantum systems, based on the symplectic formalism in phase space.

Aside from information theory, another key tool for characterising a quantum system from the point of view of its *disorder* content consists in the mathematical theory of *majorization* [5]. The latter is an algebraic theory which provides a mean to compare two probability distributions in terms of randomness. Since the density matrices describing the states of quantum mechanical systems are characterised by eigenvalues forming discrete probability distributions, majorization theory beautifully extends to the quantum realm. The power of majorization theory can, for instance, be witnessed through its connexion with both classical (Shannon) and quantum (von Neumann) entropies. Indeed, whenever a state ρ majorizes another state σ, denoted as $\rho \succ \sigma$, then σ is more disordered than ρ so that $S(\rho) \leq S(\sigma)$, where S represents the von Neumann entropy. The first application of majorization proper to quantum systems was discovered by Nielsen [6], long after the establishment of the algebraic theory. Indeed, he showed that majorization can be employed in order to compare pure bipartite entangled state or, more precisely, to investigate the possibility to transform a state into another using local operations assisted by classical communication. Since then, majorization theory has proven to be

quite a powerful tool for the study of quantum systems.

When Claude Shannon introduced the notions of entropy and capacity in his mathematical theory of communication, he proposed and partially proved an entropic inequality which allowed him to compute bounds on the capacity of non-Gaussian additive-noise channels. The inequality is based on the entropy power of a random variable, which can be understood as the variance of the normal distribution having the same Shannon differential entropy as the distribution of the original variable. With this in mind, the so-called entropy power inequality asserts that the entropy power of the sum of independent random variables is at least the sum of their entropy powers. In an attempt to compute the classical capacities of certain types of quantum channels acting on bosonic systems, Guha conjectured the so-called entropy photon-number inequality (EPnI) [7], which can be understood as an similar relation to the entropy power inequality, but describing quantum states. Similarly to the entropy power, the entropy photon-number $\mathfrak{N}(\rho)$ of a quantum state ρ describes the mean number of photons of the Gaussian thermal state having the same von Neumann entropy as the original state ρ. If one inputs two quantum states ρ_a and ρ_b in a beam splitter U_η^{BS} of transmittance η, then according to the entropy photon-number inequality, the state $\rho_c = \mathrm{Tr}_2\left[U_\eta^{\mathrm{BS}}\left(\rho_a \otimes \rho_b\right) U_\eta^{\mathrm{BS}\dagger}\right]$ at one output of the beam splitter satisfies

$$\mathfrak{N}(\rho_c) \geq \eta\mathfrak{N}(\rho_a) + (1-\eta)\mathfrak{N}(\rho_b). \tag{1.1}$$

If one defines the map $\Phi_\eta\left[\rho_a, \rho_b\right] = \mathrm{Tr}_2\left[U_\eta^{\mathrm{BS}}\left(\rho_a \otimes \rho_b\right) U_\eta^{\mathrm{BS}\dagger}\right]$, the entropy photon-number inequality can be restated as

$$S\left(\Phi_\eta\left[\rho_a, \rho_b\right]\right) \geq S\left(\Phi_\eta\left[\tau_a, \tau_b\right]\right), \tag{1.2}$$

where τ_a and τ_b are two thermal Gaussian states having the same entropies as ρ_a and ρ_b, respectively. Despite a large research effort, the entropy photon-number inequality remains a conjecture as of today.

In the first part of this thesis, we introduce the theoretical background as well as the various notions on which the present work is based. In Chapter 2, we begin by presenting the mathematical framework of classical and quantum information theory, focusing on the concept of disorder through the notions of entropy and majorization. We assume that the reader is familiar with the basic notions of quantum mechanics. Chapter 3 is dedicated to the description of bosonic quantum systems, with a particular emphasis on Gaussian bosonic systems. We end the first part of the manuscript with Chapter 4, in which we present both the entropy power and entropy photon-number inequalities in more details, introducing continuous majorization in the context of the former. This first part does not contain any of the new results introduced in

the context of this thesis. Experienced researchers in quantum information in general can skip Chapter 2, while researchers familiar with continuous-variable quantum information need not read Chapter 3. One of the reasons we have chosen to write a detailed introductory part is in the hope that it will serve future students in the field.

Bosonic quantum systems are characterised by an infinite-dimensional Hilbert space. As a consequence, their physical states are described by density operators with an infinite eigenspectrum, making the treatment of some theoretical problems involving bosonic systems rather involved. In some interesting cases, this difficulty can be circumvented by means of the symplectic formalism applied to the phase-space description of quantum systems, whose central object is the Wigner function. The latter provides us with a representation of quantum states which happens to be completely equivalent to the characterisation provided by density matrices. Whenever one has to deal with Gaussian states and transformations, the symplectic formalism in phase space provides us with a powerful mathematical tool for their study, based on finite-dimensional matrices. However, as soon as either of the state or the transformation is non-Gaussian, the analytical description based on the symplectic formalism becomes powerless. Nevertheless, the investigation of non-Gaussian systems is clearly imperative, since it has been demonstrated in recent years that non-Gaussian resources are essential in order to perform various information-theoretic tasks, as depicted for instance by several Gaussian no-go theorems [8–10]. In Chapter 5 of this thesis, we develop a theoretical framework that is particularly suited to Gaussian unitaries describing both the passive and active linear coupling between bosonic modes (*i.e.*, all Bogoliubov transformations). Our technique relies on the generating function of Gaussian matrix elements in Fock space, which can be expressed in a simple form involving Gaussian states and the symplectic formalism, while it enables accessing intrinsically non-Gaussian features such as multi-photon transition probabilities. Our method also brings forth an elegant relation between the two fundamental two-mode Gaussian unitaries, *i.e.*, the beam splitter and the two-mode squeezer. We demonstrate that either one of the two Gaussian unitaries can be obtained by performing a *partial time reversal* on the other, a fact which can be used in order to describe the behaviour of one of the objects by exploiting knowledge about the other.

In Chapter 6, we make use of our method in order to investigate quantum interferences, which are a cornerstone of quantum physics. Over the last years, there has been a vigorous activity on harnessing multi-mode multi-photon interferences as it may be a key for implementing future quantum technologies with photonic integrated devices, see *e.g.* [11]. This is also significant in connection with the boson sampling paradigm [12], which builds on the computational hardness of simulating the coherent propagation of many identical bosons through a multimode linear interferometer and holds the promise of substantiating the advantage of quantum com-

puters [13–16]. More generally, this has led to a revived interest for quantum interferometry going beyond the celebrated Hong-Ou-Mandel effect [17], *e.g.*, the generalised bunching effect in linear networks [18], the signatures of nonclassicality in a multimode interferometer [19], the observation of intrinsically 3-photon interference [20, 21], or the suppression laws in a 8-mode optical Fourier interferometer [22]. The technique we develop in Chapter 5 allows us to derive fundamental recurrence relations for the transition probabilities in a beam splitter and in a two-mode squeezer. In particular, it exhibits a suppression term that generalises the Hong-Ou-Mandel effect to many photons. Remarkably, applying this tool to active transformations, we find an analogue suppression effect that had been left unnoticed in an optical amplification medium of gain 2. The results we present in Chapters 5 and 6 are based on [b].

Another interesting problem that cannot be addressed with the symplectic formalism concerns the characterisation of non-Gaussian channels. Such a theoretical description turns out to be necessary, as several crucial quantum information processes cannot be performed solely with Gaussian transformations. With this in mind, we introduce in Chapter 7 of this thesis a new class of bosonic quantum channels. In order to take advantage of the methods developed earlier and the knowledge acquired in the context of Gaussian unitaries, we focus on channels characterised in their Stinespring dilation by such unitaries, as well as an environment that is passive. Passive states are those quantum states from which no work can be extracted using unitary transformations, making them ubiquitous in theories in which the concept of energy plays a crucial role, such as quantum thermodynamics. We call our non-Gaussian channels passive-environment Gaussian-dilatable channels. By exploiting the generating function of some specific passive-environment channels, we show how any of these maps can be written in terms of quantum-limited Gaussian bosonic channels. This provides us with a way to study passive-environment channels starting from the formalism of Gaussian maps. For instance, we make use of this method to show that the dual map of a passive-environment channel based on a beam splitter is proportional to a similar passive-environment channel based on a two-mode squeezer.

The theory of majorization was originally introduced in the framework of continuous-variable quantum systems by Guha in the context of the EPnI [7]. The idea was that in the specific case in which the environment ρ_b is thermal while the signal ρ_a has an entropy equal to zero, Equation (1.2) can be seen as a consequence of a majorization relation between the two outputs corresponding to state ρ_a and the vacuum state. In an attempt to generalise this idea, we conjecture what we call the precursor of the EPnI, which states that

$$\Phi_\eta\left[\rho_a^{\downarrow}, \rho_b^{\downarrow}\right] \succ \Phi_\eta\left[\rho_a, \rho_b\right], \tag{1.3}$$

where the map Φ_η is the same as the one appearing in Equation (1.2), and $\rho_a^{\downarrow}$ and $\rho_b^{\downarrow}$ are the

passive state with the same eigenspectra as ρ_a and ρ_b, respectively. If one fixes the environment ρ_b to be in a passive state, Equation (1.3) can be restated in terms of any passive-environment channel $\mathcal{C}^{\downarrow}$ as

$$\mathcal{C}^{\downarrow}\left[\rho^{\downarrow}\right] \succ \mathcal{C}^{\downarrow}\left[\rho\right]. \tag{1.4}$$

Inspired by the conjectured majorization relation of Equation (1.4), we ask the question whether a majorization relation is preserved in any passive-environment channel, meaning that if two states verify some majorization relation, they will verify a similar relation at the output of the channel. This happens not to be verified in general, as one can easily find counter-examples to such a statement. Nevertheless, we describe another fundamental majorization-like relation which is preserved through a passive-environment channel. To this aim, we introduce in Chapter 8 a new preorder relation on quantum states called *Fock-majorization* and denoted as $\succ_{\mathrm{F}}$. We show several properties of the latter and demonstrate that the existence of such a relation between two states implies the existence of an *amplifying*, or *heating* map connecting the two. As a consequence, it happens to be closely connected to the concept of energy of a quantum state. We then prove in Chapter 9 that Fock-majorization is precisely the relation that is preserved through any passive-environment channel, including of course Gaussian bosonic channels. In other words, we show that

$$\rho \succ_{\mathrm{F}} \sigma \quad \Rightarrow \quad \mathcal{C}^{\downarrow}[\rho] \succ_{\mathrm{F}} \mathcal{C}^{\downarrow}[\sigma], \tag{1.5}$$

which is to be compared with Equation (1.4). The two states ρ and $\rho^{\downarrow}$ are equivalent from the point of view of majorization, in the sense that both $\rho \succ \rho^{\downarrow}$ and $\rho \prec \rho^{\downarrow}$ are verified. In the context of Fock-majorization however, the two states are ordered, *i.e.*, $\rho^{\downarrow} \succ_{\mathrm{F}} \rho$. As a consequence, Equation (1.5) implies $\mathcal{C}^{\downarrow}[\rho^{\downarrow}] \succ_{\mathrm{F}} \mathcal{C}^{\downarrow}[\rho]$, which is also implied by Equation (1.4) in this particular case. Nonetheless, Equation (1.5) deals with all quantum states and does not only concern passive ones. Furthermore, it implies in particular that passive-environment channels preserve majorization among the set of passive states. We conclude the chapter by discussing the implications of Equation (1.5) in the context of the entropy photon-number inequality. The results we present in Chapters 8 and 9 are based on [a] and [c].

Finally, having developed a method for the characterisation of Gaussian unitaries applied to non-Gaussian states and having studied passive states of the harmonic oscillator in the context of passive-environment channels and Fock-majorization, we undertake the establishment of a non-trivial resource theory for bosonic quantum systems in which passive states are free. In Chapter 10, we outline the guiding lines of such a theory, in which we also choose passive Gaussian transformations, or passive interferometers, to be free operations. As a consequence, our free states are given by products of single-mode passive states transformed by any passive Gaussian unitary. In doing so, our goal is to compare our theory to one in which the central quantity

would be the work $W_{\max}^{\mathrm{l,PG}}$ which is extractable *locally* using unitary transformations, but assisted by passive Gaussian *global* unitaries. To achieve this goal, we introduce a notion of local-activity distance A_l of a state, which is the relative entropy distance of the latter to the closest free state as we defined earlier. We then compare this new quantity to the work $W_{\max}^{\mathrm{l,PG}}$. Our hope is to develop a resource theory which would connect the frameworks of continuous-variable quantum information theory and quantum thermodynamics by building on the notions of Gaussian unitaries and passive states. The results presented in Chapter 10 are based on a paper in preparation [d].

I

Preliminaries

You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one really knows what entropy really is, so in a debate you will always have the advantage.

John von Neumann to Claude Shannon,
Scientific American (1971) [23].

2

Disorder in information theory

Entropy may be understood as a measure of the disorder, uncertainty, or randomness in a physical system. As such, it is one of the most fundamental quantities derived in physics. Depending on its exact interpretation, the entropy plays a critical role in various fields of physics, from statistical mechanics and thermodynamics to classical and quantum information theory. Furthermore, the definition of entropy can be adapted to the different mathematical frameworks describing the physics of systems, characterising discrete or continuous objects, classical or quantum. Although the entropy happens to to be the only function satisfying some specific chosen properties altogether, it can still be extended to general measures which characterise the disorder in a system. This line of thought gives rise to the algebraic theory of majorization, which provides a pre-order among vectors that proves to be more general than entropic inequalities.

In this chapter, we introduce the mathematical framework of information theory, in which entropy is viewed as a measure of the concept of information contained in a physical system. We present several entropy-like functionals in the process, such as the so-called Rényi entropies. By doing so, we attempt to connect these different quantities to the notions of disorder and uncertainty, specifically through the theory of majorization. In Section 2.1, we begin by exploring the realm of discrete information theory, in which random variables are associated to discrete probability distributions. Section 2.2 is concerned with Shannon's theory in the framework of variables having a continuous probability density. It might be worthwhile to note that the concept of majorization for continuous variables has, to our knowledge, never been explored in the context of quantum information theory, to which this manuscript is dedicated. Still, an

important part of the research we carried out was exactly about this notions of continuous majorization, even though we chose not to present it in this manuscript. Furthermore, continuous majorization is interesting as such, which is why we explore it in Section 2.2. Finally, we introduce von Neumann's entropy in Section 2.3, along with the corresponding notions of quantum information theory. Majorization happens to play a crucial role in the quantum theory, specifically as a consequence of its relation with local operations and classical communication. This will lead us to present notions related to the latter, such as measures of entanglement.

This chapter is based on the very well-written book on majorization [5]. We also draw inspiration from [24] and [25]. Note that, in order to be consistent will all the existing literature, we write majorization instead of *majorisation*.

2.1 Discrete probability distributions

2.1.1 Shannon entropy

We begin by introducing the concept of Shannon entropy, a measure of the uncertainty of a random variable. Consider a discrete random variable X, whose realisations x belong to an alphabet $\mathcal{X}$, and define its probability mass function as $p(x)$. The information content $i(x)$ of a particular realisation x of $\mathcal{X}$ can be seen as the measure of the surprise one has upon learning the outcome of a random experiment [25],

$$i(x) = -\log p(x), \tag{2.1}$$

where the logarithm is chosen to be base two, meaning that the measure is done in bits. This choice of the information content is, of course, not random. It happens to behave as hoped. Indeed, it is lower for higher probability events, which tend to be expected, while it is higher for lower probability events, which tend to surprise us. Furthermore, it is additive when taking products of probabilities, which is expected from an informational measure. However, $i(x)$ is obviously a "one-shot" quantity, and cannot be a measure of the uncertainty of the random variable X as such. In order to capture this, one actually needs to take the expected value of the information content $i(X)$,

$$\mathbb{E}\{i(X)\} = -\sum_{x\in\mathcal{X}} p(x)\log p(x). \tag{2.2}$$

This is exactly the entropy $H(X)$ of the variable X.

Definition 1 (Shannon entropy)**.** *Consider a variable X whose realisations x belong to an alphabet*

$\mathcal{X}$, with probability mass function $p(x)$. The Shannon entropy of X is defined as

$$H(X) = -\sum_{x\in\mathcal{X}} p(x)\log p(x). \tag{2.3}$$

Note that there is no problem in the definition of the entropy if we rely on the fact that

$$\lim_{\varepsilon\to 0} \varepsilon\log\varepsilon = 0. \tag{2.4}$$

We already used the word uncertainty to characterise the entropy. There are actually two ways to view the latter. It can either be seen as the amount of information one gains upon learning the value of X (hence the measure of surprise after some experiment), or the amount of uncertainty one has about the variable X before the experiment.

The entropy was introduced by Shannon in his seminal work of 1948 [1], but it was already a well-established quantity in several other fields of physics at that time, even though it had never been connected to the concept of information. The word entropy was actually created by Rudolph Clausius around 1864, in the work that led him to postulate that the entropy of a closed system cannot decrease, which is nowadays known as the second law of thermodynamics [26]. In the context of his work, Clausius had therefore already associated his entropy to the concept of disorder of a thermodynamical system [27]. This was clarified by Ludwig Boltzmann in the following years, when he gave a probabilistic view of the concept of disorder of a system.

Shannon's entropy has several interesting properties, which we list hereafter.

Property 1 (Positivity)**.** *The Shannon entropy of any random variable X is non-negative, i.e.,*

$$H(X) \geq 0. \tag{2.5}$$

This is expected, since the entropy is related to the measure of uncertainty of X.

Property 2 (Concavity)**.** *The Shannon entropy of any random variable X is concave in its probability density $p(x)$.*

This means that if one mixes two random variables, the resulting variable's uncertainty will be greater than the expected uncertainty of the two initial variables. It simply reflects the fact that the process of mixing creates uncertainty.

Property 3 (Symmetry)**.** *The Shannon entropy is invariant under permutations of the realisations of any random variable X.*

This seemingly trivial fact is actually of importance, since it expresses that the entropy of X does not depend on the values of the realisations x of X, unlike other measures of uncertainty, like the variance for instance. Since we are interested in entropy-like measures in this work, we

will forget about the alphabet $\mathcal{X}$ from now on, focusing on the probability density of X instead. As a consequence, we will consider vectors of probabilities $\mathbf{p} \in \mathbb{R}^n$, whose elements p_i are non-negative and sum to one, i ranging from one to the number of elements n of the vector $\mathbf{p}$. Furthermore, we will define the entropy as

$$H(\mathbf{p}) = -\sum_{i=1}^{n} p_i \log p_i, \tag{2.6}$$

taking it to be a functional of the vector of probabilities, instead of the variable X itself. Note that we will often label the vector of probabilities $\mathbf{p}$ simply as a probability distribution in the following sections.

Since we will be using it later on, let us introduce another concept closely related to the Shannon entropy. The so-called relative entropy is a measure of the distance between two probability distributions [28].

Definition 2 (Relative entropy)**.** *The relative entropy, or Kullback-Leibler divergence between two probability distributions* $\mathbf{p} \in \mathbb{R}^n$ *and* $\mathbf{q} \in \mathbb{R}^n$ *is defined as*

$$D\left(\mathbf{p}||\mathbf{q}\right) = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i}. \tag{2.7}$$

One could describe the relative entropy as an expected log-likelihood ratio of the probabilities $\mathbf{p}$ and $\mathbf{q}$. According to its definition, it can become infinite when the support of $\mathbf{p}$ is not contained in the support of $\mathbf{q}$, or in other words, when $p_i \neq 0$ but $q_i = 0$ for some i between 1 and n. Similarly to Equation (2.4), we use the conventions $0 \log \frac{0}{0} = 0$, $0 \log \frac{0}{q_i} = 0$ and $p_i \log \frac{p_i}{0} = \infty$ for any p_i and q_i. An important property of the relative entropy concerns its positivity.

Property 4 (Positivity)**.** *The relative entropy between any two probabilities* $\mathbf{p} \in \mathbb{R}^n$ *and* $\mathbf{q} \in \mathbb{R}^n$ *is non-negative, i.e.,*

$$D\left(\mathbf{p}||\mathbf{q}\right) \geq 0. \tag{2.8}$$

It is equal to zero if and only if $\mathbf{p} = \mathbf{q}$.

Despite this, the relative entropy is not a true distance in the mathematical sense, as it is not symmetric (in its two arguments $\mathbf{p}$ and $\mathbf{q}$), and does not satisfy the triangle inequality. Unlike the Shannon entropy, it is a convex quantity.

Property 5 (Joint convexity)**.** *The relative entropy is jointly convex; that is, if* $\left(\mathbf{p}^{(1)}, \mathbf{q}^{(1)}\right)$ *and* $\left(\mathbf{p}^{(2)}, \mathbf{q}^{(2)}\right)$ *are two pairs of probability distributions, then*

$$D\left(\lambda\mathbf{p}^{(1)} + (1-\lambda)\mathbf{p}^{(2)}||\lambda\mathbf{q}^{(1)} + (1-\lambda)\mathbf{q}^{(2)}\right) \leq \lambda D\left(\mathbf{p}^{(1)}||\mathbf{q}^{(1)}\right) + (1-\lambda) D\left(\mathbf{p}^{(2)}||\mathbf{q}^{(2)}\right), \tag{2.9}$$

for all λ *such that* $0 \leq \lambda \leq 1$.

2.1.2 Rényi entropies

Until now, we focused on the best-known entropy, the Shannon entropy. It is however in our interest for the present dissertation to introduce other specific measures of uncertainty. The Rényi entropies, which generalise the Shannon entropy, will be crucial in understanding the fundamental relation between entropies and the theory of majorization. These functionals were put forward by Alfréd Rényi in his seminal paper of 1961 [29] in which he was investigating other approaches to derive the Shannon entropy.

Definition 3 (Rényi entropy)**.** *For any vector of probabilities* $\mathbf{p} \in \mathbb{R}^n$*, the Rényi entropy of order* $a \in (0,1) \cup (1,\infty)$ *is defined as*

$$H_a(\mathbf{p}) = \frac{1}{1-a} \log \left(\sum_{i=1}^{n} p_i^a \right). \tag{2.10}$$

The Shannon entropy of $\mathbf{p}$ can be recovered from the Rényi entropies of $\mathbf{p}$, as intended by Alfréd Rényi. This can be done by taking the limit

$$\lim_{a \to 1} H_a(\mathbf{p}) = H(\mathbf{p}). \tag{2.11}$$

From now on, we will extend the definition of H_a to $a = 1$, by setting

$$H_1(\mathbf{p}) := H(\mathbf{p}). \tag{2.12}$$

Apart from this connection, there are actually other interesting special cases of Rényi's uncertainty function. The Hartley entropy, or max-entropy,

$$H_0(\mathbf{p}) := \lim_{a \to 0} H_a(\mathbf{p}) = \log n, \tag{2.13}$$

represents the information revealed after picking a sample from a set of finite length n, uniformly at random. As such, it coincides with all the Rényi entropies (as well as the Shannon entropy) in the case of a uniform probability distribution. Similarly, the min-entropy

$$H_\infty(\mathbf{p}) := \lim_{a \to \infty} H_a(\mathbf{p}) = -\log \max_i p_i, \tag{2.14}$$

is yet another way to evaluate the uncertainty contained in a variable. These two extreme cases are of particular interest because of the behaviour of the Rényi entropy when its order a changes. This is encompassed in the following property.

Property 6. *The Rényi entropy of a fixed probability distribution* $\mathbf{p}$ *is non-increasing in its order* a.

This actually means that the two extremal values for the Rényi entropies of a fixed probability distribution are given by the min-entropy and the max-entropy, as their names actually suggest. As an obvious consequence, they constitute bounds for the more interesting Shannon entropy. We now list some useful properties of the Rényi entropies.

Property 7 (Positivity). *The Rényi entropies of any probability distribution* $\mathbf{p}$ *are positive for all* $\alpha \geq 0$*, i.e.,*

$$H_\alpha(\mathbf{p}) \geq 0, \quad \forall \alpha \geq 0. \tag{2.15}$$

We already know that the Shannon entropy is always positive, meaning that this property simply extends to all Rényi entropies. On the other hand, the property of concavity of the former is not simply transferred to all $\alpha \geq 0$ [30].

Property 8 (Concavity). *The Rényi entropies of any probability distribution* $\mathbf{p}$ *are concave in* $\mathbf{p}$ *for* $\alpha \in [0, 1]$.

This property is not verified for $\alpha > 1$. Rényi entropies of these orders are neither convex nor concave in general. This actually makes them less likely to be chosen as uncertainty measures, since concavity would be expected from such functionals. Nevertheless, we will show later that all Rényi entropies are still good candidates for measures of disorder, a concept closely related to uncertainty. Finally, we just mention that all Rényi entropies are also symmetrical, which can obviously be seen from their definition, but is still an important fact.

Property 9 (Symmetry). *The Rényi entropies of the probability distribution* $\mathbf{p}$ *are invariant under permutations of* $\mathbf{p}$.

One can always generalise Definition 2 to Rényi entropies. This can be done as follows [29].

Definition 4 (Rényi divergence). *The Rényi divergence of order* $\alpha \in [0, 1) \cup (1, \infty)$ *between two probability distributions* $\mathbf{p} \in \mathbb{R}^n$ *and* $\mathbf{q} \in \mathbb{R}^n$ *is defined as*

$$D_\alpha(\mathbf{p}||\mathbf{q}) = \frac{1}{\alpha - 1} \log \left(\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha} \right). \tag{2.16}$$

It becomes infinite if $\alpha > 1$ and if the support of $\mathbf{p}$ is not a subset of the support of $\mathbf{q}$. Like the relative entropy, it has some useful properties [31].

Property 10 (Positivity). *The Rényi divergence of order* $\alpha \in [0, 1) \cup (1, \infty)$ *between any two probabilities* $\mathbf{p} \in \mathbb{R}^n$ *and* $\mathbf{q} \in \mathbb{R}^n$ *is non-negative, i.e.,*

$$D_\alpha(\mathbf{p}||\mathbf{q}) \geq 0. \tag{2.17}$$

For $\alpha > 0$*, it is equal to zero if and only if* $\mathbf{p} = \mathbf{q}$*. For* $\alpha = 0$*,* $D_\alpha(\mathbf{p}||\mathbf{q}) = 0$ *if and only if the support of* $\mathbf{q}$ *is a subset of the support of* $\mathbf{p}$.

It is jointly convex for some a.

Property 11 (Joint convexity)**.** *The Rényi divergence of order $a \in [0, 1)$ is jointly convex; that is, if $\left(\mathbf{p}^{(1)}, \mathbf{q}^{(1)}\right)$ and $\left(\mathbf{p}^{(2)}, \mathbf{q}^{(2)}\right)$ are two pairs of probability distributions, then*

$$D_a\left(\lambda\mathbf{p}^{(1)} + (1-\lambda)\mathbf{p}^{(2)} || \lambda\mathbf{q}^{(1)} + (1-\lambda)\mathbf{q}^{(2)}\right) \leq \lambda D_a\left(\mathbf{p}^{(1)} || \mathbf{q}^{(1)}\right) + (1-\lambda) D_a\left(\mathbf{p}^{(2)} || \mathbf{q}^{(2)}\right),$$

for all λ such that $0 \leq \lambda \leq 1$.

The Rényi divergence happens to be a crucial tool with various applications which can be found in, for instance, hypothesis testing and multiple source adaptation [31].

2.1.3 Majorization relations and Schur-convex functions

The Rényi entropies (including the Shannon entropy) we have presented until now can be seen as functionals which measure the uncertainty of a probability distribution $\mathbf{p} \in \mathbb{R}^n$. We explained that this uncertainty is related to the surprise one would gain upon learning the value of a realisation of some alphabet $\mathcal{X}$ whose random variable X had a probability mass function given by $\mathbf{p}$. The functionals all happen to be symmetric. Furthermore, some of them are concave, but not all. We will now see that they are however all related to the concept of disorder, through the mathematical theory of majorization [32], which provides a means to compare two probability distributions in terms of disorder or randomness. Before clarifying what we mean by disorder, let us begin by introducing the algebraic theory of majorization. In order to do so, we first need to fix some notations. Let $\mathbf{p}$ be a probability distribution vector. We will denote by $\mathbf{p}^{\downarrow}$ the vector containing the elements of $\mathbf{p}$ sorted in non-increasing order. With this in mind, we give the definition of a majorization relation [5].

Definition 5 (Majorization)**.** *Let $\mathbf{p} \in \mathbb{R}^n$ and $\mathbf{q} \in \mathbb{R}^n$ be two vectors of dimension n. We say that $\mathbf{p}$ majorizes $\mathbf{q}$, i.e. $\mathbf{p} \succ \mathbf{q}$, if*

$$\sum_{i=1}^{k} p_i^{\downarrow} \geq \sum_{i=1}^{k} q_i^{\downarrow}, \quad k = 1, \ldots, n, \tag{2.18}$$

with equality when $k = n$.

Let us give a few remarks about the last definition.

Remark 1. The definition of majorization can be applied to any two vectors whose elements are real. The vectors do not necessarily have to be probability distributions.

Remark 2. If the two vectors which are being compared have a different number of elements, we may always append zeros to the vector with the lowest dimension in order to compare them in terms of majorization.

Remark 3. If $\mathbf{p}$ and $\mathbf{q}$ are probability distributions, the equality for $k = n$ in Equation (2.18) is always satisfied.

It is important to realise that majorization only provides a pre-order on vectors, in the sense that if $\mathbf{p} \not\succ \mathbf{q}$, this does not necessarily mean that $\mathbf{p} \prec \mathbf{q}$. When both $\mathbf{p} \not\succ \mathbf{q}$ and $\mathbf{p} \not\prec \mathbf{q}$ are satisfied, $\mathbf{p}$ and $\mathbf{q}$ are said to be incomparable. When two vectors $\mathbf{p}$ and $\mathbf{q}$ satisfy both $\mathbf{p} \succ \mathbf{q}$ and $\mathbf{p} \prec \mathbf{q}$, we will say that they are equivalent in terms of majorization, and will denote this by $\mathbf{p} \equiv \mathbf{q}$. It will simply mean that they have the same spectrum.

We mentioned earlier that majorization was related both to uncertainty measures like the Rényi entropies and to the concept of disorder. We may now explain its connection to the latter. This can be understood using the notion of doubly-stochastic, or bistochastic matrix.

Definition 6 (Bistochastic matrix)**.** *A matrix $\mathbf{D} \in \mathbb{R}^{n\times n}$ is said to be bistochastic if all its elements are non-negative, and if its columns and rows all sum to one, i.e.,*

$$D_{ij} \geq 0 \; \forall i,j = 1,\ldots,n, \quad \sum_{i=1}^{n} D_{ij} = 1, \; \forall j = 1,\ldots,n, \quad \sum_{j=1}^{n} D_{ij} = 1, \; \forall i = 1,\ldots,n.$$

The set of bistochastic matrices of a given dimension is convex and its extremal points are given by permutation matrices of the same dimension. Consequently, we have the following theorem.

Theorem 1 (Birkhoff's theorem)**.** *Any doubly-stochastic matrix $\mathbf{D} \in \mathbb{R}^{n\times n}$ can be decomposed as a convex combination of permutation matrices, i.e.,*

$$\mathbf{D} = \sum_{k} \omega_k \mathbf{\Pi}_k, \tag{2.19}$$

where $\{\omega_k\}$ represents a probability distribution and the $\mathbf{\Pi}_k$ are n-dimensional permutation matrices.

Any bistochastic matrix can therefore be written as a convex mixture of permutation matrices.

In order to understand why majorization allows one to compare probability distributions in terms of disorder, let us introduce an alternative way of detecting majorization, which is given by the following theorem [33].

Theorem 2 (Hardy, Littlewood and Pólya)**.** *Let $\mathbf{p} \in \mathbb{R}^n$ and $\mathbf{q} \in \mathbb{R}^n$ be two probability distributions. $\mathbf{p}$ majorizes $\mathbf{q}$ if and only if there exists a bistochastic matrix $\mathbf{D}$ such that*

$$\mathbf{q} = \mathbf{D}\mathbf{p}. \tag{2.20}$$

If we combine Theorem 2 with Birkhoff's theorem, we end up with the fact that, for any two probability distributions $\mathbf{p} \in \mathbb{R}^n$ and $\mathbf{q} \in \mathbb{R}^n$, $\mathbf{p}$ majorizes $\mathbf{q}$ if and only if we can relate them

through a mixture of permutations, *i.e.*,

$$\mathbf{p} \succ \mathbf{q} \quad \Leftrightarrow \quad \exists \{\omega_k\} \text{ s.t. } \mathbf{q} = \sum_k \omega_k \mathbf{\Pi}_k \mathbf{p}, \tag{2.21}$$

where $\{\omega_k\}$ is a probability distribution and the $\mathbf{\Pi}_k$ are n-dimensional permutation matrices. This last equation clearly shows the relation between disorder and majorization. Indeed, we see that if $\mathbf{p}$ majorizes $\mathbf{q}$, then $\mathbf{q}$ can be obtained by applying random permutations to $\mathbf{p}$, making $\mathbf{q}$ more disordered than $\mathbf{p}$.

As we mentioned already, majorization theory is closely related to entropy-like measures. They, or rather their negatives, are actually part of a larger set of functions which preserve the ordering of majorization. These functions, which we thereby define, where introduced by Schur in 1923 [34].

Definition 7 (Schur-convex function)**.** *A real-valued function G defined on a set $\mathcal{R} \in \mathbb{R}^n$ is said to be Schur-convex on $\mathcal{R}$ if*

$$\mathbf{p} \succ \mathbf{q} \quad \text{on } \mathcal{R} \quad \Rightarrow \quad G(\mathbf{p}) \geq G(\mathbf{q}). \tag{2.22}$$

The negative of a Schur-convex function is called Schur-concave. As their name would suggest, some Schur-convex functions can be built using convex functions. This is the content of the following theorem [33, 34].

Theorem 3. *If $g : \mathbb{R} \to \mathbb{R}$ is convex, then*

$$G(\mathbf{p}) = \sum_{i=1}^{n} g(p_i) \tag{2.23}$$

is Schur-convex on $\mathbb{R}^n$.

The function $g(x) = x \log x$ is easily shown to be convex. Combining this fact with Theorem 3 allows us to conclude that the Shannon entropy is Schur-concave. This implies the following corollary.

Corollary 1. *If $\mathbf{p} \in \mathbb{R}^n$ and $\mathbf{q} \in \mathbb{R}^n$ are two probability distributions,*

$$\mathbf{p} \succ \mathbf{q} \quad \Rightarrow \quad H(\mathbf{p}) \leq H(\mathbf{q}). \tag{2.24}$$

Corollary 1 demonstrates the clear connection between the Shannon entropy and disorder. This is consistent with the fact that the process of mixing increases the Shannon entropy. As a matter of fact, Corollary 1 can be strengthened using the relative entropy between the two probabilities [35].

Theorem 4. *If* $\mathbf{p} \in \mathbb{R}^n$ *and* $\mathbf{q} \in \mathbb{R}^n$ *are two probability distributions,*

$$\mathbf{p} \succ \mathbf{q} \quad \Rightarrow \quad H(\mathbf{q}) - H(\mathbf{p}) \geq D\left(\mathbf{p}^{\downarrow} || \mathbf{q}^{\downarrow}\right). \tag{2.25}$$

This shows that not only is majorization "stronger" than a Shannon entropy inequality (2.24) in the sense that it implies it, it even implies something stronger.

Theorem 3 can actually be generalised by relaxing the form of the Schur-convex function G. The following theorem was essentially proved by Schur [34] for a restricted domain of the Schur-convex function.

Theorem 5. *If* $G : \mathbb{R}^n \to \mathbb{R}$ *is symmetric and convex, then* G *is Schur-convex.*

The Rényi entropies of order $a \in [0, 1)$ cannot be written as a sum over concave functions, like the Shannon entropy. They are however all concave, as mentioned before. Since they are all symmetric, the following Corollary is true.

Corollary 2. *If* $\mathbf{p} \in \mathbb{R}^n$ *and* $\mathbf{q} \in \mathbb{R}^n$ *are two probability distributions,*

$$\mathbf{p} \succ \mathbf{q} \quad \Rightarrow \quad H_a(\mathbf{p}) \leq H_a(\mathbf{q}), \quad \forall a \in [0, 1]. \tag{2.26}$$

The other Rényi entropies are neither convex, nor concave. Corollary 2 can nevertheless be generalised as well.

Theorem 6. *If* $\mathbf{p} \in \mathbb{R}^n$ *and* $\mathbf{q} \in \mathbb{R}^n$ *are two probability distributions,*

$$\mathbf{p} \succ \mathbf{q} \quad \Rightarrow \quad H_a(\mathbf{p}) \leq H_a(\mathbf{q}), \quad \forall a \in [0, \infty). \tag{2.27}$$

This actually means that Theorem 5 is only an implication, as a Schur-convex function is not necessarily convex, although it is always symmetric. It would be interesting to obtain an inequality similar to the one of Theorem 4, using the Rényi divergence of Definition 4. One can however find an example which demonstrates that this is not possible [36]. Still, the generalisation can be done by considering the so-called Tsallis entropies (which we choose not to define here) and some corresponding measures of divergence [36].

Theorem 6 relates all our uncertainty functions, the Rényi entropies, to the concept of disorder inherent to majorization. One could ask the question whether all the inequalities on the Rényi entropies imply a majorization relation. In other words, can Relation (2.27) be generalised to an equivalence instead of an implication? This happens to be impossible, as one can

always find a couple of probability distributions $\mathbf{p} \in \mathbb{R}^n$ and $\mathbf{q} \in \mathbb{R}^n$ which are incomparable, but whose Rényi entropies follow the same inequality for all $a \in [0, \infty)$. It actually means that the knowledge of all the Rényi entropies of $\mathbf{p}$ and $\mathbf{q}$ is not sufficient to conclude whether the two vectors are comparable. It is however possible to check that $G(\mathbf{p}) \geq G(\mathbf{q})$ for some specific Schur-convex functions G in order to investigate whether $\mathbf{p} \succ \mathbf{q}$. These specific functions of the vectors of probabilities typically output the cumulated sums of Equation (2.18). This fact is contained in the following theorem, which was proved by Hardy, Littlewood and Pólya [33, 37].

Theorem 7. *Let $\mathbf{p} \in \mathbb{R}^n$ and $\mathbf{q} \in \mathbb{R}^n$ be two probability distributions. The inequality*

$$\sum_{i=1}^{n} g(p_i) \geq \sum_{i=1}^{n} g(q_i) \tag{2.28}$$

holds for all continuous convex functions $g : \mathbb{R} \to \mathbb{R}$ if and only if $\mathbf{p} \succ \mathbf{q}$.

One needs the information on all the cumulated sums (Equation (2.18)) of two vectors $\mathbf{p}$ and $\mathbf{q}$ in order to compare them in terms of majorization, while the information on the Rényi entropies is not enough in general for such a task. It can however be exploited on the context of so-called catalytic majorization, which is defined as follows [38].

Definition 8 (Catalytic majorization)**.** *Let $\mathbf{p} \in \mathbb{R}^n$ and $\mathbf{q} \in \mathbb{R}^n$ be two vectors of dimension n. We say that $\mathbf{p}$ trump-majorizes $\mathbf{q}$, i.e. $\mathbf{p} \succ_{\mathrm{T}} \mathbf{q}$, if there exists a vector $\mathbf{c} \in \mathbb{R}^m$ for some $m \geq 0$, such that $\mathbf{p} \otimes \mathbf{c} \succ \mathbf{q} \otimes \mathbf{c}$.*

One understands the reason it is called catalytic majorization, as the catalyst $\mathbf{c}$ in Definition 8 remains unchanged when comparing $\mathbf{p} \otimes \mathbf{c}$ and $\mathbf{q} \otimes \mathbf{c}$. The role of $\mathbf{c}$ will be made clearer when catalytic majorization will be exploited in the context of quantum information theory (see Section 2.3.3). The importance of Rényi entropies in this context can now be witnessed through the following theorem [39], which generalises Theorem 6.

Theorem 8. *If $\mathbf{p} \in \mathbb{R}^n$ and $\mathbf{q} \in \mathbb{R}^n$ are two probability distributions,*

$$\mathbf{p} \succ_{\mathrm{T}} \mathbf{q} \quad \Leftrightarrow \quad H_a(\mathbf{p}) \leq H_a(\mathbf{q}), \quad \forall a \geq 1. \tag{2.29}$$

Theorem 8 was originally expressed in terms of l_p norms, which are essentially equivalent to the Rényi entropies.

Until now, all the definitions and theorems we presented applied to vectors with a finite dimension n. Nevertheless, the theory of majorization nicely adapts to the infinite-dimensional case. Thus, Definition 5 of majorization in terms of cumulated sums stays unchanged, with

n "simply" going to infinity. Bistochastic matrices, however, should be swapped for column-stochastic and row-substochastic matrices in this context. For completeness, we state the corresponding definition.

Definition 9 (Column-stochastic and row-substochastic matrix)**.** *A infinite matrix* $\mathbf{D}$ *is said to be column-stochastic and row-substochastic if all its elements are non-negative, and if its columns all sum to one and its rows all sum to a value less than one, i.e.,*

$$D_{ij} \geq 0 \; \forall i, j = 1, \ldots, \quad \sum_{i=1}^{\infty} D_{ij} = 1, \; \forall j = 1, \ldots, \quad \sum_{j=1}^{\infty} D_{ij} \leq 1, \; \forall i = 1, \ldots. \tag{2.30}$$

The most important theorem concerning majorization is arguably Theorem 2. It was generalised to the infinite-dimensional case in [40] and [41], as follows.

Theorem 9. *Let* $\mathbf{p}$ *and* $\mathbf{q}$ *be two infinite probability distributions.* $\mathbf{p}$ *majorizes* $\mathbf{q}$ *if and only if there exists a column-stochastic and row-substochastic matrix* $\mathbf{D}$ *such that*

$$\mathbf{q} = \mathbf{D}\mathbf{p}. \tag{2.31}$$

The interested reader can find a nice review of these notions in [42], which also generalises some other well known majorization results to infinite dimensions.

2.2 Continuous probability densities

2.2.1 Differential Shannon entropy

Although the meaning of continuous variables should be understood from its denomination, we begin by clarifying some concepts to which it relates. Take X to be a random variable with cumulative distribution function

$$F(x) = \Pr\left(X \leq x\right). \tag{2.32}$$

We will say that the random variable X is continuous whenever F is continuous. If F is differentiable almost everywhere, we define the probability density function for X as $f(x) = F'(x)$. It verifies

$$\int_{-\infty}^{\infty} \mathrm{d}x \, f(x) = 1. \tag{2.33}$$

The differential entropy came as an attempt by Shannon to extend the idea of his Shannon entropy to continuous variables. He defined it as follows [24].

Definition 10 (Differential entropy)**.** *The differential entropy* $h(f)$ *of a continuous random vari-*

able X with density $f(x)$ is defined as

$$h(f) = -\int_{-\infty}^{\infty} \mathrm{d}x\, f(x) \log f(x). \tag{2.34}$$

As we did in the discrete case, we took the differential entropy to be a function of the density $f(x)$ rather than the variable X itself, as it only depends on $f(x)$. Furthermore, the log is taken to base 2 (this can be generalised to other bases).

Remark 4. The definition of the differential entropy obviously only applies to situations in which both the density $f(x)$ and the integral given by $h(f)$ exist.

The differential entropy is an extension of the Shannon entropy to the continuous case. The definitions of the two quantities are similar, with discrete summations being replaced by integrations when going from the Shannon entropy to the differential one. Although we expect them to share some properties, the differential entropy is not always positive, even for a genuine random variable, unlike in the discrete case. It is however still concave.

Property 12 (Concavity)**.** *The differential entropy of any random variable X is concave in its density $f(x)$.*

Another interesting property concerns the entropy of a vector of random variables $\mathbf{X}$ which is transformed by a matrix A.

Property 13. *Suppose a vector of random variables $\mathbf{X}$ with density f is transformed into $A\mathbf{X}$ with density f_A, where A is a matrix. The entropy of the new variable is given by*

$$h(f_A) = h(f) + \log|\det[A]|, \tag{2.35}$$

where the log *is taken to base 2.*

In the last equation, $\det[A]$ stands for the determinant of a matrix A. This property will be interesting in the context of Bosonic Gaussian transformations, which are characterised in phase space by a matrix whose determinant equals one.

It is in our interest to introduce the definition of continuous relative entropy, as it will be useful in the context of entropy power inequalities, which will be introduced in Chapter 4. It is as follows.

Definition 11 (Continuous relative entropy)**.** *The continuous relative entropy between two random variables X and Y with respective densities f_X and f_Y is defined as*

$$D(f_X||f_Y) = \int_{-\infty}^{\infty} \mathrm{d}x\, f_X(x) \log \frac{f_X(x)}{f_Y(x)}. \tag{2.36}$$

As in the discrete case, it is always positive.

Property 14 (Positivity)**.** *The continuous relative entropy between any two continuous random variables X and Y with respective densities f_X and f_Y is non-negative, i.e.,*

$$D(f_X||f_Y) \geq 0. \tag{2.37}$$

It is equal to zero if and only if $f_X(x) = f_Y(x)$, for almost all $x \in \mathbb{R}$ (i.e. all x outside a set of Lesbegue measure zero).

2.2.2 Continuous Rényi entropies

In this section, we simply define the Rényi entropies in the continuous case, since they will serve as interesting examples when describing continuous majorization.

Definition 12 (Continuous Rényi entropy)**.** *The continuous Rényi entropy $h_\alpha(f)$ of order $\alpha \in (0,1) \cup (1,\infty)$ of a continuous random variable X with density $f(x)$ is defined as*

$$h_\alpha(f) = \frac{1}{1-\alpha} \log \left(\int_{-\infty}^{\infty} \mathrm{d}x\, f^\alpha(x) \right). \tag{2.38}$$

Like in the discrete case, we extend the definition of h_α to $\alpha = 1$, by setting

$$h_1(f) := h(f). \tag{2.39}$$

2.2.3 Continuous majorization relations and Schur-convex functions

As we did for the Shannon entropy, we are now going to extend the concept of majorization to continuous variables. In the discrete case, it was applied to probability vectors which had first been reordered. In order to apply the theory to continuous probability densities, these should similarly be reordered in some way. In order to be able to do so, we begin by presenting the concept of rearrangement of a non-negative function, which is based on the rearrangement of a set [43].

Definition 13 (Rearrangement of a Borel set)**.** *For a Borel set A with volume $|A|$, one can define its spherically decreasing symmetric rearrangement $A^\downarrow$ by*

$$A^\downarrow = B(0,r), \tag{2.40}$$

where $B(0,r)$ stands for the open ball of radius r centred at the origin with volume $|A|$.

We choose the convention that if $|A| = 0$, then $A^\downarrow = \emptyset$. Figure 2.2.1 depicts an example of rearrangement of a set in $\mathbb{R}^2$.

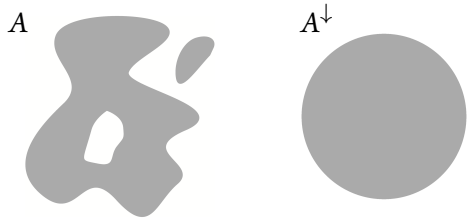

Figure 2.2.1: Example of rearrangement of a Borel set in $\mathbb{R}^2$. The figure on the left represents a Borel set A. The figure on the right represents its rearrangement $A^{\downarrow}$, which is simply a disk having the same area as A.

The rearrangement of a non-negative function can be understood from its layer cake representation [44]. Before presenting the latter, let us introduce a functional which will be of importance later on. For any non negative real-valued measurable function f defined on a n-dimensional Euclidean space $\mathbb{R}^n$, we define [5]

$$m_f(t) = |\{x : f(x) > t\}|, \quad t \geq 0. \tag{2.41}$$

This function represents the volume of the set of elements x such that $f(x) > t$. The layer cake representation of f is the formula

$$f(x) = \int_0^\infty 1_{\{y \in \mathbb{R}^n | f(y) \geq t\}}(x) \mathrm{d}t, \tag{2.42}$$

where 1_A denotes the indicator function of a subset $A \subseteq \mathbb{R}^n$. This can be obtained from the fact that

$$1_{\{y \in \mathbb{R}^n | f(y) \geq t\}}(x) = 1_{[0, f(x)]}(t), \tag{2.43}$$

and the identity

$$f(x) = \int_0^{f(x)} \mathrm{d}t = \int_0^\infty 1_{[0, f(x)]}(t) \mathrm{d}t. \tag{2.44}$$

Alternatively, we can write

$$f(y) = \int_0^\infty \{y \in B_t\} \, \mathrm{d}t, \tag{2.45}$$

where

$$\{y \in B_t\} = \begin{cases} 1 & \text{if} \quad y \in B_t, \\ 0 & \text{else}, \end{cases} \tag{2.46}$$

and $B_t = \{x : f(x) > t\}$. Now that we know how the layer cake representation works, we are able to define the rearrangement of a non-negative function [43].

Definition 14 (Spherically decreasing symmetric rearrangement of a non-negative function)**.** *For a measurable non-negative function f, one can define its spherically decreasing symmetric rear-*

rangement $f^{\downarrow}$ *by*

$$f^{\downarrow}(y) = \int_0^{\infty} \left\{ y \in B_t^{\downarrow} \right\} \mathrm{d}t, \tag{2.47}$$

where

$$\left\{ y \in B_t^{\downarrow} \right\} = \begin{cases} 1 & \text{if} \quad y \in B_t^{\downarrow} \\ 0 & \text{else} \end{cases} \tag{2.48}$$

and $B_t = \{x : f(x) > t\}$.

By construction, this kind of rearrangement moves the mass of the function towards the origin. As an illustration, Figure 2.2.2 shows an example of such an operation, applied to a non-negative function in $\mathbb{R}$. As a second illustration, Figure 2.2.3 shows two examples of the

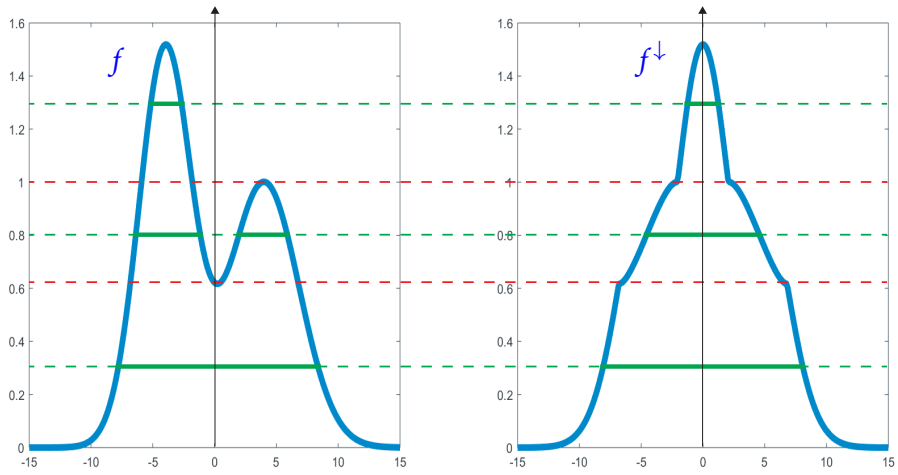

Figure 2.2.2: Example of spherically decreasing symmetric rearrangement in one dimension. The figure on the left represents a non-negative function defined f on $\mathbb{R}$. The figure on the right represents its spherically decreasing symmetric rearrangement $f^{\downarrow}$. As one can see, some discontinuities in the derivatives can in general appear when constructing the rearrangement. These discontinuities are highlighted by the red dotted lines. The green plain lines represent the fact that the volumes of the level sets of the function remain unchanged.

spherically decreasing symmetric rearrangement, applied to two-dimensional probability distributions. The latter actually represent the Wigner function of two normalised extremal Fock-passive states, which will be defined in Section 7.1, in the context of bosonic quantum systems. The definition of the rearrangement can actually be extended to negative functions, but they would have to have a finite support in order to do so. Note that there is an alternative definition of the rearrangement of a non-negative function, which produces an asymmetric non-negative function of a single real argument. We refer the interested reader to Appendix A.1 for such a

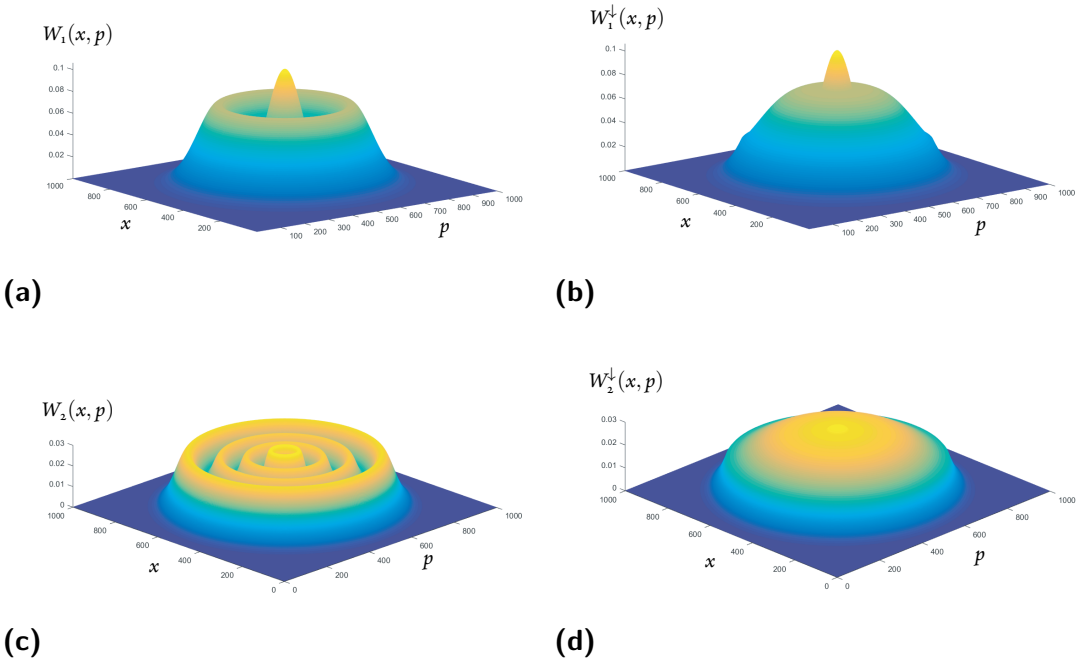

Figure 2.2.3: **(a)** Wigner function of the third normalised extremal Fock-passive state. **(b)** Spherically decreasing symmetric rearrangement of the Wigner function of the third normalised extremal Fock-passive state. **(c)** Wigner function of the eighth normalised extremal Fock-passive state. **(d)** Spherically decreasing symmetric rearrangement of the Wigner function of the eighth normalised extremal Fock-passive state. As one can see, in the case of **(b)** and **(d)**, the mass of the functions has been moved towards the origin.

definition.

Using one's intuition, it is possible to understand that

$$\{x : f(x) > t\}^{\downarrow} = \{x : f^{\downarrow}(x) > t\}. \tag{2.49}$$

This simply means that a non-negative function and its spherically decreasing symmetric rearrangement have the same rearranged level sets. As a consequence, one can see that

$$m_f(t) = m_{f^{\downarrow}}(t), \quad \forall\, t \geq 0. \tag{2.50}$$

Another direct implication is that two non-negative functions f and g have the same rearrangements, *i.e.*, $f^{\downarrow} = g^{\downarrow}$, if and only if they satisfy

$$m_f(t) = m_g(t), \quad \forall\, t \geq 0. \tag{2.51}$$

In the discrete case, the Rényi entropies H_a played a special role in the context of the theory of majorization. One of their simplest properties was that they were symmetrical. This can be retrieved in the context of continuous distributions, as Rényi entropies are preserved by rear-

rangements.

Lemma 1. *For any* $a \in (0, \infty)$,

$$h_a(f) = h_a(f^{\downarrow}). \tag{2.52}$$

Indeed, in the continuous case, the spherically decreasing symmetric rearrangement plays the role of the operation which transforms vectors into their reordered versions through permutations. Lemma 1 is actually a particular case of the following theorem involving convex functions.

Theorem 10. *Let* f *be a probability density and* $\varphi(x)$ *be a convex function defined on the non-negative real line such that* $\varphi(0) = 0$ *and it is continuous at* 0. *Then*

$$\int \varphi(f(x))\mathrm{d}x = \int \varphi(f^{\downarrow}(x))\mathrm{d}x, \tag{2.53}$$

provided that the two integrals are well-defined.

An elegant proof of Theorem 10 can be found in [43].

Using the concept of a decreasing rearrangement, we are now able to introduce majorization theory for continuous variables. Similarly to discrete majorization, continuous majorization is a partial order on probability densities [43].

Definition 15 (Continuous majorization)**.** *For probability densities* f *and* g *on* $\mathbb{R}^n$, *we say that* f *majorizes* g, *i.e.* $f \succ g$ *if*

$$\int_{\{x:||x||<r\}} f^{\downarrow}(x)\mathrm{d}x \geq \int_{\{x:||x||<r\}} g^{\downarrow}(x)\mathrm{d}x, \quad \forall\, r \geq 0. \tag{2.54}$$

Like in the discrete case, majorization consists in comparing all the "cumulated" integrals of the rearranged probability densities. Note that, by construction, two probability densities f and g which verify

$$m_f(t) = m_g(t), \quad \forall\, t \geq 0, \tag{2.55}$$

will satisfy $f \succ g$ and $f \prec g$, meaning that they will be equivalent in terms of majorization, since they will have the same rearrangements. This is comparable to the situation in which two probabilities have the same spectrum in the discrete case. We will be denoting this by $f \equiv g$ in the following.

In the context of continuous majorization, the notion of Schur-convex function can be extended as follows [5].

Theorem 11. *Let $\varphi(x)$ be a convex function defined on the non-negative real line such that $\varphi(0) = 0$ and it is continuous at 0. Consider two probability densities f and g. If $f \succ g$ then*

$$\int \varphi(f(x))\mathrm{d}x \geq \int \varphi(g(x))\mathrm{d}x, \tag{2.56}$$

provided that both sides are well-defined.

One could ask the question whether majorization can still be clearly related to the concept of disorder in the continuous case. Using Jensen inequality, we show in appendix A.2 that when a density g can be related to functions which are equivalent to a density f, through a continuous mixture, f and g obey a majorization relation. This is encompassed in the following Lemma.

Lemma 2. *If f and g are probability densities, and there exists a distribution $K : \mathbb{R} \to \mathbb{R}_{\geq 0}$ such that $\int_{\mathbb{R}} K = 1$ and*

$$g(x) = \int_{-\infty}^{\infty} \mathrm{d}y\, K(y) f_y(x), \quad \forall x \in \mathbb{R}, \tag{2.57}$$

where f_y has the same rearrangement as f, for all $y \in \mathbb{R}$, then $f \succ g$.

Like in the discrete case, this is an intuitive way of considering that g is more disordered than f. It turns out it is possible to define a so-called doubly stochastic function for continuous variables [5].

Definition 16 (Bistochastic function)**.** *A function $k : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_{\geq 0}$ is said to be bistochastic if*

$$\int_{-\infty}^{\infty} \mathrm{d}y\, k(x, y) = 1, \ \forall x \in \mathbb{R}, \quad \int_{-\infty}^{\infty} \mathrm{d}x\, k(x, y) = 1, \ \forall y \in \mathbb{R}. \tag{2.58}$$

We restricted the definition to a function $k : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_{\geq 0}$, but it can easily be generalised to functions $k : \mathcal{A} \times \mathcal{A} \to \mathbb{R}_{\geq 0}$, where $\mathcal{A}$ is any set on which the probability densities we consider are defined. One can readily use Jensen's inequality to prove the following theorem.

Theorem 12. *If f and g are probability densities, and there exists a bistochastic function $k : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_{\geq 0}$ such that*

$$g(x) = \int_{-\infty}^{\infty} \mathrm{d}y\, k(x, y) f(y), \quad \forall x \in \mathbb{R}, \tag{2.59}$$

then $f \succ g$.

Note that unlike in the discrete case, Theorem 12 is not an equivalence.

2.3 Extension to quantum information theory

2.3.1 Quantum entropies

In 1932, von Neumann introduced a mathematical formalism of quantum mechanics [2]. He had already associated an entropy quantity to a statistical operator in 1927 [45], but later discussed it in more details in his book on the quantum theory. Before Shannon's work, the concept of entropy had no connection with the notion of information. von Neumann actually defined his entropy in the framework of a thought experiment on the ground of phenomenological thermodynamics [26].

A quantum state will in general be described by a density operator (or density matrix) ρ, *i.e.* a positive semi-definite Hermitian operator of trace one in a Hilbert space $\mathcal{H}$ of the system [46]. In this context, the von Neumann entropy is defined as follows [2].

Definition 17 (von Neumann entropy)**.** *The von Neumann entropy of a density matrix ρ is defined as*

$$S(\rho) = -\mathrm{Tr}\left[\rho \log \rho\right]. \tag{2.60}$$

Remark 5. In Equation (2.60), the function $x \log x$ of an operator is well-defined by functional calculus.

If the state ρ is written using its spectral decomposition

$$\rho = \sum_i \lambda_i \left|\psi_i\right\rangle \left\langle\psi_i\right|, \tag{2.61}$$

where the $\left\{\left|\psi_i\right\rangle\right\}_{i\in\mathbb{N}}$ represent a basis of pure states, one can readily check that its von Neumann entropy can be computed to be the Shannon entropy of the classical probability distribution $\boldsymbol{\lambda}$, the vector containing all the λ_i for $i \in \mathbb{N}$, *i.e.*,

$$S(\rho) = H(\boldsymbol{\lambda}). \tag{2.62}$$

The reason the state ρ we are attempting to describe is usually mixed, and written in the form of Equation (2.61), is that it can actually be in any of the pure states $\left|\psi_i\right\rangle \left\langle\psi_i\right|$. We may however not know exactly which of the $\left|\psi_i\right\rangle \left\langle\psi_i\right|$ is the actual state of our system, which is why we rely on a distribution $\{\lambda_i\}$ to describe it. The von Neumann entropy captures exactly this idea by evaluating how uncertain we are of the exact pure state of our system. As such, this uncertainty is actually completely classical. If there was a way to know in which pure state our system is, the von Neumann entropy of ρ would always be zero.

The von Neumann entropy shares many properties with its classical counterpart. We list some of these properties below.

Property 15 (Positivity). *The von Neumann entropy of any density matrix ρ is non-negative, i.e.,*

$$S(\rho) \geq 0. \tag{2.63}$$

It is zero if and only if the state ρ is pure.

This is in accordance with the interpretation of von Neumann's functional we gave above.

Property 16 (Concavity). *The von Neumann entropy is a concave function of its inputs, i.e., for any probability distribution $\{p_i\}$, and corresponding density matrices ρ_i,*

$$S\left(\sum_i p_i \rho_i\right) \geq \sum_i p_i S(\rho_i). \tag{2.64}$$

Similarly to what we explained earlier, the expression $\sum_i p_i \rho_i$ is a signature of the fact that we do not know the exact state of our system, which might be described by any of the ρ_i. As such, not only does it express our ignorance due to the ρ_i themselves, there is actually a contribution due to our ignorance of the index i. Consequently, our uncertainty of the mixture $\sum_i p_i \rho_i$ should logically be greater than the average uncertainty of the states ρ_i, which is what the property of concavity of the von Neumann entropy captures.

Suppose a quantum system is comprised of two subsystem A and B. One can ask the question whether one can compare the von Neumann entropies of the latter with the whole system AB. This can be done using the notion of quantum conditional entropy, which we define now [47].

Definition 18 (Quantum conditional entropy). *Consider a quantum system in a state ρ_{AB}, composed of two subsystems in respective states ρ_A and ρ_B. The quantum conditional entropy is defined as*

$$S\left(\rho_A|\rho_B\right)_{\rho_{AB}} = S\left(\rho_{AB}\right) - S\left(\rho_A\right). \tag{2.65}$$

The reason we introduce this notion is because it will later serve as an interesting way to compare the von Neumann entropy with the concept of majorization in the quantum realm.

The goal of the present section is to generalise entropy measures, among which the Shannon entropy, to the quantum setting. It seems natural to try and do the same for the relative entropy, given how powerful a tool it is in classical information theory. This can be done as follows [46].

Definition 19 (Quantum relative entropy). *The quantum relative entropy between two density operators ρ and σ is defined as*

$$S\left(\rho||\sigma\right) = \mathrm{Tr}\left[\rho \log \rho\right] - \mathrm{Tr}\left[\rho \log \sigma\right]. \tag{2.66}$$

Remark 6. According to our conventions, whenever $\mathrm{supp}(\rho) \cap \ker(\sigma) \neq 0$, $S\left(\rho||\sigma\right) = \infty$.

Again, many properties of the relative entropy can be extended to the quantum realm.

Property 17 (Klein's inequality)**.** *The quantum relative entropy between any two states ρ and σ is non-negative, i.e.,*

$$S(\rho||\sigma) \geq 0. \tag{2.67}$$

It is equal to zero if and only if $\rho = \sigma$.

Property 18 (Joint convexity)**.** *The quantum relative entropy is jointly convex in its arguments; that is, if $\left(\rho^{(1)}, \sigma^{(1)}\right)$ and $\left(\rho^{(2)}, \sigma^{(2)}\right)$ are two pairs of density matrices, then*

$$S\left(\lambda\rho^{(1)} + (1-\lambda)\rho^{(2)}||\lambda\sigma^{(1)} + (1-\lambda)\sigma^{(2)}\right) \leq \lambda S\left(\rho^{(1)}||\sigma^{(1)}\right) + (1-\lambda)S\left(\rho^{(2)}||\sigma^{(2)}\right), \tag{2.68}$$

for all λ such that $0 \leq \lambda \leq 1$.

Finally, it is in our interest in the present dissertation to extend the Rényi entropies to the quantum realm as well. As expected, they will play an interesting role in the generalisation of majorization to the quantum theory.

Definition 20 (Quantum Rényi entropy)**.** *For any density matrix ρ, the quantum Rényi entropy of order $a \in (0,1) \cup (1,\infty)$ is defined as*

$$S_a(\rho) = \frac{1}{1-a} \log\left(\mathrm{Tr}\left[\rho^a\right]\right). \tag{2.69}$$

Like in the classical realm, the von Neumann entropy can be recovered from Expression (2.69) by taking the limit for $a \to 1$. As a consequence, we extend Definition 20 to $a = 1$ by setting

$$S_1(\rho) := S(\rho). \tag{2.70}$$

Since the quantum Rényi entropy of any state ρ can be seen as the Rényi entropy of the probability distribution given by its eigenvalues, the properties we listed in section 2.1.2 trivially extend to the quantum case. We hereby present them for completeness.

Property 19. *The quantum Rényi entropy of a fixed density matrix ρ is non-increasing in its order a.*

Property 20 (Positivity)**.** *The quantum Rényi entropies of any density matrix ρ are positive for all $a \geq 0$, i.e.,*

$$S_a(\rho) \geq 0, \quad \forall a \geq 0. \tag{2.71}$$

Property 21 (Concavity)**.** *The quantum Rényi entropies of any density matrix ρ are concave in ρ for $a \in (0,1]$.*

It is interesting to mention that the property of symmetry of the Rényi entropies (among which the Shannon entropy) can be generalised to the quantum realm by actually considering any unitary transformation.

Property 22 (Unitary invariance)**.** *The quantum Rényi entropies of the density matrix ρ are invariant under any unitary transformation applied to ρ.*

2.3.2 Majorization relations for density matrices

We will denote by $\mathbf{\Sigma}(\rho)$ the vector whose elements constitute the spectrum of the density matrix ρ. Since $\mathbf{\Sigma}(\rho)$ is a well-defined probability distribution, the definition of majorization can naturally be extended to quantum states through their spectra.

Definition 21 (Majorization for quantum states)**.** *Let ρ and σ be two density matrices. We say that ρ majorizes σ, i.e. $\rho \succ \sigma$, whenever $\mathbf{\Sigma}(\rho) \succ \mathbf{\Sigma}(\sigma)$.*

Like in the classical case, we will say that two states ρ and σ are equivalent in terms of majorization when both $\rho \succ \sigma$ and $\rho \prec \sigma$ are satisfied. We will denote this as $\rho \equiv \sigma$.

As we showed, it was rather trivial to generalise the definitions of the Shannon and Rényi entropies to quantum information theory. Definition 21 does so for majorization, but we would like to have a direct generalisation of Definition 5. In order to do this, consider a state ρ written in its spectral decomposition as

$$\rho = \sum_{i=1}^{n} \lambda_i^{\downarrow} \left|\psi_i\right\rangle \left\langle\psi_i\right|, \tag{2.72}$$

meaning that the $\left\{\left|\psi_i\right\rangle\right\}_{i\in\mathbb{N}}$ have been chosen so that $\lambda_i^{\downarrow} \geq \lambda_{i+1}^{\downarrow}$ for all $i = 1, \ldots, n-1$. Define $Q(\rho)_k$ as

$$Q(\rho)_k = \sum_{i=1}^{k} \left|\psi_i\right\rangle \left\langle\psi_i\right|, \tag{2.73}$$

which is a projector on the k first eigenvectors of ρ corresponding to its k highest eigenvalues. We then have the following trivial lemma.

Lemma 3. *If ρ and σ are two density matrices, then $\rho \succ \sigma$ if and only if*

$$\mathrm{Tr}\left[Q(\rho)_k\, \rho\right] \geq \mathrm{Tr}\left[Q(\sigma)_k\, \sigma\right], \quad k = 1, \ldots, n. \tag{2.74}$$

Our goal in introducing such an obvious lemma is to stress that, unlike in the definition of the quantum Rényi entropies, one needs to diagonalise both the states ρ and σ in order to compare them in terms of majorization. This actually reminds us of the fact that one needs to reorder

the elements of a probability distribution in order to use it in the context of Definition 5 (it is actually the reason why the diagonalisation needs to be done in the quantum case).

When looking at Definition 21, one may wonder about the purpose of such a generalisation to the quantum realm, since the theory seems to be dealing solely with the classical probability distribution of the eigenvalues of a quantum state. As we are going to show later, majorization theory and the corresponding notion of disorder constitute a powerful tool for the study of some primary, purely quantum resource. For this reason, a characterisation of majorization for quantum states seems important enough in itself. One of the most relevant properties of the theory is arguably the one relating it to the notion of disorder, through the concept of mixture of permutations of Equation (2.21). This raises the question whether such a mixture can be generalised when it involves quantum objects. The following theorem, due to Uhlmann, provides us with a non-trivial answer [46, 48–50].

Theorem 13 (Uhlmann's theorem)**.** *Consider two density matrices ρ and σ. The relation $\rho \succ \sigma$ holds if and only if there exists a probability distribution $\{p_i\}$ and unitary matrices U_i such that*

$$\sigma = \sum_i p_i U_i \rho U_i^\dagger . \tag{2.75}$$

Equation (2.75) is a generalisation of the fact that when two probability vectors are related by a majorization relation, one of them can be obtained from the other by applying a mixture of permutations. As a consequence, it allows us to make sense of the concept of disorder in quantum information theory. Furthermore, Theorem 13 is an elegant explanation of the fact that majorization relations arise frequently in quantum mechanics, which can be seen as a result of the fundamental role played by unitarity in the quantum theory.

The mixture of unitaries of Equation (2.75) is a particular case of unital channels. A channel is a linear, completely positive and trace-preserving map (see Section 3.4.1 for the detailed definition of a channel), while a unital map can be defined as follows for finite-dimensional systems [51].

Definition 22 (Unital map)**.** *Denote as $\mathfrak{T}(\mathcal{H})$ the space of all operators in a finite-dimensional Hilbert space $\mathcal{H}$ equipped with the trace norm. A map $\mathcal{M} : \mathfrak{T}(\mathcal{H}) \to \mathfrak{T}(\mathcal{H})$ is called unital if*

$$\mathcal{M}[\mathbb{I}] = \mathbb{I}, \tag{2.76}$$

where $\mathbb{I}$ represents the identity operator in $\mathfrak{T}(\mathcal{H})$.

A unital channel can alternatively be called a bistochastic completely positive map (which is by definition trace-preserving). Whenever two states are related by such a channel, they can be ordered using a majorization relation. Note that a unital channel cannot always be written as a mixture of unitaries. The following theorem can however be proved for finite dimensions [51].

Theorem 14. *Denote by M_n the set of $n \times n$ density matrices. A channel $\mathcal{C} : M_n \to M_n$ is unital if and only if it can be written as an affine combination of unitaries, i.e.,*

$$\mathcal{C}[\rho] = \sum_i a_i U_i \rho U_i^\dagger, \quad \forall \rho \in M_n, \tag{2.77}$$

where the a_i are all real and sum to one, and each $U_i \in M_n$ is unitary.

As we already explained, von Neumann did not connect his concept of entropy to information theory. The relation can best be understood through Schumacher's noiseless channel coding theorem [52], which quantifies the resources needed to perform quantum data compression [46]. Consider a source producing orthogonal states $|\psi_j\rangle$ with probabilities p_j. Schumacher's theorem tells us that it may be compressed down to the Shannon entropy $H(\mathbf{p})$. However, if the states $|\psi_j\rangle$ are not orthogonal, Schumacher's theorem certifies that the source may actually be compressed down to the von Neumann entropy $S\left(\sum_j p_j \left|\psi_j\right\rangle \left\langle\psi_j\right|\right)$, which may be less than $H(\mathbf{p})$. The entropic inequality relating $S\left(\sum_j p_j \left|\psi_j\right\rangle \left\langle\psi_j\right|\right)$ and $H(\mathbf{p})$ is actually a consequence of a stronger relation involving majorization theory. It is encompassed in the Schur-Horn theorem [34], which is as follows.

Theorem 15 (Schur-Horn theorem)**.** *If H is an $n \times n$ Hermitian matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ and diagonal elements $h_1, \ldots, h_n$, then*

$$\boldsymbol{\lambda} \succ \mathbf{h}. \tag{2.78}$$

2.3.3 Measures of entanglement and LOCC

One of the most important resources put forward by quantum mechanics resides in the notion of entanglement of two particles [53]. This "spooky" action at a distance was first noticed by Einstein, Podolsky and Rosen in their famous paper of 1935 [54], and originally called *Verschränkung* (now translated as entanglement) by Schrödinger [55]. Einstein, Podolsky and Rosen argued that this particular effect was a result of the fact that the quantum description of physical reality was not complete, and needed to be extended using some *hidden variables*. Bell later proved that the predictions of quantum mechanics cannot be explained by any local hidden variable theory [56]. The major consequence is that entanglement cannot be simulated within any classical formalism. A mixed state ρ of n systems will be called separable if it can be written as a convex combination of product states, *i.e.,*

$$\rho = \sum_i p_i \rho_i^{(1)} \otimes \ldots \otimes \rho_i^{(n)}, \tag{2.79}$$

where $\rho_i^{(j)}$ is a state of subsystem j. A state which is not separable is called entangled [57]. Entanglement is nowadays considered as the most substantial resource of quantum mechanics. Its

first application as such was discovered by Ekert, who explained in 1991 that it could be used in the context of quantum key distribution [58]. Entanglement has since been shown to be a resource for a plethora of applications, such as quantum dense coding [59], quantum teleportation [60] and others. One then understands the importance of being able to quantify, or measure entanglement as a resource. Several measures of entanglement have been proposed in the last 20 years. As it happens, quantum correlations in a quantum system share a close connection with the disorder or uncertainty of the corresponding subsystems. Thus, many of the measures put forward are based on the concept of entropy. Furthermore, all of these quantities (the ones based on entropy) coincide whenever the quantum system is in a pure state.

Definition 23 (Entropy of entanglement)**.** *The entropy of entanglement of a bipartite pure state* $|\psi\rangle$ *is defined as*

$$E(|\psi\rangle) = S(\rho), \tag{2.80}$$

where ρ *is the state of one of the subsystems of* $|\psi\rangle$*.*

In Equation (2.80), it does not matter which subsystem we consider, as the two subsystems of a pure state have the same entropy (as a matter of fact, they share the same spectrum). The connection between entanglement and disorder goes deeper than what is exhibited by Definition 23. This can be witnessed by introducing the concept of local operations with classical communication (LOCC) [61], under which entanglement should never increase. Any function which does not increase on average under LOCCs is called an entanglement monotone. Since entropy seems to play a particular role in the theory of entanglement, one could ask whether the latter could be investigated in the framework of the theory of majorization. As it happens, the latter can be linked to the concept of LOCCs through the following result proved by Nielsen [6].

Theorem 16. *A pure bipartite state* $|\varphi\rangle$ *transforms to another pure bipartite state* $|\psi\rangle$ *using local operations and classical communication, i.e.* $|\varphi\rangle \xrightarrow{\text{LOCC}} |\psi\rangle$*, if and only if*

$$\rho_\psi \succ \rho_\varphi, \tag{2.81}$$

where ρ_ψ *and* ρ_φ *are respective states of subsystems of* $|\psi\rangle$ *and* $|\varphi\rangle$*.*

A consequence of Theorem 16 is that all Schur-concave functions of the state ρ_ψ of any subsystem of a pure bipartite state $|\psi\rangle$ are entanglement monotones, the entropy of entanglement of Definition 23 being such a function. As it happens, there are situations in which entanglement can help LOCCs, without being consumed. Indeed, it is possible to find states $|\varphi\rangle$ and $|\psi\rangle$ such that $|\varphi\rangle$ cannot be transformed to $|\psi\rangle$, while the transformation can be performed with the help of a catalyst state $|\chi\rangle$, as

$$|\varphi\rangle \otimes |\chi\rangle \xrightarrow{\text{LOCC}} |\psi\rangle \otimes |\chi\rangle \,. \tag{2.82}$$

This is called entanglement-assisted LOCC, which we denote as

$$|\varphi\rangle \xrightarrow{\text{ELOCC}} |\psi\rangle\,. \tag{2.83}$$

If one exploits the theory of catalytic majorization through Theorem 8, it can be shown that the transformation of Equation (2.83) is possible if and only if [39]

$$\rho_\psi \succ_T \rho_\varphi, \tag{2.84}$$

where ρ_ψ and ρ_φ are respective states of subsystems of $|\psi\rangle$ and $|\varphi\rangle$.

Theorem 16 and Equation (2.84) are some ways to compare the entanglement of two bipartite states by comparing the disorder of their respective subsystems. Majorization can however also be used as a witness of the presence of entanglement in any mixed state of a system. This fact can already be anticipated by considering the relation between separable states and the notion of quantum conditional entropy defined in Equation (2.65). Indeed, the following Lemma holds [62].

Lemma 4. *Consider a quantum system in a bipartite state ρ_{12}, composed of two subsystems in respective states ρ_1 and ρ_2, where $\rho_1 = \mathrm{Tr}_2[\rho_{12}]$ and $\rho_2 = \mathrm{Tr}_1[\rho_{12}]$. If ρ_{12} is separable, then*

$$S\left(\rho_1|\rho_2\right)_{\rho_{12}} \geq 0 \quad \textit{and} \quad S\left(\rho_2|\rho_1\right) \geq 0. \tag{2.85}$$

As it happens, the conditional entropy of a quantum system can be negative, if quantum correlations are involved. Equation (2.85) can actually be seen as an implication of the following theorem, proved by Nielsen and Kempe [63].

Theorem 17. *If a bipartite state ρ_{12} is separable, then*

$$\rho_1 \succ \rho_{12} \quad \textit{and} \quad \rho_2 \succ \rho_{12}, \tag{2.86}$$

where $\rho_1 = \mathrm{Tr}_2[\rho_{12}]$ and $\rho_2 = \mathrm{Tr}_1[\rho_{12}]$.

As expressed by Nielsen and Kempe in the title of their paper [63], separable states are more disordered globally than locally. The criterion they present is a witness of entanglement, in the sense that if a bipartite mixed state does not majorize its two subsystems, than it has to be entangled. Theorem 17 was further strengthened by Hiroshima as follows [64].

Theorem 18. *If a bipartite state ρ_{12} is not distillable, then*

$$\rho_1 \succ \rho_{12} \quad \textit{and} \quad \rho_2 \succ \rho_{12}, \tag{2.87}$$

where $\rho_1 = \mathrm{Tr}_2[\rho_{12}]$ and $\rho_2 = \mathrm{Tr}_1[\rho_{12}]$.

Entanglement distillation represents the process of extracting pure maximally entangled states from several copies of a given entangled state. A state which is not distillable is one from which no pure entangled state can be extracted, even though it may be entangled. In particular, a separable quantum state can be seen as a specific case of a state which is not distillable.

3

Bosonic quantum systems

Bosons make up one of the two fundamental classes of particles in the universe. In the quantum theory, a system of such bodies is composed of a collection of non-interacting indistinguishable particles described by Bose-Einstein statistics. In this chapter, we introduce the mathematical formalism of bosonic systems. We focus particularly on the so-called Gaussian systems, which are of great importance in quantum optics and continuous-variables quantum information theory. Gaussian states are easy to produce and manipulate experimentally. Furthermore, they can be mathematically described by only using the first two statistical moments of the quadrature operators in phase space. Similarly, Gaussian transformations are ubiquitous in quantum optics set-ups, and need only a few parameters for their theoretical description. We begin by introducing general bosonic systems in Section 3.1. In Section 3.2, we present the phase-space formalism of quantum states, as an alternative but equivalent representation to the one described in Chapter 2. Finally, section 3.3 is devoted to the detailed characterisation of Gaussian states and unitaries, while Section 3.4 concerns the generalisation of the latter to Gaussian channels.

Our overview relies on the very nicely written review of Gaussian quantum information [65]. Note that in the following, we choose natural units so that the value of Planck's constant $\hbar$ is set to 2. As a consequence, the variance of the ground state of our bosonic system is normalised to 1.

3.1 Bosonic systems in a nutshell

The state of a quantum system can always be described using a density matrix ρ in a Hilbert space; that is, a Hermitian operator with a discrete spectrum. As such, the eigenvalues of ρ can be used in order to encode information in some way. A quantum system is called a continuous variables system when its *relevant* degrees of freedom are associated to observables with continuous spectra, which might alternatively be used to encode information. A collection of N bosonic modes is an archetype of such a quantum system [65, 66]. In the framework of quantum optics, each of these canonical modes corresponds to a quantised radiation mode of the electromagnetic field, and is mathematically modelled by a quantum harmonic oscillator. The whole bosonic system is associated with a tensor-product Hilbert space $\mathcal{H} = \bigotimes_{k=1}^{N} \mathcal{H}_k$, each of the $\mathcal{H}_k$ corresponding to one of the individual modes, which are furthermore described by bosonic field operators $\hat{a}_k$ and $\hat{a}_k^\dagger$, respectively called annihilation and creation operator. These operators satisfy the bosonic commutation relations

$$\left[\hat{a}_k, \hat{a}_l^\dagger\right] = \delta_{kl}, \quad \left[\hat{a}_k, \hat{a}_l\right] = 0, \quad \left[\hat{a}_k^\dagger, \hat{a}_l^\dagger\right] = 0, \quad k, l = 1, \ldots, N. \tag{3.1}$$

They can be used to define the Hamiltonian of the complete bosonic system,

$$\hat{H} = \sum_{k=1}^{N} \hat{H}_k, \quad \hat{H}_k = \hbar\omega_k \left(\hat{a}_k^\dagger \hat{a}_k + \frac{1}{2}\right), \tag{3.2}$$

where each $\hat{H}_k$ is the Hamiltonian of the individual mode k, ω_k is the frequency of excitation in mode k, and $\hat{n}_k = \hat{a}_k^\dagger \hat{a}_k$ represents the so-called number operator in mode k.

The eigenstates of each of the number operators $\hat{n}_k$ form a countable basis for each of the corresponding infinite dimensional Hilbert spaces $\mathcal{H}_k$. Each of these eigenbasis is called a Fock basis, or number basis, and is labelled $\{|n\rangle_k\}_{n\in\mathbb{N}_0}$, $\mathbb{N}_0$ being the set of all natural numbers (including zero). Note that we choose to omit the index k of $|n\rangle_k$ in the following, since the tensor product notations will allow us to differentiate between the different Hilbert spaces. As one would expect, the eigenvalues of $\hat{n}$ correspond to the different boson numbers n, *i.e.*,

$$\hat{n}\,|n\rangle = n\,|n\rangle\,. \tag{3.3}$$

The effect of the annihilation and creation operators on the elements of the Fock basis is also well defined, and is such that $\hat{a}\,|0\rangle = 0$ and

$$\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle\,, \quad \forall n \geq 1, \tag{3.4}$$

$$\hat{a}^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle\,, \quad \forall n \geq 0. \tag{3.5}$$

The bosonic field operators can be used to derive other mathematical objects of great importance in specific bosonic system (*e.g.*, Gaussian ones, see later). The so-called quadrature field operators $\hat{q}_k$ and $\hat{p}_k$ are related to the creation and annihilation operators as

$$\hat{q}_k = \hat{a}_k + \hat{a}_k^\dagger, \quad \hat{p}_k = i(\hat{a}_k^\dagger - \hat{a}_k). \tag{3.6}$$

They equivalently completely characterise bosonic quantum systems. As a consequence, they satisfy the canonical commutation relations

$$\left[\hat{q}_k, \hat{p}_l\right] = 2i\delta_{kl}, \tag{3.7}$$

where we chose natural units so that $\hbar = 2$. For future convenience, we arrange the quadrature field operators in the vector

$$\hat{\mathbf{x}} = (\hat{q}_1, \hat{p}_1, \ldots, \hat{q}_N, \hat{p}_N)^{\mathrm{T}}, \tag{3.8}$$

and define the so-called N-mode symplectic form $\mathbf{\Omega}$ as

$$\mathbf{\Omega} = \bigoplus_{k=1}^{N} \boldsymbol{\omega}, \quad \boldsymbol{\omega} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{3.9}$$

In this case, Equation (3.7) can be equivalently rewritten

$$\left[\hat{x}_k, \hat{x}_l\right] = 2i\Omega_{kl}. \tag{3.10}$$

The bosonic field operators $\hat{a}$ and $\hat{a}^\dagger$, and the quadrature field operators $\hat{q}$ and $\hat{p}$ will all play an important role later on. However, the first two will prove more suitable in some situations (when being applied on Fock states for instance) while the other two will definitely make calculations more practical in other cases (when investigating Gaussian systems in phase space).

3.2 State space versus phase space representation

In Section 2.3, which concerned the description of a theory of disorder applied to quantum information theory, we always described the state of our quantum system using a density matrix ρ on a Hilbert space $\mathcal{H}$. Furthermore, most of the definitions and theorems we presented applied to density operators of a fixed finite dimension n, meaning that the Hilbert space itself was of finite dimension. We explained at the end of Section 2.1 that these notions could be generalised to infinite dimensional systems, although dealing with the latter does not seem very handy. In the conclusion of Section 2.2, which was about a characterisation of disorder in the case of continuous objects, *i.e.*, continuous probability distributions, we hinted at the fact that such functionals could be employed in order to overcome such a difficulty. One actually needs to generalise

these "classical" objects so that they can be used to characterise quantum systems, especially continuous variables systems, which are characterised by infinite dimensional Hilbert spaces. As a matter of fact, apart from density operators, there exists a completely equivalent representation of quantum states in terms of quasiprobability distributions, which are defined over a real symplectic space, called phase space [65], in opposition with the state space of density matrices. These generalised distributions can be introduced using the so-called Weyl operator

$$\hat{D}(\boldsymbol{\xi}) = \exp\left(i\hat{\mathbf{x}}^{\mathrm{T}}\boldsymbol{\Omega}\boldsymbol{\xi}\right), \tag{3.11}$$

where $\boldsymbol{\xi} \in \mathbb{R}^{2N}$. A quantum state described by a density operator ρ can equivalently be characterised by a Wigner characteristic function

$$\chi_\rho(\boldsymbol{\xi}) = \mathrm{Tr}\left[\rho\hat{D}(\boldsymbol{\xi})\right], \tag{3.12}$$

which is itself equivalent (in terms of representation of a quantum state) to its Fourier transform $W(\mathbf{x})$, called the Wigner function of the state. It is defined as follows.

Definition 24 (Wigner function)**.** *The Wigner function of a state ρ can be defined through its characteristic function χ_ρ as*

$$W_\rho(\mathbf{x}) = \frac{1}{(2\pi)^{2N}}\int_{\mathbb{R}^{2N}} \mathrm{d}^{2N}\boldsymbol{\xi}\, \exp\left(-i\mathbf{x}^{\mathrm{T}}\boldsymbol{\Omega}\boldsymbol{\xi}\right)\chi_\rho(\boldsymbol{\xi}). \tag{3.13}$$

The continuous variables $\mathbf{x} \in \mathbb{R}^{2N}$ which appear in Equation (3.13) can be seen as the different eigenvalues of the quadrature field operators $\hat{\mathbf{x}}$. In order for the Wigner function to characterise an actual quantum state, it has to take real values, this being a consequence of the fact that the corresponding density matrix is Hermitian. However, it does not have to be positive, meaning that it is not an actual probability distribution, but rather a quasiprobability distribution. Note that the fact that it can take negative values is not necessarily a signature of its quantum "nature". Some quantum states have a Wigner function which is positive everywhere (*e.g.*, Gaussian states, see below), even though they enjoy properties that clearly go beyond classical mechanics, such as superposition and entanglement. Since a density matrix is of trace 1, the Wigner function is normalised to 1, *i.e.*,

$$\int_{\mathbb{R}^{2N}} \mathrm{d}^{2N}\mathbf{x}\, W_\rho(\mathbf{x}) = 1. \tag{3.14}$$

Figure 3.2.1 exhibits an example of a Wigner function taking negative values, namely the Fock state with 5 photons.

Mathematically, the Wigner function (as well as the corresponding characteristic function) can be completely characterised by the statistical moments of the quantum state. As we will see

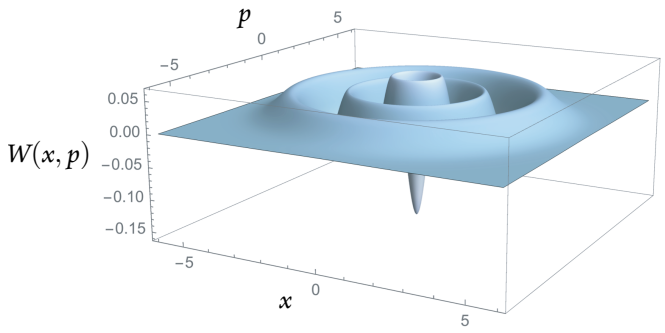

Figure 3.2.1: The Wigner function of the Fock state with 5 photons. As expected, it takes negative values, making it a quasiprobability distribution.

later, the first two moments are of paramount importance in the study of some specific quantum states [65]. The first one, called the displacement vector, is defined as

$$\bar{\mathbf{x}} = \langle \hat{\mathbf{x}} \rangle_\rho = \mathbf{Tr}\left[\hat{\mathbf{x}}\rho\right], \tag{3.15}$$

while the second moment, or covariance matrix $\mathbf{V}$, is defined through its elements

$$V_{ij} = \frac{1}{2}\left\langle \left\{ \Delta\hat{x}_i, \Delta\hat{x}_j \right\} \right\rangle_\rho, \quad \Delta\hat{x}_i = \hat{x}_i - \langle \hat{x}_i \rangle_\rho, \tag{3.16}$$

where $\{A, B\}$ represents the anticommutator between the operators A and B. Note from its definition that the covariance matrix is real and symmetric. We explained earlier that the Wigner function should be real and normalised, as a consequence of the fact that the corresponding density matrix is Hermitian and normalised. This is obviously not enough to guaranty that the Wigner function characterises a physical state, as the density matrix furthermore satisfies an uncertainty principle. This condition is difficult to verify for a Wigner function. However, it implies a rather simple condition at the level of the covariance matrix, that is [67],

$$\mathbf{V} + i\mathbf{\Omega} \geq 0. \tag{3.17}$$

Among other things, Equation (3.17) implies the positive definiteness of the covariance matrix of a quantum state.

Before ending this section, let us just give the definition of the purity of a quantum state.

Definition 25 (Purity)**.** *The purity μ_ρ of a quantum state ρ is defined as*

$$\mu_\rho = \mathrm{Tr}[\rho^2]. \tag{3.18}$$

As one would expect, it is readily seen from its definition that the purity is one for a pure state, while it is less than one for a state which is in a non-trivial mixture. The purity can be rewritten

in terms of the Wigner function as

$$\mu_\rho = (4\pi)^N \int_{\mathbb{R}^{2N}} \mathrm{d}^{2N}\mathbf{x} \left(W_\rho(\mathbf{x})\right)^2 . \tag{3.19}$$

3.3 From Gaussian unitaries to Gaussian quantum states

3.3.1 Gaussian unitaries and symplectic transformations

A unitary matrix U is a square matrix such that $U^{-1} = U^\dagger$. When applied on a density matrix ρ, it mathematically describes a reversible operation which does not modify the spectrum of ρ, but only changes its eigenbasis. A Gaussian unitary is one which can be generated from a Hamiltonian which is a second-order polynomial in the field operators [65]. Let us arrange the annihilation and creation operators of the corresponding N-mode bosonic system in the vectors

$$\hat{\mathbf{a}} = \left(\hat{a}_1, \ldots, \hat{a}_N\right)^{\mathrm{T}}, \quad \hat{\mathbf{a}}^\dagger = \left(\hat{a}_1^\dagger, \ldots, \hat{a}_N^\dagger\right)^{\mathrm{T}}. \tag{3.20}$$

Any Gaussian unitary U_{G} can be written in the form

$$U_{\mathrm{G}} = \exp\left[-i\hat{H}\right], \quad \hat{H} = i\left(\hat{\mathbf{a}}^\dagger \mathbf{c} + \hat{\mathbf{a}}^\dagger \mathbf{C}^{(1)} \hat{\mathbf{a}} + \hat{\mathbf{a}}^\dagger \mathbf{C}^{(2)} \hat{\mathbf{a}}^{\dagger \mathrm{T}}\right) + \mathrm{H.c.}, \tag{3.21}$$

where $\mathbf{c} \in \mathbb{C}^N$, $\mathbf{C}^{(1)}$ and $\mathbf{C}^{(2)}$ are two $N \times N$ complex matrices, and H.c. stands for Hermitian conjugate. While the action of a Gaussian unitary might be difficult to characterise in state space, as the dimension of the latter's Hilbert space is infinite, it produces a rather elementary transformation in phase space. Indeed, the corresponding operation on the quadrature field operators amounts to a simple affine mapping [65]

$$\hat{\mathbf{x}} \to \mathbf{S}\hat{\mathbf{x}} + \mathbf{d}, \tag{3.22}$$

where $\mathbf{d} \in \mathbb{R}^{2N}$ and $\mathbf{S}$ is an $N \times N$ real matrix. Obviously, this map should preserve the commutation relations (3.10). This will be the case if the matrix $\mathbf{S}$ is symplectic, *i.e.*,

$$\mathbf{S}\boldsymbol{\Omega}\mathbf{S}^{\mathrm{T}} = \boldsymbol{\Omega}, \tag{3.23}$$

$\boldsymbol{\Omega}$ being the symplectic form defined in Equation (3.9). Clearly, the eigenvalues $\mathbf{x}$ of the quadrature operators $\hat{\mathbf{x}}$ must follow the same rules. As a consequence, the evolution of the first two statistical moments $\bar{\mathbf{x}}$ and $\mathbf{V}$ which corresponds to a Gaussian unitary evolution in state space can be shown to be of the form

$$\bar{\mathbf{x}} \to \mathbf{S}\bar{\mathbf{x}} + \mathbf{d}, \quad \mathbf{V} \to \mathbf{S}\mathbf{V}\mathbf{S}^{\mathrm{T}}. \tag{3.24}$$

3.3.2 Definition of a Gaussian state

Gaussian quantum states ϱ are those whose Wigner function is given by a normal distribution, *i.e.*,

$$W_\varrho(\mathbf{x}) = \frac{1}{(2\pi)^N}\frac{1}{\sqrt{\det \mathbf{V}_\varrho}} \exp\left(-\frac{1}{2}(\mathbf{x}-\bar{\mathbf{x}}_\varrho)^{\mathrm{T}} \mathbf{V}_\varrho^{-1}(\mathbf{x}-\bar{\mathbf{x}}_\varrho)\right), \tag{3.25}$$

where $\bar{\mathbf{x}}_\varrho$ represents the displacement vector of ϱ, while $\mathbf{V}_\varrho$ is its covariance matrix. In particular, Gaussian states are completely characterised by their first two statistical moments, making them rather easy to study when adopting the phase space representation. If a Gaussian unitary is applied on a Gaussian state, the latter remains Gaussian. A direct consequence of this is the fact that the effect of a Gaussian unitary on a Gaussian state is completely characterised by Equations (3.24), making the study of such transformations on such states quite practical in phase space. When introducing Wigner functions, we mentioned that these quasiprobability distributions can be negative in general, although it is not necessarily a signature of the "quantumness" of the state characterised by the Wigner function in question. By this, we mean that the state should not necessarily have a negative Wigner function in order to exhibit quantum characteristics. A precise illustration of such a phenomenon is found in Gaussian states. Since their Wigner function is, by definition, a normal distribution, it is positive. Gaussian states will however often exhibit a purely quantum behaviour, as we will see later. As a matter of fact, pure Gaussian states happen to be the only pure quantum states with a positive Wigner function [68]. For any Gaussian state ϱ, the purity defined in Equation (3.19) depends only on the second statistical moment of the state. It can readily be computed to be

$$\mu_\varrho = \frac{1}{\sqrt{\det[\mathbf{V}_\varrho]}}. \tag{3.26}$$

The most simple pure Gaussian state, which we introduce in the next section, is the so-called vacuum state. All pure Gaussian states can be generated by applying a chosen Gaussian unitary on the vacuum state. After introducing the latter, as well as a relevant class of mixed Gaussian states (thermal states), we will present specific important Gaussian unitaries, as well as the Gaussian states they generate upon action on the vacuum state.

3.3.3 Archetypes of Gaussian unitaries and states

3.3.3.1 Vacuum state and thermal states

The vacuum state $|0\rangle$ is the most fundamental Gaussian bosonic state, as it corresponds to the ground state of each quantised mode of a bosonic field. As such, it has zero photons (if one considers the electromagnetic field), and is an eigenvector of the annihilation operator, with eigenvalue zero, *i.e.*,

$$\hat{a}|0\rangle = 0. \tag{3.27}$$

In the natural notations we chose, the covariance matrix of the vacuum state is normalised to the identity, $\mathbf{V}_{|0\rangle\langle 0|} = \mathbb{1}_2$, and it has a zero displacement vector. Its Wigner function can be found to be

$$W_{|0\rangle\langle 0|}(q,p) = \frac{1}{2\pi}\exp\left(-\frac{q^2+p^2}{2}\right). \tag{3.28}$$

The vacuum state saturates the uncertainty relation of Equation (3.17), meaning that it has the minimum product of variances in position and momentum reachable by a physically acceptable quantum state. It actually saturates several continuous variables uncertainty relations, some of which are based on the differential entropies of the marginals of the Wigner function. For a very concise review on continuous variables uncertainty relations, see *e.g.* Reference [69].

In the statistical mechanics of a quantum mechanical system, a system in thermal equilibrium is described by a so-called Kubo–Martin–Schwinger (KMS) [70], or Gibbs state. In the context of bosonic systems, it corresponds to a thermal Gaussian state, defined as

$$\tau_\varepsilon = \frac{\mathrm{e}^{-\beta\hat{n}}}{\mathrm{Tr}\left[\mathrm{e}^{-\beta\hat{n}}\right]}, \quad \beta = -\ln\varepsilon \geq 0, \tag{3.29}$$

where $\hat{n}$ is the number operator (which is, up to a constant, the Hamiltonian of one mode of the bosonic system defined in Equation (3.2)), and β can be interpreted as the inverse temperature of the system in equilibrium. As it happens, the thermal state is the bosonic state which maximises the von Neumann entropy (2.60) for a fixed energy. Define the function $g : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ ($\mathbb{R}_{\geq 0}$ being the set of all non-negative real numbers) as

$$g(x) = (x+1)\log(x+1) - x\log(x), \tag{3.30}$$

where the *log* is taken to the same base as the one considered in the von Neumann entropy (2.60). If the mean number of photons of a thermal state is $\bar{n}$, its von Neumann entropy can be computed to be $g(\bar{n})$. A more explicit form of the thermal state τ_ε can be expressed in the Fock basis as

$$\tau_\varepsilon = (1-\varepsilon)\sum_{n=0}^{\infty}\varepsilon^n |n\rangle\langle n|, \quad 0 \leq \varepsilon \leq 1, \tag{3.31}$$

which explains why we chose to parametrise it with $\varepsilon = \mathrm{e}^{-\beta}$. It will sometimes be in our interest to parametrise it using its mean number of photons, which is why we will also define it as

$$\zeta_{\bar{n}} = \sum_{n=0}^{\infty}\frac{\bar{n}^n}{(\bar{n}+1)^{n+1}}|n\rangle\langle n|, \quad \bar{n} \geq 0. \tag{3.32}$$

The relation between Equations (3.31) and (3.32) is readily computed to be such that

$$\varepsilon = \frac{\bar{n}}{\bar{n}+1}, \quad \bar{n} = \frac{\varepsilon}{1-\varepsilon}. \tag{3.33}$$

The thermal state is actually a generalisation of a vacuum state, as the former reduces to the latter when it has a zero mean number of photon. The displacement vector of the thermal state is the null vector, while its covariance matrix is proportional to the identity, and is given by

$$\mathbf{V}_{\zeta_{\bar{n}}} = (2\bar{n} + 1)\mathbb{1}_2. \tag{3.34}$$

Let us add that the thermal state is the only type of one-mode phase-invariant Gaussian state. A one-mode phase-invariant state ρ_{d} is one which is not affected by any one-mode rotation in phase space. As such, it is diagonal in the Fock basis, *i.e.*,

$$\rho_{\mathrm{d}} = \sum_i \lambda_i \left|i\right\rangle \left\langle i\right|, \tag{3.35}$$

and obviously has a Wigner function which is rotation-invariant. An example of such a phase-invariant state is found in the Fock state with 5 photons, shown in Figure 3.2.1.

As mentioned already, any pure Gaussian state can be obtained by applying a Gaussian unitary on the vacuum state. As a matter of fact, this can be further generalised by stating that any one-mode mixed Gaussian state can be obtained by applying a Gaussian unitary to a thermal Gaussian state. This is actually a result of Williamson's theorem, as we will see in Section 3.3.5, in which we provide a way to decompose any Gaussian state in terms of thermal states. Before doing so, we are going to introduce elementary examples of one or two-mode Gaussian unitaries, along with the states they generate upon action on the vacuum state, as they play a major role in quantum optics.

3.3.3.2 Displacement unitary and coherent state

The only Gaussian unitary generated by a Hamiltonian which is a first order polynomial in the bosonic field operators is of the form

$$D_\alpha = \exp\left[\alpha\hat{a}^\dagger - \alpha^*\hat{a}\right], \tag{3.36}$$

where $\alpha \in \mathbb{C}$. This so-called displacement operator can be understood as the complex version of the Weyl operator introduced earlier. Its action on the field operators in phase space is characterised by the equation

$$\hat{a} \to \hat{a} + \alpha. \tag{3.37}$$

As a trivial consequence, the displacement operator does not modify the covariance matrix of a state it acts on, while it transforms its displacement vector $\bar{\mathbf{x}}$ as

$$\bar{\mathbf{x}} \to \bar{\mathbf{x}} + \mathbf{d}_\alpha, \tag{3.38}$$

with $\mathbf{d}_\alpha = (q, p)^\mathrm{T}$, such that $\alpha = (q + ip)/2$. In the case of Gaussian states, the displacement unitary increases their energy when it moves them away from the origin of phase space.

The vacuum state is obviously not the only eigenstate of the annihilation operator. By displacing it, one generates the so-called coherent states defined as

$$|\alpha\rangle = D_\alpha |0\rangle , \tag{3.39}$$

which verify $\hat{a} |\alpha\rangle = \alpha |\alpha\rangle$. Since the displacement unitary does not modify the second statistical moment, and since the vacuum has a null displacement vector, the first two moments of the coherent state are found to be

$$\bar{\mathbf{x}}_{|\alpha\rangle\langle\alpha|} = 2 \left(\Re(\alpha), \Im(\alpha)\right)^\mathrm{T} \quad \text{and} \quad \mathbf{V}_{|\alpha\rangle\langle\alpha|} = \mathbb{1}_2. \tag{3.40}$$

The coherent state can be decomposed as a superposition of Fock states as

$$|\alpha\rangle = \mathrm{e}^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle . \tag{3.41}$$

Different coherent states $|\alpha\rangle$ and $|\beta\rangle$ are not orthogonal, although their overlap verifies

$$\langle\beta|\alpha\rangle = \mathrm{e}^{-\frac{1}{2}(|\alpha|^2+|\beta|^2)} \mathrm{e}^{\beta^* \alpha}. \tag{3.42}$$

As such, coherent states form an overcomplete continuous basis, in which one can derive the closure relation

$$\frac{1}{\pi} \int \mathrm{d}^2\alpha \, |\alpha\rangle \langle\alpha| = \mathbb{I}, \tag{3.43}$$

$\mathbb{I}$ being the identity operator.

3.3.3.3 Squeezing unitary and squeezed state

The first one-mode Gaussian unitary generated by a Hamiltonian which is quadratic in the bosonic field operators is the so-called squeezing operator, defined as

$$U_r^S = \exp\left[\frac{r}{2}\left(\hat{a}^2 - \hat{a}^{\dagger 2}\right)\right], \tag{3.44}$$

where $r \in \mathbb{R}$ is the squeezing parameter. When applied to a state ρ, the squeezing unitary decreases the variance of one of ρ's quadrature. Since the uncertainty relation (3.17) must be verified, the other quadrature of ρ has to be increased accordingly. In the Heisenberg picture, the annihilation operator is transformed by the squeezing operator as

$$\hat{a} \to (\cosh r)\hat{a} - (\sinh r)\hat{a}^\dagger. \tag{3.45}$$

Consequently, the quadrature operators are transformed as

$$\hat{\mathbf{x}} \to \mathbf{S}_r \hat{\mathbf{x}}, \tag{3.46}$$

where

$$\mathbf{S}_r = \begin{pmatrix} e^{-r} & 0 \\ 0 & e^{r} \end{pmatrix}. \tag{3.47}$$

The last two relations show how the uncertainty property verified by the two quadratures is conserved upon the action of the squeezing operator on a quantum state. Unlike the displacement unitary, the squeezing operator does not change the displacement vector of a state, only its covariance matrix. As a consequence, it can also increase the energy of a Gaussian state, but in a different way.

The state which is generated by the action of a squeezing unitary on a vacuum state is simply called squeezed vacuum state, and can be expressed in the Fock basis as [71]

$$U_r^S |0\rangle = \frac{1}{\sqrt{\cosh r}} \sum_{n=0}^{\infty} \frac{\sqrt{(2n)!}}{2^n n!} \tanh r^n |2n\rangle . \tag{3.48}$$

Its displacement vector is the same as the vacuum, *i.e.*, the null vector, while its covariance matrix is found to be

$$\mathbf{V}_{U_r^S|0\rangle\langle 0|U_r^{S\dagger}} = \mathbf{S}_r \mathbf{S}_r^{T} = \begin{pmatrix} e^{-2r} & 0 \\ 0 & e^{2r} \end{pmatrix}. \tag{3.49}$$

3.3.3.4 Phase rotation unitary

We already introduced the Gaussian unitary which induces a phase rotation. Indeed, the free Hamiltonian (3.2) of the bosonic systems happens to generate such a unitary. If one forgets about the zero-point energy of the free Hamiltonian, one can define the phase rotation operator as

$$U_\theta^R = \exp\left[-i\theta \hat{a}^\dagger \hat{a}\right]. \tag{3.50}$$

In the Heisenberg picture, it corresponds to applying the transformation

$$\hat{a} \to e^{i-\theta} \hat{a}, \tag{3.51}$$

which corresponds to a simple rotation of the (q, p) plane in phase space. Indeed, the quadrature operators are transformed as

$$\hat{\mathbf{x}} \to \mathbf{R}_\theta \hat{\mathbf{x}}, \tag{3.52}$$

where

$$\mathbf{R}_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \tag{3.53}$$

One can readily understand that the phase rotation unitary does not affect the vacuum state. However, it changes displaced and squeezed vacuum states, resulting in the generation of other types of pure Gaussian states.

3.3.3.5 Beam-splitter unitary

The most important two-mode Gaussian unitary for bosonic systems is the so-called beam splitter. As its name suggests, it is the simplest example of an interferometer, which is ubiquitous in quantum optics. The beam-splitter unitary is defined as

$$U_\eta^{\mathrm{BS}} = \exp\left[\theta\left(\hat{a}^\dagger \hat{b} - \hat{a}\hat{b}^\dagger\right)\right], \quad \eta = \cos^2\theta \in [0,1], \tag{3.54}$$

where η represents its transmittance (note that it is in our interest to parametrise the unitary using the transmittance rather than θ). Since we only have two modes, we choose to write $\hat{a}, \hat{b}$ instead of $\hat{a}_1, \hat{a}_2$ for the annihilation operators of the two modes, for convenience. The beam splitter acts in the Heisenberg picture as

$$\begin{cases} U_\eta^{\mathrm{BS}\dagger}\, \hat{a}\, U_\eta^{\mathrm{BS}} = \sqrt{\eta}\, \hat{a} + \sqrt{1-\eta}\, \hat{b}, \\ U_\eta^{\mathrm{BS}\dagger}\, \hat{b}\, U_\eta^{\mathrm{BS}} = -\sqrt{1-\eta}\, \hat{a} + \sqrt{\eta}\, \hat{b}. \end{cases} \tag{3.55}$$

As it is the realisation of the energy-conserving unitary acting on bosonic systems, the beam splitter does not modify a couple of vacua, *i.e.*, $U_\eta^{\mathrm{BS}} |0,0\rangle = |0,0\rangle$. In other words, $|0,0\rangle$ is an eigenvector of U_η^{BS} for any transmittance η (it is actually an eigenvector of its Hamiltonian, with eigenvalue zero). As a matter of fact, it can more generally be shown that a product of coherent states remains so when going through a beam splitter. Some energy will however be transmitted from one mode to the other in general (if $\eta \neq 1/2$), meaning that the parameters of the two coherent states will change. It can also be proven that if a product of thermal states is fed into a beam splitter, the resulting state is not a product state any more, but remains separable, meaning that the unitary induces no quantum correlations in this particular case. Furthermore, the two subsystems each remains in a thermal Gaussian state when evolving through the beam splitter. In Figure 3.3.1, we show the representation of an optical beam splitter, as it will be used in sketches appearing later on.

Figure 3.3.1: Representation of an optical beam splitter of transmittance η.

3.3.3.6 Two-mode squeezing unitary and Gaussian EPR state

Another two-mode Gaussian unitary essential for the study of bosonic systems is the two-mode squeezer, which is defined as

$$U_\lambda^{\mathrm{TMS}} = \exp\left[\frac{r}{2}\left(\hat{a}\hat{b} - \hat{a}^\dagger\hat{b}^\dagger\right)\right], \quad \lambda = \tanh^2 r \in [0,1]. \tag{3.56}$$

In this case, we choose to parametrise the unitary with the parameter λ instead of the squeezing r. The two-mode squeezer unitary U_λ^{TMS} models the generation of pairs of entangled photons by parametric amplification due to the pumping of a non-linear crystal. The reason it is called a two-mode squeezer is simply because it involves one-mode squeezers along with a two-mode interaction effected by a beam splitter. Indeed, as depicted in Figure 3.3.2, it can be obtained by first applying a balanced beam splitter, before applying a squeezing on one mode and the corresponding anti-squeezing on the other mode, and finally adding a second balanced beam splitter.

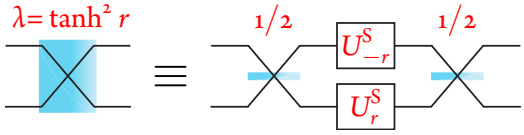

Figure 3.3.2: Representation of a two-mode squeezer of parameter λ.

In the Heisenberg picture, the annihilation operators undergo the transformation

$$\begin{cases} U_\lambda^{\mathrm{TMS}\dagger}\,\hat{a}\,U_\lambda^{\mathrm{TMS}} = \cosh(r)\,\hat{a} + \sinh(r)\,\hat{b}^\dagger, \\ U_\lambda^{\mathrm{TMS}\dagger}\,\hat{b}\,U_\lambda^{\mathrm{TMS}} = \sinh(r)\,\hat{a}^\dagger + \cosh(r)\,\hat{b}, \end{cases} \quad \lambda = \tanh^2 r. \tag{3.57}$$

When a two-mode squeezer is applied on a couple of vacua, one obtains the two-mode squeezed vacuum state, also known as Einstein-Podolski-Rosen (EPR) state $|\psi_r^{\mathrm{EPR}}\rangle$, which can be written in the Fock basis as

$$|\psi_r^{\mathrm{EPR}}\rangle = \frac{1}{\cosh r}\sum_{n=0}^{\infty}\tanh^n r\,|n,n\rangle\,. \tag{3.58}$$

It will sometimes be useful to write the EPR state in terms of the parameter $\lambda = \tanh^2 r$, using the notation

$$|\varphi_\lambda^{\mathrm{EPR}}\rangle = \sqrt{1-\lambda}\sum_{n=0}^{\infty}(\sqrt{\lambda})^n\,|n,n\rangle\,. \tag{3.59}$$

The two-mode squeezed vacuum state has a zero displacement vector, and its covariance matrix is given by

$$\mathbf{V}_{|\psi_r^{\mathrm{EPR}}\rangle\langle\psi_r^{\mathrm{EPR}}|} = \begin{pmatrix} (\cosh 2r)\mathbb{1}_2 & (\sinh 2r)\mathbf{Z} \\ (\sinh 2r)\mathbf{Z} & (\cosh 2r)\mathbb{1}_2 \end{pmatrix}, \tag{3.60}$$

where we defined the matrix

$$\mathbf{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.61}$$

Note that each subsytem of the two-mode squeezed state is in a thermal state τ_λ, of mean number of photons

$$\bar{n} = \frac{\lambda}{1-\lambda} = \frac{\tanh^2 r}{1-\tanh^2 r} = (\tanh^2 r)(\cosh^2 r) = \sinh^2 r. \tag{3.62}$$

As such, the two-mode squeezed state can be understood as the purification of a thermal state.

3.3.4 Bloch-Messiah decomposition of canonical unitaries

Any N-mode Gaussian unitary can actually be decomposed in terms of the ones introduced in Section 3.3.3. As a consequence, these archetypes of Gaussian unitaries can be thought of as building blocks for all N-mode Gaussian unitaries. This is known as the Euler decomposition [72], or the Bloch-Messiah reduction [73]. There are two types of Gaussian unitaries: passive and active. A passive unitary preserves the energy, or photon number, of the state it trasnforms. In phase space, it corresponds to a symplectic matrix $\mathbf{S}$ which preserves the trace of the covariance matrix $\mathbf{V}$ of the state, *i.e.* [65],

$$\mathrm{Tr}\left[\mathbf{S}\mathbf{V}\mathbf{S}^{\mathrm{T}}\right] = \mathrm{Tr}\left[\mathbf{V}\right], \tag{3.63}$$

which happens when the symplectic matrix is orthogonal, *i.e.*,

$$\mathbf{S}^{\mathrm{T}} = \mathbf{S}^{-1}. \tag{3.64}$$

The beam splitter described in Section 3.3.3.5 is an example of such a passive unitary, while the latter is more generally described by a multiport interferometer. An active unitary does not conserve the energy of the state it acts on. The symplectic matrix which describes it in phase space is not trace-preserving, and cannot be orthogonal. An example of such an object is readily found in the squeezing operator introduced in Section 3.3.3.3. Indeed, in order for a squeezing operation to be performed in practice, one needs to pump a non-linear cristal with a laser. If one takes into account the pump, the energy of the whole system is of course conserved.

Any symplectic matrix $\mathbf{S}$ can be decomposed using the canonical symplectic matrices introduced in Section 3.3.3. Indeed, it can always be written as

$$\mathbf{S} = \mathbf{I}^{(2)} \left(\bigoplus_{k=1}^{N} \mathbf{S}_{r_k} \right) \mathbf{I}^{(1)}, \tag{3.65}$$

where $\mathbf{I}^{(1)}$ and $\mathbf{I}^{(2)}$ are symplectic and orthogonal matrices, and all the $\mathbf{S}_{r_k}$ are squeezing ma-

trices defined in Equation 3.47. In state space, the corresponding Gaussian unitary U_G can be decomposed as

$$U_G = U_2^{PI} \left(\bigotimes_{k=1}^{N} U_{r_k}^{S} \right) U_1^{PI}. \tag{3.66}$$

Each of the $U_{r_k}^S$ is a squeezing unitary defined in Equation 3.44, while U_1^{PI} and U_2^{PI} correspond to multiport passive interferometers. These can be constructed by considering only (two-mode) beam splitters defined in Section 3.3.3.5, as well as phase rotations.

3.3.5 Thermal decomposition of Gaussian states

We already mentioned (and showed using examples) that any one-mode Gaussian state can be generated by applying a one-mode Gaussian unitary on a thermal Gaussian state. This can actually be generalised to any N-mode Gaussian state by considering the so-called thermal decomposition of Gaussian states. The latter is based on Williamson's theorem [74], which states that any positive-definite real matrix of even dimension can be put in a diagonal form using a symplectic transformation. In particular, the theorem can be applied to covariance matrices [65].

Theorem 19 (Williamson's theorem)**.** *For any N-mode covariance matrix* $\mathbf{V}$*, there exists a symplectic matrix* $\mathbf{S}$ *such that*

$$\mathbf{V} = \mathbf{S}\mathbf{V}_{\boldsymbol{\varepsilon}}\mathbf{S}^T, \quad \boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_N) \in \mathbb{R}^N, \tag{3.67}$$

where $\mathbf{V}_{\boldsymbol{\varepsilon}}$ *is a diagonal matrix which can be written as a direct sum*

$$\mathbf{V}_{\boldsymbol{\varepsilon}} = \bigoplus_{k=1}^{N} \left(\frac{1+\varepsilon_k}{1-\varepsilon_k} \right) \mathbb{1}_2. \tag{3.68}$$

Note that in this case, the subscript $\boldsymbol{\varepsilon}$ of $\mathbf{V}_{\boldsymbol{\varepsilon}}$ is a vector of real elements, in opposition with the notation $\mathbf{V}_\rho$ for the covariance matrix of a state ρ (as in Equation (3.34) for instance). We will be using the two notations in the following, the difference will be understood from the context. The matrix $\mathbf{V}_{\boldsymbol{\varepsilon}}$ is called the Williamson form of $\mathbf{V}$, and the N positive quantities

$$\nu_k = \frac{1+\varepsilon_k}{1-\varepsilon_k} \tag{3.69}$$

are called the symplectic eigenvalues of $\mathbf{V}$. They can be obtained by computing the modulus of each of the $2N$ real eigenvalues of $i\boldsymbol{\Omega}\mathbf{V}$.

The covariance matrix $\mathbf{V}_{\boldsymbol{\varepsilon}}$ actually corresponds to a tensor product $\Gamma_{\boldsymbol{\varepsilon}}$ of thermal states τ_{ε_k}

defined in Equation (3.31), *i.e.*,

$$\Gamma_{\boldsymbol{\varepsilon}} = \bigotimes_{k=1}^{N} \tau_{\varepsilon_k} = \bigotimes_{k=1}^{N} \left[(1-\varepsilon_k) \sum_{n_k=0}^{\infty} \varepsilon_k^{n_k} \left| n_k \right\rangle \left\langle n_k \right| \right], \tag{3.70}$$

meaning that in our notations,

$$\mathbf{V}_{\boldsymbol{\varepsilon}} = \mathbf{V}_{\Gamma_{\boldsymbol{\varepsilon}}}. \tag{3.71}$$

Having this in mind, one can express Williamson's theorem for density operators, or the thermal decomposition for Gaussian states. According to the latter, any N-mode Gaussian state can be obtained by starting from a tensor product of thermal states, whose parameters are exactly the ε_k appearing in the symplectic eigenvalues of Equation 3.69, before applying some N-mode unitary. In other words, we have the following corollary.

Corollary 3 (Thermal decomposition of Gaussian states)**.** *Any N-mode Gaussian state ϱ of displacement vector $\bar{\mathbf{x}}$ and covariance matrix $\mathbf{V}$ can be written as*

$$\varrho = D_{\bar{\mathbf{x}}}\, U_{\mathrm{G}}\, \Gamma_{\boldsymbol{\varepsilon}}\, U_{\mathrm{G}}^{\dagger}\, D_{\bar{\mathbf{x}}}^{\dagger}, \tag{3.72}$$

where $\Gamma_{\boldsymbol{\varepsilon}}$ is a tensor product of Gaussian thermal states τ_{ε_k}, U_{G} is a general N-mode Gaussian unitary of the form defined in Equation (3.66), and $D_{\bar{\mathbf{x}}}$ is a unitary displacing the state so that its final displacement vector is $\bar{\mathbf{x}}$.

One understands that the part $U_{\mathrm{G}}\, \Gamma_{\boldsymbol{\varepsilon}}\, U_{\mathrm{G}}^{\dagger}$ in Equation (3.72) is the translation of Equation (3.67) to state space. Figure 3.3.3 summarises the thermal decomposition.

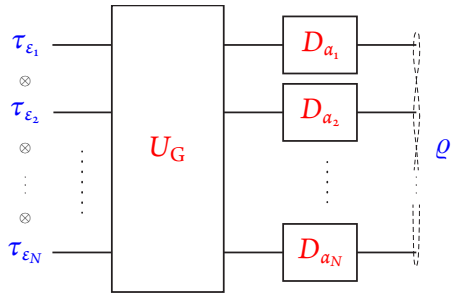

Figure 3.3.3: Representation of the thermal decomposition of an N-mode Gaussian state ϱ.

The thermal decomposition allows one to write any N-mode state as a global Gaussian unitary applied on a product of thermal states. In order to generate an entangled state, one obviously needs to apply an entangling unitary to the product of thermal states. Consider a pure Gaussian state shared by two parties, say Alice and Bob, each one having N modes. The thermal decomposition tells us that both of them can transform their state into a product of thermal states, by applying a local unitary on their side. It can be shown that by applying the right unitaries, they can end up sharing a product of N two-mode squeezed states, whose subsystems on

Alice and Bob's side are both in the same product of thermal states. This is encompassed in the following theorem [75].

Theorem 20 (Standard form of $N \times N$ pure Gaussian states)**.** *Any pure Gaussian state $|\phi\rangle$ of $N \times N$ modes shared by two parties* A *and* B *can be transformed by local unitary Gaussian operations into a state which is a tensor product of N pure two-mode squeezed states, i.e.,*

$$|\phi\rangle_{AB} = \left(D^{(A)} \otimes D^{(B)}\right)\left(U_G^{(A)} \otimes U_G^{(B)}\right)\left(\bigotimes_{k=1}^{N} |\varphi_{\lambda_k}^{EPR}\rangle_{AB}\right), \tag{3.73}$$

where

$$|\varphi_{\lambda}^{EPR}\rangle_{AB} = \sqrt{1-\lambda}\sum_{n=0}^{\infty}(\sqrt{\lambda})^n |n\rangle_A |n\rangle_B, \tag{3.74}$$

$D^{(A)}$ *and* $D^{(B)}$ *are N-mode displacements, and* $U_G^{(A)}$ *and* $U_G^{(B)}$ *are Gaussian unitaries of the form defined in Equation* (3.66).

Each of the two-mode squeezed states in the tensor product of Equation (3.73) is shared by both Alice and Bob. The situation is summarised in Figure 3.3.4. The standard form of $N \times N$ pure Gaussian states becomes useful when one needs to study the entanglement of a state shared by Alice and Bob, such as $|\phi\rangle_{AB}$ defined in (3.73). Indeed, since the latter is related to the product of two-mode squeezed states via local operations, the two states $|\phi\rangle_{AB}$ and $\bigotimes_{k=1}^{N} |\varphi_{\lambda_k}^{EPR}\rangle_{AB}$ have the same entanglement (as measured by the entropy of entanglement defined in Equation 2.80).

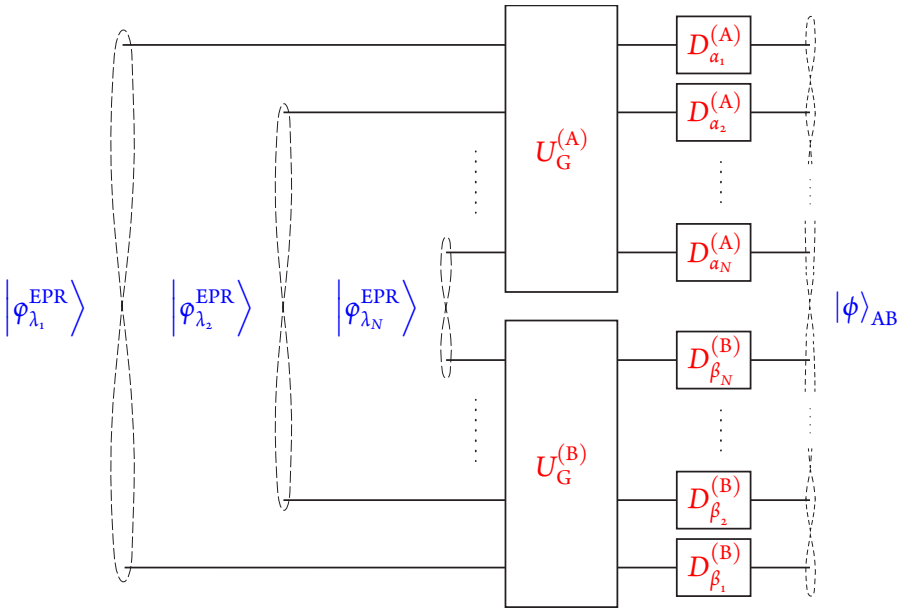

Figure 3.3.4: Representation of the standard form of an $N \times N$ pure Gaussian state ϕ.

3.4 Gaussian bosonic quantum channels

3.4.1 Definition of a Gaussian channel

Unitary transformations, such as the ones effected by the Gaussian unitaries introduced in Section 3.3, are reversible. In general, an operation acting on a quantum state does not have to be so. It should however map density matrices to density matrices. Such a quantum channel can always be described by a linear and completely positive and trace-preserving (CPTP) map [76].

Definition 26 (Quantum channel)**.** *Denote as $\mathfrak{T}(\mathcal{H})$ the space of all operators in a Hilbert space $\mathcal{H}$ equipped with the trace norm. A quantum channel $\mathcal{C} : \mathfrak{T}(\mathcal{H}) \to \mathfrak{T}(\mathcal{H})$ is*

(i) linear, i.e.,

$$\mathcal{C}\left[\sum_i c_i X_i\right] = \sum_i c_i \mathcal{C}[X_i], \quad c_i \in \mathbb{C}\,\forall i, X_i \in \mathfrak{T}(\mathcal{H})\,\forall i; \tag{3.75}$$

(ii) trace-preserving, i.e.,

$$\mathrm{Tr}[\mathcal{C}[X]] = \mathrm{Tr}[X], \quad \forall X \in \mathfrak{T}(\mathcal{H}); \tag{3.76}$$

(iii) completely positive, i.e.,

$$Y \geq 0 \quad \Rightarrow \quad (\mathcal{C} \otimes \mathbb{1}_n)[Y] \geq 0, \quad \forall Y \in \mathfrak{T}\left(\mathcal{H} \otimes \tilde{\mathcal{H}}\right), \tag{3.77}$$

where $\mathfrak{T}\left(\mathcal{H} \otimes \tilde{\mathcal{H}}\right)$ is the space of all operators in a tensored Hilbert space $\mathcal{H} \otimes \tilde{\mathcal{H}}$ and $\tilde{\mathcal{H}}$ is an n-dimensional Hilbert space for all $n = 1, 2, \ldots$

It is in our interest to also introduce the definition of an adjoint map, or dual map [76], as we will need it later. It is as follows.

Definition 27. *Denote as $\mathfrak{B}(\mathcal{H})$ the space $\mathfrak{T}(\mathcal{H})$ equipped with the operator norm. For every linear, positive and trace-preserving map $\mathcal{C} : \mathfrak{T}(\mathcal{H}) \to \mathfrak{T}(\mathcal{H})$, one can define the adjoint map $\mathcal{C}^\dagger : \mathfrak{B}(\mathcal{H}) \to \mathfrak{B}(\mathcal{H})$ through the formula*

$$\mathrm{Tr}[\mathcal{C}[X]Y] = \mathrm{Tr}\left[X\mathcal{C}^\dagger[Y]\right], \quad X \in \mathfrak{T}(\mathcal{H}), Y \in \mathfrak{B}(\mathcal{H}). \tag{3.78}$$

The following properties can be proven for the adjoint map of a linear, completely positive and trace-preserving map.

Theorem 21. *The adjoint map $\mathcal{C}^\dagger : \mathfrak{B}(\mathcal{H}) \to \mathfrak{B}(\mathcal{H})$ of a CPTP map $\mathcal{C} : \mathfrak{T}(\mathcal{H}) \to \mathfrak{T}(\mathcal{H})$ verifies the following properties.*

(i) $\mathcal{C}^\dagger$ is linear;

(ii) $\mathcal{C}^\dagger$ *is completely positive;*

(iii) $\mathcal{C}^\dagger$ *is unital (at least for finite dimensions, Definition 22).*

One understands that a channel can be seen as a linear, completely positive trace-preserving map in the Schrödinger picture, while its dual can be seen as a linear, completely positive unital map in the Heisenberg picture.

Gaussian channels are linear CPTP maps which transform Gaussian quantum states into other such states. Particular cases of such channels can precisely be found in the Gaussian unitaries presented in Section 3.3. Take $\mathcal{H}_S$ to be the Hilbert space associated to a bosonic quantum system, and $\mathcal{D}(\mathcal{H}_S)$ to be the set of density operators on $\mathcal{H}_S$. Any Gaussian channel $\mathcal{G} : \mathcal{D}(\mathcal{H}_S) \to \mathcal{D}(\mathcal{H}_S)$ acting on $\rho_S \in \mathcal{D}(\mathcal{H}_S)$ can be decomposed as

$$\mathcal{G}[\rho_S] = \mathrm{Tr}_E \left[U_G \left(\rho_S \otimes \varrho_E \right) U_G^\dagger \right], \tag{3.79}$$

where ϱ_E is a Gaussian state acting in the Hilbert space $\mathcal{H}_E$ associated to another bosonic system (typically, the environment), U_G is a Gaussian unitary acting on the tensored Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_E$, and $\mathrm{Tr}_E[\bullet]$ represents the partial trace over the system corresponding to $\mathcal{H}_E$ (the environment). According to Equation 3.79, in order to reproduce the effect of $\mathcal{G}$ on ρ_S, one should couple the latter with an ancillary Gaussian state ϱ_E of the environment through a Gaussian unitary U_G, before discarding the part of the resulting state corresponding to the environment. If ϱ_E is taken to be pure, one obtains the Stinespring dilation of a Gaussian bosonic channel [65]. One could always choose to discard the part corresponding to the main system at the output of the Gaussian unitary, retrieving the output state of the environment. In that case, one obtains the effect of the so-called complementary channel $\tilde{\mathcal{G}}$, *i.e.*,

$$\tilde{\mathcal{G}}[\rho_S] = \mathrm{Tr}_S \left[U_G \left(\rho_S \otimes \varrho_E \right) U_G^\dagger \right], \tag{3.80}$$

where $\mathrm{Tr}_S[\bullet]$ represents the partial trace over the system S. This is summarised in Figure 3.4.1.

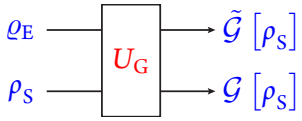

Figure 3.4.1: Representation of a Gaussian bosonic quantum channel $\mathcal{G}$ acting on ρ_S. U_G is a Gaussian unitary and ϱ_E is a Gaussian state. $\tilde{\mathcal{G}}$ represent the channel complementary to $\mathcal{G}$.

Having in mind that any Gaussian state is completely characterised by its two first statistical moments, and that such a state remains Gaussian when evolving through a Gaussian channel, one would expect the latter to be completely characterised by a finite number of parameters as

well. This is indeed the case, as the action of an N-mode Gaussian channel over an arbitrary Gaussian state ϱ of displacement vector $\bar{\mathbf{x}}$ and covariance matrix $\mathbf{V}$ can always be described in phase space as [65]

$$\bar{\mathbf{x}} \to \mathbf{T}\bar{\mathbf{x}} + \mathbf{d}, \quad \mathbf{V} \to \mathbf{T}\mathbf{V}\mathbf{T}^{\mathrm{T}} + \mathbf{N}, \tag{3.81}$$

where $\mathbf{d} \in \mathbb{R}^{2N}$ is a displacement vector, while $\mathbf{T}$ and $\mathbf{N} = \mathbf{N}^{\mathrm{T}}$ are $2N \times 2N$ real matrices, which must satisfy the condition

$$\mathbf{N} + i\boldsymbol{\Omega} - i\mathbf{T}\boldsymbol{\Omega}\mathbf{T}^{\mathrm{T}} \geq 0, \tag{3.82}$$

in order for the channel to verify the condition of complete positivity.

3.4.2 General form of one-mode Gaussian channels

One-mode Gaussian bosonic channels are of particular importance when its comes to the study of continuous-variables systems. Unlike one-mode Gaussian unitaries, they characterise the evolution of one-mode bosonic states when these are subject to the noise introduced by the environment with which they interact. As a consequence, the entropies (and spectrum) of the states transformed by noisy channels evolve as well. An arbitrary one-mode Gaussian channel is fully characterised by the maps of Equation (3.81), where $\mathbf{d} \in \mathbb{R}^2$ and $\mathbf{T}$ and $\mathbf{N} = \mathbf{N}^{\mathrm{T}}$ are 2×2 real matrices, which must satisfy

$$\mathbf{N} \geq 0, \quad \det[\mathbf{N}] \geq (\det[\mathbf{T}] - 1)^2 . \tag{3.83}$$

As it happens, any one-mode Gaussian channel $\mathcal{G}$ can be decomposed as [77]

$$\mathcal{G}[\bullet] = U_{\mathrm{G}}^{(2)} \left(\mathfrak{G} \left[U_{\mathrm{G}}^{(1)} \bullet U_{\mathrm{G}}^{(1)\dagger} \right] \right) U_{\mathrm{G}}^{(2)\dagger}, \tag{3.84}$$

where $U_{\mathrm{G}}^{(1)}$ and $U_{\mathrm{G}}^{(2)}$ are one-mode Gaussian unitaries, and $\mathfrak{G}$ is a specific one-mode Gaussian channel called the canonical form, as shown in Figure 3.4.2. As explained in [65], the effect of

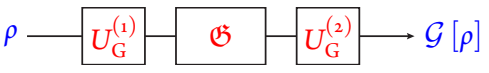

Figure 3.4.2: Decomposition of any one-mode Gaussian bosonic quantum channel $\mathcal{G}$ acting on a state ρ. $U_{\mathrm{G}}^{(1)}$ and $U_{\mathrm{G}}^{(2)}$ are one-mode Gaussian unitaries, while $\mathfrak{G}$ is the canonical form of a one-mode Gaussian bosonic channel.

the canonical form $\mathfrak{G}$ can be completely characterised by the 2×2 diagonal matrices $\mathbf{T}_{\mathfrak{G}}$ and $\mathbf{N}_{\mathfrak{G}}$, which depend on three parameters $\{\kappa, r, \bar{n}\}$ that are preserved by the action of Gaussian unitaries, *i.e.*, the generalised transmissivity

$$\kappa = \det\left[\mathbf{T}_{\mathfrak{G}}\right] \in \mathbb{R}, \tag{3.85}$$

the rank of the channel

$$r = \min\left(\text{rank}\left[\mathbf{T}_{\mathfrak{G}}\right], \text{rank}\left[\mathbf{N}_{\mathfrak{G}}\right]\right) \in \{0, 1, 2\}, \tag{3.86}$$

where rank $[\mathbf{T}_{\mathfrak{G}}]$ represents the rank of the matrix $\mathbf{T}_{\mathfrak{G}}$, and the thermal number $\bar{n} \geq 0$, defined by

$$\bar{n} = \begin{cases} \sqrt{\det\left[\mathbf{N}_{\mathfrak{G}}\right]}, & \text{for } \kappa = 1, \\ \dfrac{\sqrt{\det\left[\mathbf{N}_{\mathfrak{G}}\right]}}{2|1-\kappa|} - \dfrac{1}{2}, & \text{for } \kappa \neq 1. \end{cases} \tag{3.87}$$

The canonical form $\mathfrak{G}$ can be dilated using a three-mode Gaussian unitary $\tilde{U}_G$ parametrised by κ and r, and a two-mode squeezed state $|\varphi_\varepsilon^{EPR}\rangle$, where

$$2\bar{n} + 1 = \frac{1+\varepsilon}{1-\varepsilon}, \tag{3.88}$$

as depicted in Figure 3.4.3. Furthermore, the three parameters $\{\kappa, r, \bar{n}\}$ can be exploited in or-

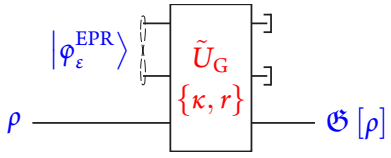

Figure 3.4.3: Dilation of the canonical form $\mathfrak{G}$ acting on ρ. $\tilde{U}_G$ is a three-mode Gaussian unitary. As one can see, the two modes which were initially in an EPR state at the input of $\tilde{U}_G$ are both traced out at its output.

der to classify the different possible canonical forms $\mathfrak{G}$. The resulting classification is shown in Table 3.4.1. The description of the different canonical classes is explained in [65]. However, we summarise it here for completeness. Class A_1 represents completely depolarising channels, as these simply replace the input state by the thermal state of the environment. Classes A_2 and B_1 involve canonical forms which transform the quadratures asymmetrically. The most interesting classes are the remaining four. Class B_2 represents the so-called classical noise channels, while class C(Loss) involves CPTP maps known as lossy channels, C(Amp) contains the forms known as amplifying channels, and D is associated with the so-called phase conjugate channels. In the following section, we characterise these last four classes in more details, as they are of particular interest for our work.

Apart from class B_2, all the other classes can be dilated using a three-mode unitary of the form $\tilde{U}_G = U_G \otimes \mathbb{I}$, where U_G acts only on the input state ρ and one mode of the two-mode squeezed state $|\varphi_\varepsilon^{EPR}\rangle$ in Figure 3.4.3. Consequently, the other mode of the two-mode squeezed state remains unchanged, and is no longer needed in the description of the canonical form $\mathfrak{G}$. Since the remaining mode of $|\varphi_\varepsilon^{EPR}\rangle$ is obviously in a thermal state τ_ε, the canonical forms of

all the classes but B_2 can be described by a single-mode thermal state interacting with the input state ρ via a two-mode Gaussian unitary U_G, as shown in Figure 3.4.4. As one would expect, the

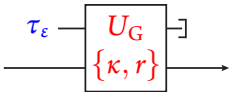

Figure 3.4.4: Dilation of the canonical form $\mathfrak{G}$ acting on ρ for all classes but B_2. U_G is a two-mode Gaussian unitary, while the environment τ_ε is in a thermal state.

thermal state τ_ε, where ε is related to the thermal noise $\bar{n}$ through Equation (3.88), parametrises the noise added by the channel. Let us mention that, when dealing with the dilation of a one-mode Gaussian bosonic channel, the mode of the system is often referred to as the signal mode, in analogy with quantum optics settings. We will often be using this terminologie later on.

3.4.3 Phase-insensitive and phase-conjugate one-mode Gaussian channels

A phase-insensitive one-mode Gaussian channel $\mathcal{G}_I$ is such that

$$\mathcal{G}_I\left[U_\theta^R \rho U_\theta^{R\dagger}\right] = U_\theta^R \mathcal{G}_I[\rho] U_\theta^{R\dagger}, \tag{3.89}$$

for any phase rotation U_θ^R and any quantum state ρ, while a phase-conjugate one mode Gaussian channel $\mathcal{G}_C$ satisfies

$$\mathcal{G}_C\left[U_\theta^R \rho U_\theta^{R\dagger}\right] = U_{-\theta}^R \mathcal{G}_C[\rho] U_{-\theta}^{R\dagger}, \tag{3.90}$$

Class	κ	r	$\mathbf{T}_\mathfrak{G}$	$\mathbf{N}_\mathfrak{G}$
A_1	0	0	0	$(2\bar{n}+1)\mathbb{1}_2$
A_2	0	1	$\frac{\mathbb{1}_2+Z}{2}$	$(2\bar{n}+1)\mathbb{1}_2$
B_1	1	1	$\mathbb{1}_2$	$\frac{\mathbb{1}_2-Z}{2}$
B_2	1	2	$\mathbb{1}_2$	$\bar{n}\mathbb{1}_2$
C(Loss)	$(0,1)$	2	$\sqrt{\kappa}\mathbb{1}_2$	$(1-\kappa)(2\bar{n}+1)\mathbb{1}_2$
C(Amp)	$(1,\infty)$	2	$\sqrt{\kappa}\mathbb{1}_2$	$(\kappa-1)(2\bar{n}+1)\mathbb{1}_2$
D	$(-\infty,0)$	2	$\sqrt{-\kappa}\mathbb{1}_2$	$(1-\kappa)(2\bar{n}+1)\mathbb{1}_2$

Table 3.4.1: Classification of canonical one-mode Gaussian bosonic channels. The first column represents the canonical class, which is specified by the possible values of the two parameters κ and r (second and third columns). The fourth and fifth column shows the expression of the two diagonal matrices $\mathbf{T}_\mathfrak{G}$ and $\mathbf{N}_\mathfrak{G}$. Note that one needs to specify the value of $\bar{n}$ as well in order to completely characterise the canonical form $\mathfrak{G}$. However, this is not needed in order to fix the canonical class to which $\mathfrak{G}$ belongs.

As such, both of these types of channels output a phase-invariant state $\mathcal{G}_{\mathrm{I,C}}[\rho_{\mathrm{d}}]$ for any input state ρ_{d} that is phase-invariant. They can be decomposed using the canonical form of Equation 3.84 as

$$\mathcal{G}_{\mathrm{I,C}}[\bullet] = \mathbf{R}_{\theta_2}\left(\mathfrak{G}\left[\mathbf{R}_{\theta_1}\bullet\mathbf{R}_{\theta_1}^{\dagger}\right]\right)\mathbf{R}_{\theta_2}^{\dagger}. \tag{3.91}$$

Compared to the decomposition of Equation 3.84, $U_{\mathrm{G}}^{(1)}$ and $U_{\mathrm{G}}^{(2)}$ can only be rotations (in opposition with displacement or one-mode squeezing operators), as exhibited by Equation 3.91. The most important one-mode Gaussian channels are the actual canonical forms involved in classes B_2, C(Loss), C(Amp) and D shown in Table 3.4.1, which we characterise in more details in the following.

3.4.3.1 Classical-noise channel

Class B_2 describes the so-called classical-noise channels (or Gaussian channels with additive classical noise). These phase-insensitive channels transform the quadratures as

$$\hat{\mathbf{x}} \to \hat{\mathbf{x}} + \boldsymbol{\xi}, \tag{3.92}$$

where $\boldsymbol{\xi}$ is a Gaussian noise with covariance matrix $\bar{n}\mathbb{1}_2$. In state space, the effect of the classical-noise channel can be written as

$$\rho \mapsto \int \mathrm{d}^2\alpha\, Q_{\bar{n}}(\alpha) D_\alpha \rho D_\alpha^{\dagger}, \tag{3.93}$$

where the D_α are displacement operators, and $Q_{\bar{n}}(\alpha)$ is a Gaussian distribution defined as

$$Q_{\bar{n}}(\alpha) = \frac{1}{\pi\bar{n}} \exp\left[-\frac{|\alpha|^2}{\bar{n}}\right]. \tag{3.94}$$

In other words, the channel randomly displaces the input state according to the Gaussian distribution $Q_{\bar{n}}(\alpha)$. As a result, the displacement vector of the input state is not changed, while its covariance matrix $\mathbf{V}$ is transformed according to

$$\mathbf{V} \to \mathbf{V} + \bar{n}\mathbb{1}_2. \tag{3.95}$$

3.4.3.2 Lossy channel

Class C(Loss) describes the so-called lossy channels, which represent the basic model to describe communication lines such as optical fibres. In this work, we denote the lossy channels by $\mathcal{B}_\eta^{(\varepsilon)}$. Notice that the superscript is surrounded by parentheses, which is a reminder of the fact that it is a real (non-negative) number. This is important, as we will define other channels using a discrete parameter later in this work. The lossy channel is phase-insensitive, and its effect on

a state ρ in state space can be written as

$$\mathcal{B}_\eta^{(\varepsilon)}[\rho] = \mathrm{Tr}_2\left[U_\eta^{\mathrm{BS}}\left(\rho \otimes \tau_\varepsilon\right) U_\eta^{\mathrm{BS}\dagger}\right], \tag{3.96}$$

where U_η^{BS} is a beam splitter of transmittance η, while τ_ε is a thermal state of parameter ε. Note that $\mathrm{Tr}_2[\bullet]$ means that the second subsystem is traced out at the output of the beam splitter. Here, the transmittance η is exactly the parameter κ of Equation (3.85) characterising class C(Loss). As depicted in Figure 3.4.5, Equation (3.96) simply models the fact that the input state ρ of the signal mode interacts with a thermal state of the environment through a beam splitter, before the environment mode is discarded. In phase space it means that the quadra-

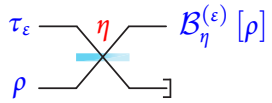

Figure 3.4.5: Lossy channel $\mathcal{B}_\eta^{(\varepsilon)}$. The input state interacts with a thermal state of parameter ε through a beam splitter of transmittance η. The environment is then discarded.

tures $\hat{\mathbf{x}}$ of the input state are transformed according to

$$\hat{\mathbf{x}} \to \sqrt{\eta}\hat{\mathbf{x}} + \sqrt{1-\eta}\hat{\mathbf{x}}_{\mathrm{th}}, \tag{3.97}$$

where $\hat{\mathbf{x}}_{\mathrm{th}}$ represent the quadratures of the thermal state τ_ε. The effect of the lossy channel on the first two statistical moments of the input state can be shown to be

$$\bar{\mathbf{x}} \to \sqrt{\eta}\bar{\mathbf{x}}, \quad \mathbf{V} \to \eta\mathbf{V} + (1-\eta)\frac{1+\varepsilon}{1-\varepsilon}\mathbb{1}_2, \tag{3.98}$$

since the number of photon of the thermal state in the environment of the dilation of the lossy channel is $\bar{n} = \varepsilon/(1-\varepsilon)$. If the environment is taken to be in the vacuum state, as shown in Figure 3.4.6, one obtains the so-called pure-loss channel $\mathcal{B}_\eta^{(0)}$, which we also label as $\mathcal{B}_\eta$.

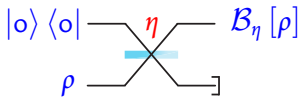

Figure 3.4.6: Pure-loss channel $\mathcal{B}_\eta = \mathcal{B}_\eta^{(0)}$. The input state interacts with a vacuum state through a beam splitter of transmittance η. The environment is then discarded.

3.4.3.3 Amplifier channel

Class C(Amp) describes the so-called amplifier channels, which amplify input signals while adding some thermal noise. We will denote the amplifier channels by $\mathcal{A}_G^{(\varepsilon)}$. Again, the superscript is surrounded by parentheses, since it is a real (non-negative) number. Like the lossy

channel, the amplifier channel is phase-insensitive. It is deifned as

$$\mathcal{A}_G^{(\varepsilon)}[\rho] = \mathrm{Tr}_2\left[U_\lambda^{\mathrm{TMS}}\left(\rho \otimes \tau_\varepsilon\right) U_\lambda^{\mathrm{TMS}\dagger}\right], \quad \lambda = \frac{G-1}{G}, \tag{3.99}$$

where U_λ^{TMS} is a two-mode squeezer of parameter λ, while τ_ε is a thermal state of parameter ε. Note that we chose to parametrise the channel using a gain $G > 1$ rather than the parameter λ of the two-mode squeezer U_λ^{TMS}. This gain is exactly the parameter κ of Equation (3.85) defining the class C(Amp), *i.e.*, $\kappa = G$ in this case. As shown in Figure 3.4.7, it means that the input state ρ of the signal mode interacts with a thermal state of the environment (called the idler mode) through a two-mode squeezer, before the idler mode is discarded. In phase space, the

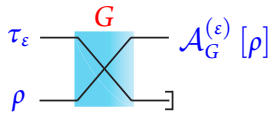

Figure 3.4.7: Amplifier channel $\mathcal{A}_G^{(\varepsilon)}$. The input state interacts with a thermal state of parameter ε through a two-mode squeezer U_λ^{TMS} of gain $G = 1/(1-\lambda)$. The environment is then discarded.

quadratures $\hat{\mathbf{x}}$ of the input state evolve following

$$\hat{\mathbf{x}} \to \sqrt{G}\hat{\mathbf{x}} + \sqrt{G-1}\hat{\mathbf{x}}_{\mathrm{th}}, \tag{3.100}$$

where $\hat{\mathbf{x}}_{\mathrm{th}}$ represent the quadratures of the thermal state τ_ε. The first two statistical moments of the input state are transformed according to

$$\bar{\mathbf{x}} \to \sqrt{G}\bar{\mathbf{x}}, \quad \mathbf{V} \to G\mathbf{V} + (G-1)\frac{1+\varepsilon}{1-\varepsilon}\mathbb{1}_2. \tag{3.101}$$

If the idler mode is taken to be in the vacuum state, as shown in Figure 3.4.8, one obtains the so-called quantum-limited amplifier $\mathcal{A}_G^{(0)}$, which we also label as $\mathcal{A}_G$ in the following.

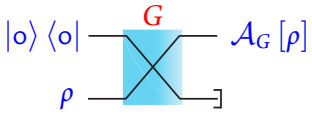

Figure 3.4.8: Quantum-limited amplifier $\mathcal{A}_G = \mathcal{A}_G^{(0)}$. The input state interacts with a vacuum state through a two-mode squeezer U_λ^{TMS} of gain $G = 1/(1-\lambda)$. The environment is then discarded.

3.4.3.4 Phase-conjugate channel

Class D describes the simplest phase-conjugate channels, which are associated with negative transmissivities κ (we simply call them phase-conjugate channels in the following, as they are

the ones we happen to be interested in). They will be denoted by $\tilde{\mathcal{A}}_G^{(\varepsilon)}$ in this work. They correspond to the complementary channel (See Equation 3.80) of the amplifier. As such, they are defined as

$$\tilde{\mathcal{A}}_G^{(\varepsilon)}[\rho] = \mathrm{Tr}_1\left[U_\lambda^{\mathrm{TMS}}\left(\rho \otimes \tau_\varepsilon\right) U_\lambda^{\mathrm{TMS}\dagger}\right], \quad \lambda = \frac{G-1}{G}, \tag{3.102}$$

where U_λ^{TMS} is a two-mode squeezer of parameter λ, while τ_ε is a thermal state of parameter ε. In this case, the partial trace $\mathrm{Tr}_1[\bullet]$ is performed over the signal mode at the output of the two-mode squeezer. Similarly to the case of the amplifier channel, we chose to parametrise the phase-conjugate channel using the gain G rather than the parameter λ of the two-mode squeezer U_λ^{TMS}. This time, the relationship between the parameter defining class D and the gain is such that $\kappa = -(G-1)$. As shown in Figure 3.4.9, Equation (3.102) models the fact that the input state ρ of the signal mode interacts with a thermal state of the idler mode through a two-mode squeezer, before the signal mode is discarded. In phase space, the quadratures $\hat{\mathbf{x}}$ of the input

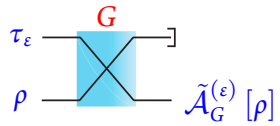

Figure 3.4.9: Phase-conjugate channel $\tilde{\mathcal{A}}_G^{(\varepsilon)}$. The input state interacts with a thermal state of parameter ε through a two-mode squeezer U_λ^{TMS} of gain $G = 1/(1-\lambda)$. The signal mode is then discarded.

state are transformed according to

$$\hat{\mathbf{x}} \to \sqrt{G}\hat{\mathbf{x}} + \sqrt{G-1}\hat{\mathbf{x}}_{\mathrm{th}}, \tag{3.103}$$

where $\hat{\mathbf{x}}_{\mathrm{th}}$ represent the quadratures of the thermal state τ_ε. The phase-conjugate channel transforms the first two statistical moments of the input state as

$$\bar{\mathbf{x}} \to \sqrt{G-1}\bar{\mathbf{x}}, \quad \mathbf{V} \to (G-1)\mathbf{V} + G\frac{1+\varepsilon}{1-\varepsilon}\mathbb{1}_2. \tag{3.104}$$

Finally, if the idler mode is taken to be in the vacuum state, as shown in Figure 3.4.10, one obtains the so-called quantum-limited phase-conjugate channel $\tilde{\mathcal{A}}_G^{(0)}$, which we also label as $\tilde{\mathcal{A}}_G$ in the following.

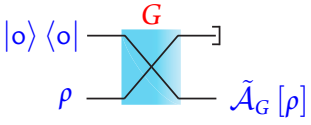

Figure 3.4.10: Quantum-limited phase-conjugate channel $\tilde{\mathcal{A}}_G = \tilde{\mathcal{A}}_G^{(0)}$. The input state interacts with a vacuum state through a two-mode squeezer U_λ^{TMS} of gain $G = 1/(1-\lambda)$. The signal mode is then discarded.

3.4.4 Quantum-limited decomposition of one-mode Gaussian channels

Among the canonical classes having a channel rank $r = 2$ (Equation (3.86)), the four classes B_2, C(Loss), C(Amp) and D can be discriminated using the generalised transmissivity κ of Equation (3.85). The channels with a non-negative value of κ, *i.e.*, the lossy channels and amplifier channels, are sometimes called gauge-covariant, while the phase-conjugate channels, whose value of κ is negative, are often called gauge-contravariant. Apart from the generalised transmissivity, the channels of the four important classes mentioned above are characterised by the thermal number $\bar{n}$. As already mentioned, the latter is exactly the photon number of the thermal state appearing in the environment mode of the dilation depicted in Figure 3.4.4. Consequently, this thermal number $\bar{n}$ can be related to what is thought of as the noise introduced by the channel. We define this noise as

$$\aleph = \sqrt{\det\left[\mathbf{N}_{\mathfrak{G}}\right]} = \begin{cases} |1-\kappa|(2\bar{n}+1), & \text{for C(Loss), C(Amp) and D,} \\ \bar{n}, & \text{for } B_2. \end{cases} \tag{3.105}$$

For the channels of classes C(Loss), C(Amp) and D to be physical, their noise must satisfy Equation (3.83), which is equivalent to stating that $\aleph \geq |1-\kappa|$. This is simply a consequence of the fact that the thermal number $\bar{n}$ should be non-negative. When it is zero, the inequality is saturated, meaning that the channel is quantum-limited. Figure 3.4.11, inspired from [78], provides a representation of the most important one-mode phase-insensitive Gaussian channels. It is interesting to mention that the channels for which $\aleph \geq |\kappa| + 1$ can be shown to be

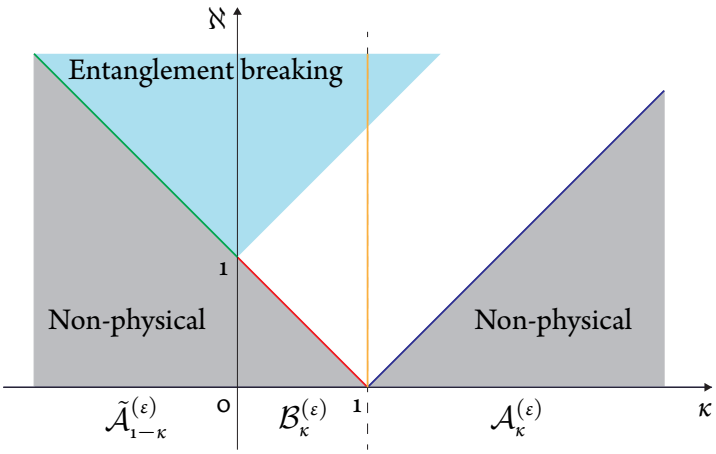

Figure 3.4.11: Classification of one-mode phase-insensitive Gaussian channels. The grey zones represent non-physical channels, which do not verify $\aleph \geq |1-\kappa|$. The parameter ε of the channels can be related to the noise $\aleph$ using Equations (3.105) and (3.88). The red line represents pure-loss channels $\mathcal{B}_\kappa$, the blue one quantum-limited amplifiers $\mathcal{A}_\kappa$, and the green one quantum-limited phase-conjugate channels $\tilde{\mathcal{A}}_{1-\kappa}$. The orange line represents general classical-noise channels. The light blue zone represents entanglement-breaking channels, characterised by $\aleph \geq |\kappa| + 1$.

entanglement-breaking [79], as enough noise is introduced in order for quantum correlations to break down. As a consequence, all phase-conjugate channels are entanglement-breaking.

When two Gaussian bosonic channels are concatenated, the resulting map can be shown to be a Gaussian bosonic channel as well [80]. This property can be exploited in order to prove that any phase-insensitive (phase-conjugate) channel can be obtained by concatenating a pure-loss channel and a quantum-limited amplifier (quantum-limited phase-conjugate channel) [81]. In appendix C, we show that any lossy channel $\mathcal{B}_{\eta}^{(\varepsilon)}$ can always be written as

$$\mathcal{B}_{\eta}^{(\varepsilon)} = \mathcal{A}_{G_0} \circ \mathcal{B}_{\eta_0}, \quad \bar{n} = \frac{\varepsilon}{1-\varepsilon}, \tag{3.106}$$

where $\eta_0 \in [0,1]$ and $G_0 \geq 1$ are chosen so that

$$\begin{cases} \eta = G_0 \eta_0, \\ \bar{n} = \dfrac{G_0 - 1}{1 - \eta_0 G_0}, \end{cases} \quad \Leftrightarrow \quad \begin{cases} \eta_0 = \dfrac{1-\varepsilon}{1-\eta\varepsilon}\eta, \\ G_0 = \dfrac{1-\eta\varepsilon}{1-\varepsilon}. \end{cases} \tag{3.107}$$

Similarly, any amplifier channel $\mathcal{A}_{G}^{(\varepsilon)}$ can always be obtained as

$$\mathcal{A}_{G}^{(\varepsilon)} = \mathcal{A}_{G_0} \circ \mathcal{B}_{\eta_0}, \quad \bar{n} = \frac{\varepsilon}{1-\varepsilon}, \tag{3.108}$$

where $\eta_0 \in [0,1]$ and $G_0 \geq 1$ satisfy

$$\begin{cases} G = G_0 \eta_0, \\ \bar{n} = \dfrac{G_0(1-\eta_0)}{\eta_0 G_0 - 1}, \end{cases} \quad \Leftrightarrow \quad \begin{cases} \eta_0 = \dfrac{1-\varepsilon}{G-\varepsilon}G, \\ G_0 = \dfrac{G-\varepsilon}{1-\varepsilon}. \end{cases} \tag{3.109}$$

Finally, any phase-conjugate channel $\tilde{\mathcal{A}}_{G}^{(\varepsilon)}$ can always be obtained as

$$\tilde{\mathcal{A}}_{G}^{(\varepsilon)} = \tilde{\mathcal{A}}_{G_0} \circ \mathcal{B}_{\eta_0}, \quad \bar{n} = \frac{\varepsilon}{1-\varepsilon}, \tag{3.110}$$

where $\eta_0 \in [0,1]$ and $G_0 \geq 1$ satisfy

$$\begin{cases} G = \eta_0(G_0 - 1) + 1, \\ \bar{n} = \dfrac{(G_0-1)(1-\eta_0)}{1 + (G_0-1)\eta_0}, \end{cases} \quad \Leftrightarrow \quad \begin{cases} \eta_0 = \dfrac{(G-1)(1-\varepsilon)}{\varepsilon + G - 1}, \\ G_0 = \dfrac{G}{1-\varepsilon}. \end{cases} \tag{3.111}$$

3.4.5 Master equations for one-mode Gaussian channels

A general quantum channel represents the evolution of an open quantum system, as it corresponds to the interaction of the system (the signal mode) with some environment. As a result,

the quantum dynamics of the system can obviously not be reproduced in terms of a unitary evolution, in contrast to the case of a closed system [82]. In many cases, it turns out to be convenient to formulate the dynamics of the open system by means of an equation of motion for the density matrix, a quantum master equation. This description turns out to be very useful in the case of Gaussian channels, particularly for one-mode gauge-covariant channels $\mathcal{G}_\kappa^{(\varepsilon)}$. The latter can be shown to possess a semi-group structure [83], and can consequently be represented as a one-parameter linear CPTP map

$$\rho(t) = \mathcal{G}_\kappa^{(\varepsilon)}[\rho] = e^{t\mathcal{L}}[\rho], \quad t = f(\kappa) \geq 0, \tag{3.112}$$

where the continuous parameter t is some function of the generalised transmissivity κ, $t = f(\kappa)$, that can be viewed as a time which characterises the continuous action of the channel on the input state $\rho = \rho(0)$, resulting in the output state $\rho(t)$. The so-called Lindblad operator $\mathcal{L}$ is a function of the parameter ε that generates the dynamics of $\mathcal{G}_\kappa^{(\varepsilon)}$. The Lindblad operator can be decomposed as

$$\mathcal{L} = \gamma_+ \mathcal{L}_+ + \gamma_- \mathcal{L}_-, \tag{3.113}$$

with

$$\begin{cases} \mathcal{L}_+[\rho] = \hat{a}^\dagger \rho a - \frac{1}{2} a\hat{a}^\dagger \rho - \frac{1}{2} \rho a \hat{a}^\dagger, \\ \mathcal{L}_-[\rho] = a\rho\hat{a}^\dagger - \frac{1}{2}\hat{a}^\dagger a\rho - \frac{1}{2}\rho\hat{a}^\dagger a. \end{cases} \tag{3.114}$$

Equation (3.112) implies that the gauge-covariant channel $\mathcal{G}_\kappa^{(\varepsilon)}$ has a semi-group structure

$$e^{(t_1+t_2)\mathcal{L}} = e^{t_1\mathcal{L}} e^{t_2\mathcal{L}} = e^{t_2\mathcal{L}} e^{t_1\mathcal{L}}, \quad t_1, t_2 \geq 0, \tag{3.115}$$

and obeys a master equation of the form

$$\frac{\partial}{\partial t}\rho(t) = \mathcal{L}[\rho(t)]. \tag{3.116}$$

In the case of a lossy channel $\mathcal{B}_\eta^{(\varepsilon)}$, one has

$$\gamma_+ = \bar{n} = \frac{\varepsilon}{1-\varepsilon}, \quad \gamma_- = \bar{n} + 1 = \frac{1}{1-\varepsilon}, \tag{3.117}$$

while the transmissivity $\kappa = \eta$ satisfies $\eta = e^{-t}$. For an amplifier channel $\mathcal{A}_G^{(\varepsilon)}$,

$$\gamma_+ = \bar{n} + 1 = \frac{1}{1-\varepsilon}, \quad \gamma_- = \bar{n} = \frac{\varepsilon}{1-\varepsilon}, \tag{3.118}$$

while the gain $\kappa = G$ is related to the time parameter through $G = e^t$. Finally, the classical-noise channel is such that

$$\gamma_+ = 1, \quad \gamma_- = 1, \tag{3.119}$$

so that it is characterised by a Lindbladian

$$\mathcal{L}_{\circ} = \mathcal{L}_{-} + \mathcal{L}_{+}, \tag{3.120}$$

with $\bar{n} = t$.

4

Entropic inequalities for bosonic quantum systems

In this chapter, we focus on a certain type of entropic inequalities for bosonic quantum systems, specifically the so-called entropy photon-number inequality. In order to do so, we begin by introducing its classical counterpart. The entropy power inequality is one of, if not the most elegant relation introduced by Claude Shannon in 1948 in his historical paper [1]. It asserts that the entropy power of the sum of independent random variables is at least the sum of their entropy powers. It enjoys many applications, among which the derivation of bounds for the capacities of certain types of channels, or the proof of some source coding theorems. The extension of the entropy power inequality to the quantum realm resides in the entropy photon-number inequality. It was introduced by Guha in his thesis [7] in an attempt to generalise the so-called minimum output entropy conjectures applied to Gaussian bosonic quantum channels. The conjectures were brought forth in order to find expressions for the capacities of several of these channels, some of which are the bosonic broadcast channel and the bosonic multiple-access channel. Despite the general effort carried out towards the proof of the entropy photon-number inequality, it still remains a conjecture.

For an exhaustive review on the entropy power inequality, the interested reader is referred to [84]. For more information on the entropy photon-number inequality and its connection with the computation of the capacity of Gaussian bosonic channel, we suggest Guha's thesis [7], in which the conjecture was originally formulated.

4.1 The entropy power inequality and beyond

4.1.1 Stam's inequality

We begin by introducing the concept of Fisher information, as it allows one to construct an elegant proof (essentially proposed by Stam [85]) of the entropy power inequality, even though we do not present all the details of the proof itself. Consider a random variable X, and suppose that its probability density f depends on some unknown parameter θ. The Fisher information provides a way of measuring the amount of information that X carries about θ. It can be defined as follows [24].

Definition 28 (Fisher information)**.** *Consider a random variable X with a probability density f depending on some unknown parameter θ. The Fisher information J of X with respect to θ is defined as*

$$J_\theta(X) = \int_{-\infty}^{\infty} \mathrm{d}x\, f(x;\theta) \left[\frac{\partial}{\partial\theta} \ln f(x;\theta)\right]^2 . \tag{4.1}$$

Notice that like the differential entropy, the Fisher information only depends on its probability density. Now, suppose the parameter θ we introduce is some "location" parameter, so that the dependence of the density on the latter is such that $f(x;\theta) = f(x-\theta)$. In this case, it can be shown using a simple change of variables that the Fisher information does not depend on θ, and can be written

$$J(X) = \int_{-\infty}^{\infty} \mathrm{d}x\, f(x) \left[\frac{\partial}{\partial x} \ln f(x)\right]^2 , \tag{4.2}$$

or,

$$J(X) = \int_{-\infty}^{\infty} \mathrm{d}x\, \frac{1}{f(x)} \left[\frac{\partial}{\partial x} f(x)\right]^2 . \tag{4.3}$$

The Fisher information satisfies an interesting convolution inequality [85], which we state in the following theorem.

Theorem 22 (Stam's inequality)**.** *If X and Y are two independent real-valued random variables, then their Fisher information satisfy*

$$\frac{1}{J(X+Y)} \geq \frac{1}{J(X)} + \frac{1}{J(Y)}. \tag{4.4}$$

Stam's inequality is crucial in proving some important relations satisfied by Shannon's differential entropy. Indeed, the Fisher information shares a fundamental relationship with the latter, known as de Bruijn's identity, which is encompassed in the following theorem [24].

Theorem 23 (de Bruijn's identity)**.** *Let X be any random variable with a finite variance and a density f. Consider an independent normally distributed random variable W_G with zero mean and*

unit variance. Then

$$\frac{\partial}{\partial t} h_e(X + \sqrt{t} W_G) = \frac{1}{2} J(X + \sqrt{t} W_G), \tag{4.5}$$

where h_e is the differential entropy to base e. In particular, if the limit exists as $t \to 0$,

$$\left[\frac{\partial}{\partial t} h_e(X + \sqrt{t} W_G) \right] \Big|_{t=0} = \frac{1}{2} J(X). \tag{4.6}$$

Note that, unlike in previous chapters, we chose to write the differential entropy as a function of the random variable X instead of its probability density, as it makes things clearer when considering sums of variables. The message conveyed by Equation (4.6) is that when some Gaussian noise W_G of unit variance is added to a variable X, the rate of change of the entropy of the latter is directly proportional to the Fisher information of Equation (4.3). This fact actually provides a different interpretation of the Fisher information than the one given in terms of parameter estimation. Stam introduced his inequality (Theorem 22) in his proof of the so-called entropy power inequality (EPI). The proof was later further simplified and made more rigorous by Blachman [86] and others [87, 88]. The entropy power inequality (EPI) was put forth by Shannon in his seminal work on the theory of communication [1]. It states that if X and Y are two independent real-valued random variables, then

$$e^{2h_e(X+Y)} \geq e^{2h_e(X)} + e^{2h_e(Y)}. \tag{4.7}$$

Shannon first used the entropy power inequality in the classical setting in order to bound the capacity of non-Gaussian additive noise channels. His relation enjoys many applications, among which the proofs of converses of channel or source coding theorems.

4.1.2 The entropy power

Equation (4.7) can be restated in terms of the so-called entropy powers (hence the denomination entropy power inequality). In order to do so, we define them now. Consider a normally distributed random variable X_G of variance σ_X^2 and centred on zero, whose probability density f is defined similarly to Equation (3.25), *i.e.*,

$$X_G \sim \mathcal{N}(0, \sigma_X^2), \quad f(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{x^2}{2\sigma_X^2}}. \tag{4.8}$$

Its differential entropy can readily be computed to be

$$h_e(X_G) = \frac{1}{2} \ln \left(2\pi e \sigma_X^2 \right). \tag{4.9}$$

Furthermore, the normally distributed variable happens to be the one with the maximum entropy for a fixed variance. In other words, for any variable X with a variance σ_X^2, we have the

inequality

$$h_e(X) \leq h_e(X_G). \tag{4.10}$$

The entropy power of a random variable is defined as follows.

Definition 29 (Entropy power)**.** *The entropy power P of a random variable X is defined as the variance of the normally distributed variable having the same differential entropy as X, i.e.,*

$$P(X) = \frac{1}{2\pi e} e^{2h_e(X)}. \tag{4.11}$$

Obviously, the entropy power of the normally distributed variable X_G is equal to its variance,

$$P(X_G) = \sigma_X^2. \tag{4.12}$$

Equation (4.10) is equivalent to the statement

$$P(X) \leq \sigma_X^2, \tag{4.13}$$

where X has a variance σ_X^2. One way to show this is by considering the relative entropy between two random variables defined in Equation (2.36). Equation (4.11) is obviously equivalent to

$$h_e(X) = \frac{1}{2} \ln \left(2\pi e P(X)\right), \tag{4.14}$$

which is the same as Equation (4.9) for $X_G \sim \mathcal{N}(\mu_X, \sigma_X^2)$. Using this, it is easy to show that the continuous relative entropy between X and the normally distributed variable X_G is such that

$$D(X||X_G) = \frac{1}{2} \ln \frac{\sigma_X^2}{P(X)}, \tag{4.15}$$

where X and X_G both have variances σ_X^2, and we chose to write the continuous relative entropy as a function of the variables rather than their densities. The last relation is equivalent to

$$\sigma_X^2 = P(X) e^{2D(X||X_G)}. \tag{4.16}$$

Since the continuous relative entropy is always non-negative, we end up with (4.13).

A useful property of the entropy power can be obtained by considering the scaling property (2.35) of the differential entropy. Indeed, for any $a \in \mathbb{R}$, one has

$$P(aX) = \frac{1}{2\pi e} e^{2h_e(aX)} = \frac{1}{2\pi e} e^{2h_e(X)} e^{2\ln|a|}, \tag{4.17}$$

so that

$$P(aX) = |a|^2 P(X). \tag{4.18}$$

4.1.3 Equivalent forms of the entropy power inequality

Now that we defined the concept of entropy power, we can equivalently restate the entropy power inequality in terms of the latter. Indeed, equation (4.7) becomes

$$P(X+Y) \geq P(X) + P(Y). \tag{4.19}$$

As a matter of fact, using the scaling property of Equation (4.18), it can be shown to be completely equivalent to the following statement [84].

Theorem 24 (Entropy power inequality)**.** *Let X and Y be two independent random variables taking values in $\mathbb{R}$ and let $Z = aX + bY$, where $a \in \mathbb{R}$ and $b \in \mathbb{R}$. Then*

$$P(Z) \geq a^2 P(X) + b^2 P(Y). \tag{4.20}$$

Notice that if two independent random variables are both normally distributed such that $X_G \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and $Y_G \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$, the new variable $Z_G = aX_G + bY_G$ is also normally distributed, and verifies

$$Z_G \sim \mathcal{N}(a\mu_X + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2). \tag{4.21}$$

The entropy power inequality is saturated for normally distributed variables. Now, suppose that X and Y defined in Equation (4.20) have the same differential entropies as X_G and Y_G, respectively. In that case,

$$\begin{aligned} a^2 P(X) + b^2 P(Y) &= a^2 P(X_G) + b^2 P(Y_G) \\ &= a^2\sigma_X^2 + b^2\sigma_Y^2, \end{aligned}$$

so that

$$a^2 P(X) + b^2 P(Y) = P(Z_G). \tag{4.22}$$

Using this, we see that Equation (4.20) is in turn equivalent to the statement

$$P(Z) \geq P(Z_G), \tag{4.23}$$

or,

$$h_e(Z) \geq h_e(Z_G), \tag{4.24}$$

where $Z = aX + bY$, $Z_G = aX_G + bY_G$, $a \in \mathbb{R}$, $b \in \mathbb{R}$, and X_G and Y_G are the two normally distributed variables having the same differential entropies as X and Y, respectively. This is yet another way of stating the entropy power inequality. Equation (4.24) is however of particular importance, as it will present the best analogy with some generalisations we introduce later.

Finally, the entropy power inequality can be readily shown to be completely equivalent to the statement [89]

$$h_e(aX + bY) \geq a^2 h_e(X) + b^2 h_e(Y), \tag{4.25}$$

for any two independent random variables X and Y, and any $a \in \mathbb{R}, b \in \mathbb{R}$ such that $a^2 + b^2 = 1$. As mentioned already, Shannon's original proof was not entirely rigorous. The alternative proofs which made it so are actually proofs of Equation (4.25).

Table 4.1.1 presents a summary of the most important equivalent forms of the entropy power inequality. Note that all these forms can readily be generalised to the sum of an arbitrary number of rescaled variables.

Eq.	Statement	Conditions
(4.20)	$P(aX + bY) \geq a^2 P(X) + b^2 P(Y)$	\
	$\Updownarrow$	
(4.24)	$h_e(aX + bY) \geq h_e(aX_G + bY_G)$	$h_e(X) = h_e(X_G)$ $h_e(Y) = h_e(Y_G)$
	$\Updownarrow$	
(4.25)	$h_e(aX + bY) \geq a^2 h_e(X) + b^2 h_e(Y)$	$a^2 + b^2 = 1$

Table 4.1.1: Equivalent forms of the entropy power inequality. $h_e(X)$ is the Shannon differential entropy defined with a logarithm in natural basis, computed for the density f_x of the random variable X. For the three forms, we have $a \in \mathbb{R}, b \in \mathbb{R}$.

4.1.4 Beyond the entropy power inequality via rearrangements

Equation (4.24) basically states that if one consider all variables X having some fixed differential entropy, and all variables Y having some other fixed differential entropy, the optimal couple of independent variables will be given by two normally distributed variables. Here, they are optimal in the sense that they will produce the minimum differential entropy possible after a rescaling followed by a summation. Since the concept of entropy is closely related to the notion of disorder introduced in Chapter 2, it seems natural to ask whether some similar statement can be brought forward in the context of the mathematical theory of majorization. Denote by $f_X \star f_Y$ the convolution of two functions f_X and f_Y. The following theorem involving rearrangements of non-negative functions, defined in Chapter 2, was proven in [43].

Theorem 25. *If X and Y are two independent random variables with respective densities f_X and f_Y, then*

$$f_X \star f_Y \prec f_X^{\downarrow} \star f_Y^{\downarrow}. \tag{4.26}$$

Note that Theorem 25 was proven for a convolution of k densities $f_1 \star f_2 \star \ldots \star f_k$, but we only state it for a convolution of two functions here. In comparison with Equation (4.24) (in which one chooses $a = b = 1$), the message conveyed by Equation (4.26) is that if one consider all densities f_X having some fixed m_{f_X} (defined in Equation (2.41) of Chapter 2), and all densities f_Y having some other fixed m_{f_Y}, the optimal couple of independent densities will be given by two spherically decreasing symmetric functions. This time, they are optimal in the sense that they will produce the minimum disorder possible, or majorize all others, after a convolution of f_X and f_Y.

4.2 The entropy photon-number inequality

Shannon's entropy power inequality proved essential for the investigation of the capacities of noisy classical channels, as it was used to prove bounds for the latter. Similarly, the computation of the capacities of some specific quantum channels requires the proof of particular quantum entropic inequalities, akin to the entropy power inequalities. These are the entropy photon-number inequalities. As we are not interested in the capacities of quantum channels in this work, we rather choose to focus on the quantum entropic inequalities, which we introduce in the following. In order to do so, we begin by introducing the concept of entropy photon-number, analogous to the notion of entropy power of classical information theory.

4.2.1 The entropy photon-number

Consider a thermal Gaussian state $\zeta_{\bar{n}}$ defined by its mean number of photons $\bar{n}$, as defined in Equation (3.32). As already mentioned, its von Neumann entropy can be computed using the function g of (3.30) as

$$S(\zeta_{\bar{n}}) = g(\bar{n}). \tag{4.27}$$

In analogy to the definition of the entropy power with respect to a normal distribution, one can define the entropy photon-number $\mathfrak{N} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ of a one-mode bosonic quantum state ρ to be the mean number of photons of the thermal Gaussian state having the same von Neumann entropy as ρ. In other words, it is given by the inverse function of g, *i.e.*,

$$\mathfrak{N}(\rho) = g^{-1}\left(S(\rho)\right). \tag{4.28}$$

By definition, we have

$$\mathfrak{N}(\zeta_{\bar{n}}) = \bar{n}. \tag{4.29}$$

Since the thermal state maximises the entropy for a fixed energy, we have that

$$\mathfrak{N}(\rho) \leq \bar{n}, \tag{4.30}$$

for any quantum state ρ whose mean number of photons is $\mathrm{Tr}\left[\hat{n}\rho\right] = \bar{n}$.

4.2.2 The entropy photon-number inequality: a conjecture

The entropy power inequality is defined in the context of a map acting as a convolution of two rescaled random variables. A natural analogy of the latter in the quantum realm consists in taking the beam-splitter unitary. Indeed, consider a bosonic mode characterised by an annihilation operator $\hat{a}$ that interacts with another bosonic mode of annihilation operator $\hat{b}$ through a beam splitter. The operations transforming the bosonic field operators correspond to the linear Bogoliubov transformations of Equation (3.55). If one discards the mode corresponding to $\hat{b}$ at the output of the unitary by tracing it out, one is left with the other mode having evolved as

$$\hat{c} = U_{\eta}^{\mathrm{BS}\dagger}\,\hat{a}\,U_{\eta}^{\mathrm{BS}} = \sqrt{\eta}\,\hat{a} + \sqrt{1-\eta}\,\hat{b}. \tag{4.31}$$

This is similar to the linear map acting on the random variables of the previous section by rescaling them before taking their sum. It then makes sense to consider the corresponding density matrices in state space and investigate the way the von Neumann entropies of the states evolve. Since the entropies are directly connected to the entropy photon-numbers, one can also choose to study the latter. Guha conjectured the following so-called entropy photon-number inequality (EPnI) [7, 90].

Conjecture 1 (Entropy photon-number inequality [90])**.** *Let ρ_a and ρ_b be two density matrices. Let $\rho_c = \mathrm{Tr}_2\left[U_{\eta}^{\mathrm{BS}}\left(\rho_a \otimes \rho_b\right) U_{\eta}^{\mathrm{BS}\dagger}\right]$, where U_{η}^{BS} is a beam-splitting unitary of transmittance η and Tr_2 corresponds to the partial trace over the second mode. Then*

$$\mathfrak{N}(\rho_c) \geq \eta\mathfrak{N}(\rho_a) + (1-\eta)\mathfrak{N}(\rho_b). \tag{4.32}$$

The situation described in Conjecture 1 is depicted in Figure 4.2.1. As a matter of fact, if ρ_a

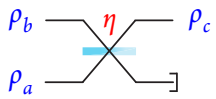

Figure 4.2.1: Set-up considered for the EPnI. The states ρ_a and ρ_b evolve in a beam splitter of transmittance η. The second mode is then discarded.

and ρ_b are both thermal Gaussian states, the entropy photon-number inequality of Equation (4.32) is saturated, *i.e.*,

$$\mathfrak{N}(\zeta_{\bar{n}_c}) = \eta\mathfrak{N}(\zeta_{\bar{n}_a}) + (1-\eta)\mathfrak{N}(\zeta_{\bar{n}_b}), \tag{4.33}$$

where

$$\zeta_{\bar{n}_c} = \mathrm{Tr}_2\left[U_{\eta}^{\mathrm{BS}}\left(\zeta_{\bar{n}_a} \otimes \zeta_{\bar{n}_b}\right) U_{\eta}^{\mathrm{BS}\dagger}\right]. \tag{4.34}$$

This is actually simply a consequence of the fact that $\zeta_{\bar{n}_c}$ is indeed thermal Gaussian (which is why we denoted it as such), and of the way energy is distributed in a beam splitter. Indeed, Equation (4.33) is exactly

$$\bar{n}_c = \eta \bar{n}_a + (1-\eta)\bar{n}_b. \tag{4.35}$$

The entropy photon-number inequality can be shown to be equivalent to the statement [7]

$$S(\rho_c) \geq S\left(\zeta_{\bar{n}_c}\right). \tag{4.36}$$

In analogy with the classical case, Guha conjectured the following third form of the entropy photon-number inequality [7],

$$S(\rho_c) \geq \eta S(\rho_a) + (1-\eta)S(\rho_b). \tag{4.37}$$

As it happens, Equation (4.37) is implied by Equation (4.32), as a consequence of the concavity of g. However, it seems to be weaker, unlike in the classical case. Equation (4.37) was actually proven to be true in [91]. Similarly to what we did in the case of the entropy power inequality, we summarise the different forms of the entropy photon-number inequality in Table 4.2.1. Note that all these forms can readily be generalised to an arbitrary number of mode entering an interferometer. Furthermore, for convenience, we define the operation of the beam splitter and partial trace as

$$\Phi_\eta\left[\rho_a, \rho_b\right] = \mathrm{Tr}_2\left[U_\eta^{\mathrm{BS}}\left(\rho_a \otimes \rho_b\right) U_\eta^{\mathrm{BS}\dagger}\right]. \tag{4.38}$$

Eq.	Statement	Conditions
(4.32)	$\mathfrak{N}\left(\Phi_\eta\left[\rho_a, \rho_b\right]\right) \geq \eta\mathfrak{N}(\rho_a) + (1-\eta)\mathfrak{N}(\rho_b)$	\
	$\Updownarrow$	
(4.36)	$S\left(\Phi_\eta\left[\rho_a, \rho_b\right]\right) \geq S\left(\Phi_\eta\left[\zeta_{\bar{n}_a}, \zeta_{\bar{n}_b}\right]\right)$	$S\left(\rho_a\right) = S\left(\zeta_{\bar{n}_a}\right)$ $S\left(\rho_b\right) = S\left(\zeta_{\bar{n}_b}\right)$
	$\Downarrow$	
(4.37)	$S\left(\Phi_\eta\left[\rho_a, \rho_b\right]\right) \geq \eta S(\rho_a) + (1-\eta)(\rho_b)$	\

Table 4.2.1: Equivalent forms of the entropy photon-number inequality. Note that $\Phi_\eta\left[\rho_a, \rho_b\right] = \mathrm{Tr}_2\left[U_\eta^{\mathrm{BS}}\left(\rho_a \otimes \rho_b\right) U_\eta^{\mathrm{BS}\dagger}\right]$, with $\eta \in [0,1]$ in the three cases.

The conjecture can be simplified by considering a fixed thermal Gaussian state $\rho_b = \zeta_{\bar{n}_b}$ on the second mode, as shown in Figure 4.2.2. In this case, the map Φ_η can be seen as a Gaussian

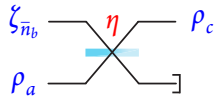

Figure 4.2.2: Set-up considered for the EPnI with a thermal Gaussian environment. The state ρ_a evolves in a lossy channel $\mathcal{B}_\eta^{(\varepsilon_b)}[\rho_a]$.

channel acting on the state ρ_a of the first mode. It is actually a lossy channel (defined in Equation (3.96)), *i.e.*,

$$\Phi_\eta\left[\rho_a, \zeta_{\bar{n}_b}\right] = \mathcal{B}_\eta^{(\varepsilon_b)}[\rho_a], \tag{4.39}$$

with $\bar{n}_b = \varepsilon_b/(1-\varepsilon_b)$. In this case, one needs to show that

$$S\left(\mathcal{B}_\eta^{(\varepsilon_b)}[\rho_a]\right) \geq S\left(\mathcal{B}_\eta^{(\varepsilon_b)}[\zeta_{\bar{n}_a}]\right), \tag{4.40}$$

with $S\left(\rho_a\right) = S\left(\zeta_{\bar{n}_a}\right)$, which is basically the simplified version of (4.36) (second row of Table 4.2.1). Equation (4.40) was first proven to be true in [78] for pure states ρ_a, meaning that one fixes $S(\rho_a) = 0$, so that $\zeta_{\bar{n}_a}$ is the vacuum state. It was later generalised in [92] for any fixed entropy of ρ_a.

Let us conclude by simply adding that all the entropy photon-number inequalities can be stated in the same way by considering a two-mode squeezer U_λ^{TMS} instead of the beam splitter U_η^{BS} of the map Φ_η.

II

Gaussian bosonic unitaries

A generating function is a device somewhat similar to a bag. Instead of carrying many little objects detachedly, which could be embarrassing, we put them all in a bag, and then we have only one object to carry, the bag.

George Pólya, Mathematics and plausible reasoning (1954) [93].

5

The generating function for Gaussian unitaries

The power of the symplectic formalism in phase space resides in the fact that purely Gaussian systems can always be represented using very few parameters. Indeed, we showed in Chapter 3 that Gaussian states are completely defined by their first two statistical moments, while Gaussian transformations need only few matrices for their characterisation. Consequently, the evolution of Gaussian states through Gaussian operations turns out rather easy to investigate. The story becomes different however, once either one of the state or the transformation is chosen to be non-Gaussian, in which case the symplectic formalism becomes powerless. The need to go beyond the framework of Gaussian systems is nevertheless imperative, since many continuous-variables quantum information processing tasks rely on non-Gaussian resources. This is reminiscent of situations where a Gaussian no-go theorem precludes the use of Gaussian resources in order to achieve a task involving Gaussian states, such as quantum entanglement distillation [8–10], quantum error correction [94], and quantum bit commitment [95].

In this chapter, we introduce a technique whose purpose is to circumvent such a difficulty. By exploiting the notion of generating function, we are able to rely on the knowledge obtained through the symplectic formalism in order to deal with the study of non-Gaussian states evolving through Gaussian unitaries. As we are going to show, the method we present will prove especially useful when including Fock states and Fock-passive states in the picture. These quantum states are of particular importance in the study of bosonic systems, since Fock states constitute the eigenbasis of the Hamiltonian of the system, while Fock-passive states play a determining role in the study of concepts like energy and work extraction.

In Section 5.1, we begin by introducing the mathematical notion of generating function, as well as its most important properties. Section 5.2 is dedicated to the calculation of some generating functions in the framework of bosonic quantum systems involving two modes, namely the generating functions of transition amplitudes and probabilities in a beam splitter and a two-mode squeezer. In the process, we point out some rather interesting facts implied by these calculations, such as the notion of partial time reversal connecting a beam splitter and a two-mode squeezer, and show how the generating functions can be used to study the asymptotic behaviour of transition probabilities for instance. Finally, we generalise the notions we deem the most interesting, by considering systems involving N modes and a passive Gaussian transformation in Section 5.3.

5.1 The generating function

5.1.1 Definition of the generating function

The generating function of an infinite sequence characterised by a discrete index provides a mean to equivalently describe the sequence using a function of a continuous parameter. There are different types of generating functions, one of them being the so-called ordinary generating function. Since we are only interested in the latter in this work, we will simply be calling it generating function (GF) thereafter. It is defined as follows [96].

Definition 30 (Generating function)**.** *Let $\{c_n\}_{n\geq 0}$ be a sequence. The generating function for the sequence is defined as*

$$g(z) := \mathcal{T}_n\left[c_n\right](z) = \sum_{n=0}^{\infty} c_n z^n, \tag{5.1}$$

where z is a complex number.

The generating function can be viewed in two different manners. On one hand, it can be seen as a formal power series in a complex number. On the other hand, it can be seen as a function of the complex variable. The last point of view often comes into play when the power series can be written in a nice closed form. The generating function (5.1) is a powerful tool as it encapsulates all information about the sequence $\{c_n\}$. It can also be defined for a sequence involving more than one index. For instance, we will write

$$g(z,w) := \mathcal{T}_{n,m}\left[c_n\right](z,w) = \sum_{n,m=0}^{\infty} c_{n,m} z^n w^m, \tag{5.2}$$

for the generating function of a sequence involving two indices n and m. In (5.2), the index n is associated with the variable z, while m is associated with w. This is represented by the fact that the subscripts n and m follow the same order as the arguments z and w in $\mathcal{T}_{n,m}\left[c_n\right](z,w)$. In

general, the generating function g of a sequence $c_{\mathbf{n}}$ involving indices arranged in a vector $\mathbf{n} \in \mathbb{N}_{\circ}^{N}$ will be written as

$$g(\mathbf{z}) := \mathcal{T}_{\mathbf{n}}\left[c_{\mathbf{n}}\right](\mathbf{z}) = \sum_{\mathbf{n}\in\mathbb{N}_{\circ}^{N}} c_{\mathbf{n}} \prod_{r=1}^{N} z_r^{n_r}. \tag{5.3}$$

Since the generating function involves power series, one may wonder about convergence issues. The following lemma addresses such a question [96].

Lemma 5. *Given a power series*

$$\sum_{n=\circ}^{\infty} c_n z^n \tag{5.4}$$

in a complex variable z, there exists an extended real number $\circ \leq R \leq \infty$ such that

- *if $|z| < R$, the series converges,*
- *if $|z| > R$, the series diverges.*

The quantity R is called the radius of convergence of the power series, and can be expressed as

$$R = \frac{1}{\limsup_{n\to\infty} |c_n|^{1/n}}. \tag{5.5}$$

It determines properties of the power series in the following way.

Lemma 6. *If g is a power series defined as*

$$g(z) = \sum_{n=\circ}^{\infty} c_n z^n \tag{5.6}$$

with a radius of convergence R, then $g(z)$ is an analytic function on the disk $|z| < R$ and has a least one singularity on the circle $|z| = R$ (the region of convergence).

5.1.2 Properties of the generating function

Since the generating function contains all the information on the corresponding sequence, many operations applied on the former can be "translated" to transformations on the latter. We list some of them hereafter [96].

Property 23 (Convolution yields multiplication)**.** *Consider two sequences $\{c_n\}$ and $\{d_n\}$. We have*

$$\mathcal{T}_n\left[\{c_\bullet * d_\bullet\}_n\right](z) = \mathcal{T}_n\left[c_n\right](z)\ \mathcal{T}_n\left[d_n\right](z)\,. \tag{5.7}$$

In the last equation, the convolution between two sequences $\{c_n\}$ and d_n is defined as

$$\{c_\bullet * d_\bullet\}_n = \sum_{m=\circ}^{n} c_m d_{n-m}. \tag{5.8}$$

Property 24 (Shifting property)**.** *Consider a sequence* $\{c_n\}$*. We have*

$$\mathcal{T}_n\left[c_{n+1}\right](z) = \frac{1}{z}\left(\mathcal{T}_n\left[c_n\right](z) - c_0\right). \tag{5.9}$$

This can be easily proven, as

$$\mathcal{T}_n\left[c_{n+1}\right](z) = \sum_{n=0}^{\infty} c_{n+1}z^n = \frac{1}{z}\left(\sum_{n=1}^{\infty} c_n z^n\right) = \frac{1}{z}\left(\sum_{n=0}^{\infty} c_n z^n - a_0\right) = \frac{1}{z}\left(\mathcal{T}_n\left[c_n\right](z) - c_0\right). \tag{5.10}$$

Property 25 (Multiplication of sequences)**.** *Consider two sequences* $\{c_n\}$ *and* $\{d_n\}$*. We have*

$$\mathcal{T}_n\left[c_n d_n\right](z) = \frac{1}{2\pi i}\oint_C d\tilde{z}\,\frac{1}{\tilde{z}}\mathcal{T}_n\left[c_n\right](\tilde{z})\,\mathcal{T}_n\left[d_n\right]\left(\frac{z}{\tilde{z}}\right), \tag{5.11}$$

where i *represents the imaginary unit, and* C *is a counter-clockwise closed path encircling the origin and entirely in the region of convergence corresponding to the product* $c_n d_n$*.*

Another interesting feature of the generating function lies in the fact that the asymptotic behaviour of a sequence $\{c_n\}$ for a growing index can be studied by analysing the asymptotic behaviour of the corresponding generating function $g(z)$ around its singularities. This is encompassed in the Tauberian theorems [97], the most famous of which being due to Hardy, Littlewood and Karamata [98].

Theorem 26 (HLK Tauberian theorem)**.** *Let* $g(z)$ *be a power series with radius of convergence equal to* 1*, satisfying*

$$g(z) \sim \frac{1}{(1-z)^a}\Lambda\left(\frac{1}{1-z}\right), \quad z \to 1, \tag{5.12}$$

for some $a \geq 0$ *with* Λ *a slowly varying function. Assume that the coefficients* $c_n = [z^n]g(z)$ *are all non-negative. Then*

$$\sum_{k=0}^{n} c_k \sim \frac{n^a}{\Gamma(a+1)}\Lambda(n), \quad n \to \infty. \tag{5.13}$$

A function Λ is said to be slowly varying at infinity if and only if, for any $\beta > 0$, one has

$$\frac{\Lambda(\beta x)}{\Lambda(x)} \to 1 \quad \text{as} \quad x \to +\infty. \tag{5.14}$$

Also, the notation $c_n = [z^n]g(z)$ means that we take the coefficient of the z^n term in $g(z) = \sum_{n=0}^{\infty} c_n z^n$.

5.2 Generating functions for two-mode Gaussian unitaries

5.2.1 Generating function for the modified transition amplitudes

The purpose of this section is to compute generating functions associated with the amplitudes involving two-mode Gaussian unitaries and Fock states. Since our goal is to exploit the symplectic formalism of Gaussian systems, we choose to calculate the 4-variate generating functions of the modified amplitudes

$$\frac{\langle n,m|\, U_\eta^{\mathrm{BS}}\, |i,k\rangle}{\sqrt{i!k!n!m!}}, \qquad \frac{\langle n,m|\, U_\lambda^{\mathrm{TMS}}\, |i,k\rangle}{\sqrt{i!k!n!m!}}, \tag{5.15}$$

where $|i\rangle$, $|k\rangle$, $|n\rangle$, and $|m\rangle$ denote Fock states, instead of the amplitudes themselves, as coherent states then naturally appear during the derivations of such generating functions. Indeed, consider the 4-dimensional sequence

$$\frac{\langle n,m|\, U\, |i,k\rangle}{\sqrt{i!k!n!m!}} \tag{5.16}$$

for some unitary U. Its 4-variate generating function can easily be written as

$$g(x,y,z,w) = \sum_{i,k,n,m} \frac{\langle n,m|\, U\, |i,k\rangle}{\sqrt{i!k!n!m!}} x^i y^k z^n w^m,$$

or,

$$g(x,y,z,w) = \left(\sum_n \frac{z^n}{\sqrt{n!}} \langle n|\right) \otimes \left(\sum_m \frac{w^m}{\sqrt{m!}} \langle m|\right) U \left(\sum_i \frac{x^i}{\sqrt{i!}} |i\rangle\right) \otimes \left(\sum_k \frac{y^k}{\sqrt{k!}} |k\rangle\right),$$

so that

$$g(x,y,z,w) = \mathrm{e}^{\frac{|x|^2+|y|^2+|z|^2+|w|^2}{2}} \langle z,w|\, U\, |x,y\rangle\,, \tag{5.17}$$

where $|x\rangle$, $|y\rangle$, $|z\rangle$, and $|w\rangle$ are coherent states defined in Equation (3.41), with the conventions shown in Figure 5.2.1. Namely, we obtain a matrix element of U between coherent states. This makes the generating function g very easy to compute when U is Gaussian, regardless of the complexity of $\langle n,m|\, U\, |i,k\rangle$.

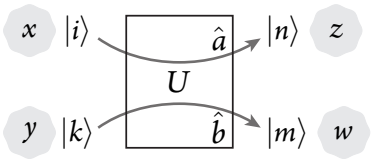

Figure 5.2.1: Conventions in the definition of $g(x,y,z,w)$.

5.2.1.1 Beam splitter

Define the generating function g_η^{BS} of the modified amplitudes involving a beam-splitter unitary as

$$g_\eta^{\mathrm{BS}}(x,y,z,w) = \sum_{i,k,n,m} \frac{\langle n,m|\, U_\eta^{\mathrm{BS}}\, |i,k\rangle}{\sqrt{i!k!n!m!}} x^i y^k z^n w^m, \tag{5.18}$$

choosing $x, y, z, w \in \mathbb{R}$. It can be rewritten in terms of coherent states in the form of Equation (5.17). In order to apply the beam splitter to the product of coherent states, we take advantage of the fact that the beam splitter does not modify a couple of vacua, and

$$U_\eta^{\mathrm{BS}}\, |x,y\rangle = U_\eta^{\mathrm{BS}} \left(D_x \otimes D_y\right) |0,0\rangle = U_\eta^{\mathrm{BS}} \left(D_x \otimes D_y\right) U_\eta^{\mathrm{BS}\dagger}\, |0,0\rangle\,, \tag{5.19}$$

where the displacement operator is defined in Equation (3.36). Now,

$$U_\eta^{\mathrm{BS}} \left(D_x \otimes D_y\right) U_\eta^{\mathrm{BS}\dagger} = U_\eta^{\mathrm{BS}} \left(\mathrm{e}^{x\hat{a}^\dagger - x^* a + y\hat{b}^\dagger - y^* b}\right) U_\eta^{\mathrm{BS}\dagger} = \mathrm{e}^{U_\eta^{\mathrm{BS}}(x\hat{a}^\dagger - x^* a + y\hat{b}^\dagger - y^* b)U_\eta^{\mathrm{BS}\dagger}}. \tag{5.20}$$

Exploiting the action of the beam splitter in the Heisenberg picture defined in Equations (3.55), we get

$$\begin{aligned} U_\eta^{\mathrm{BS}}(x\hat{a}^\dagger - x^* a + y\hat{b}^\dagger - y^* b)U_\eta^{\mathrm{BS}\dagger} = {} & (\sqrt{\eta}x + \sqrt{1-\eta}y)\hat{a}^\dagger - (\sqrt{\eta}x + \sqrt{1-\eta}y)^* a \\ & + (\sqrt{\eta}y - \sqrt{1-\eta}x)\hat{b}^\dagger - (\sqrt{\eta}y - \sqrt{1-\eta}x)^* b, \end{aligned}$$

meaning that

$$U_\eta^{\mathrm{BS}} \left(D_x \otimes D_y\right) U_\eta^{\mathrm{BS}\dagger} = \mathrm{e}^{(\sqrt{\eta}x + \sqrt{1-\eta}y)\hat{a}^\dagger - (\sqrt{\eta}x + \sqrt{1-\eta}y)^* a} \otimes \mathrm{e}^{(\sqrt{\eta}y - \sqrt{1-\eta}x)\hat{b}^\dagger - (\sqrt{\eta}y - \sqrt{1-\eta}x)^* b}, \tag{5.21}$$

or, using the definition of the displacement again,

$$U_\eta^{\mathrm{BS}} \left(D_x \otimes D_y\right) U_\eta^{\mathrm{BS}\dagger} = D_{\sqrt{\eta}x + \sqrt{1-\eta}y} \otimes D_{\sqrt{\eta}y - \sqrt{1-\eta}x}. \tag{5.22}$$

Using this information, we end up with

$$\langle z,w|\, U_\eta^{\mathrm{BS}}\, |x,y\rangle = \left\langle z,w\right|\, \sqrt{\eta}x + \sqrt{1-\eta}y, \sqrt{\eta}y - \sqrt{1-\eta}x\right\rangle. \tag{5.23}$$

Since we have $x, y, z, w \in \mathbb{R}$,

$$\langle z,w|\, U_\eta^{\mathrm{BS}}\, |x,y\rangle = \mathrm{e}^{-\frac{|\sqrt{\eta}x + \sqrt{1-\eta}y|^2 + |z|^2}{2}} \mathrm{e}^{-\frac{|\sqrt{\eta}y - \sqrt{1-\eta}x|^2 + |w|^2}{2}} \mathrm{e}^{z(\sqrt{\eta}x + \sqrt{1-\eta}y)} \mathrm{e}^{w(\sqrt{\eta}y - \sqrt{1-\eta}x)}, \tag{5.24}$$

so that

$$g_\eta^{\mathrm{BS}}(x,y,z,w) = \mathrm{e}^{\sqrt{\eta}(xz+yw) + \sqrt{1-\eta}(yz-xw)}. \tag{5.25}$$

CONSISTENCY CHECK: CONSERVATION OF ENERGY It is interesting to notice that the conservation of energy in the beam splitter can be easily verified using the generating function computed above. Define the object

$$\tilde{g}_\eta^{\mathrm{BS}}(x,y,z,w,t) = \sum_{i,k,n,m} \frac{\langle n,m|\, U_\eta^{\mathrm{BS}}\, |i,k\rangle}{\sqrt{i!k!n!m!}} x^i y^k z^n w^m t^{i+k-n-m}, \tag{5.26}$$

where we chose to add a variable t. In this case,

$$\begin{aligned}\tilde{g}_\eta^{\mathrm{BS}}(x,y,z,w,t) &= \sum_{i,k,n,m} \frac{\langle n,m|\, U_\eta^{\mathrm{BS}}\, |i,k\rangle}{\sqrt{i!k!n!m!}} (xt)^i (yt)^k \left(\frac{z}{t}\right)^n \left(\frac{w}{t}\right)^m \\ &= g_\eta^{\mathrm{BS}}(xt, yt, \frac{z}{t}, \frac{w}{t}).\end{aligned}$$

Now, using the definition of g_η^{BS}, we end up with

$$\tilde{g}_\eta^{\mathrm{BS}}(x,y,z,w,t) = g_\eta^{\mathrm{BS}}(x,y,z,w), \quad \forall t. \tag{5.27}$$

This actually means that $\tilde{g}_\eta^{\mathrm{BS}}$ as defined in Equation (5.26) does not depend on the variable t, or that the only non-zero elements in the sums of the right-hand side of Equation (5.26) verify $i+k-n-m=0$. Consequently,

$$\langle n,m|\, U_\eta^{\mathrm{BS}}\, |i,k\rangle = 0 \quad \text{if} \quad i+k \neq n+m. \tag{5.28}$$

From now on, we will define the transition amplitudes of a beam splitter with transmittance η as

$$b_n^{(i,k)} = \langle n,m|\, U_\eta^{\mathrm{BS}}\, |i,k\rangle\,, \tag{5.29}$$

noting that the index $m = i+k-n$ is redundant.

5.2.1.2 TWO-MODE SQUEEZER

Define the generating function g_λ^{TMS} of the modified amplitudes in a two-mode squeezer as

$$g_\lambda^{\mathrm{TMS}}(x,y,z,w) = \sum_{i,k,n,m} \frac{\langle n,m|\, U_\lambda^{\mathrm{TMS}}\, |i,k\rangle}{\sqrt{i!k!n!m!}} x^i y^k z^n w^m, \tag{5.30}$$

with $x,y,z,w \in \mathbb{R}$. Again, it can be rewritten using coherent states in the form of Equation (5.17). Using the same techniques as in the case of the beam splitter, we have

$$U_\lambda^{\mathrm{TMS}}\, |x,y\rangle = U_\lambda^{\mathrm{TMS}} \left(D_x \otimes D_y\right) |0,0\rangle = U_\lambda^{\mathrm{TMS}} \left(D_x \otimes D_y\right) U_\lambda^{\mathrm{TMS}\dagger} U_\lambda^{\mathrm{TMS}}\, |0,0\rangle\,, \tag{5.31}$$

where

$$U_\lambda^{\mathrm{TMS}} \left(D_x \otimes D_y\right) U_\lambda^{\mathrm{TMS}\dagger} = \mathrm{e}^{U_\lambda^{\mathrm{TMS}}(x\hat{a}^\dagger - x^* a + y\hat{b}^\dagger - y^* b)U_\lambda^{\mathrm{TMS}\dagger}}. \tag{5.32}$$

Again, from Equations (3.57),

$$\begin{aligned} U_\lambda^{\mathrm{TMS}}(x\hat{a}^\dagger - x^* a + y\hat{b}^\dagger - y^* b)U_\lambda^{\mathrm{TMS}\dagger} &= (x\cosh(r) + y^*\sinh(r))\hat{a}^\dagger \\ &\quad - (x\cosh(r) + y^*\sinh(r))^* a \\ &\quad + (y\cosh(r) + x^*\sinh(r))\hat{b}^\dagger \\ &\quad - (y\cosh(r) + x^*\sinh(r))^* b, \end{aligned}$$

where $\lambda = \tanh^2(r)$. This leads to

$$U_\lambda^{\mathrm{TMS}} \left(D_x \otimes D_y\right) U_\lambda^{\mathrm{TMS}\dagger} = D_{x\cosh(r)+y^*\sinh(r)} \otimes D_{y\cosh(r)+x^*\sinh(r)}, \tag{5.33}$$

$$U_\lambda^{\mathrm{TMS}} |x, y\rangle = D_{x\cosh(r)+y^*\sinh(r)} \otimes D_{y\cosh(r)+x^*\sinh(r)} |\varphi_\lambda^{\mathrm{EPR}}\rangle, \tag{5.34}$$

where $|\varphi_\lambda^{\mathrm{EPR}}\rangle = U_\lambda^{\mathrm{TMS}} |0, 0\rangle$ is a Gaussian two-mode squeezed state. We then have

$$\langle z, w| U_\lambda^{\mathrm{TMS}} |x, y\rangle = \langle z, w| D_{x\cosh(r)+y^*\sinh(r)} \otimes D_{y\cosh(r)+x^*\sinh(r)} |\varphi_\lambda^{\mathrm{EPR}}\rangle. \tag{5.35}$$

Since all our coherent states, and therefore displacements, are characterised by real parameters, we have that $D_\alpha D_\beta = D_{\alpha+\beta}$, meaning that

$$\begin{aligned} \langle z, w| U_\lambda^{\mathrm{TMS}} |x, y\rangle &= \langle 0, 0| D_{x\cosh(r)+y^*\sinh(r)-z} \otimes D_{y\cosh(r)+x^*\sinh(r)-w} |\varphi_\lambda^{\mathrm{EPR}}\rangle \\ &= \left\langle z - x\cosh(r) - y^*\sinh(r), w - y\cosh(r) - x^*\sinh(r)\right| \varphi_\lambda^{\mathrm{EPR}}\rangle. \end{aligned}$$

The overlap between a coherent state and a Fock state is such that [99]

$$\langle \alpha| n\rangle = \mathrm{e}^{-\frac{\alpha^2}{2}} \frac{\alpha^n}{\sqrt{n!}}, \quad \alpha \in \mathbb{R} \tag{5.36}$$

so that

$$\begin{aligned} \left\langle \alpha, \beta\right| \varphi_\lambda^{\mathrm{EPR}}\rangle &= \sqrt{1-\lambda}\,\mathrm{e}^{-\frac{\alpha^2+\beta^2}{2}} \sum_{n=0}^{\infty} \frac{(\sqrt{\lambda}\alpha\beta)^n}{n!}, \quad \alpha, \beta \in \mathbb{R}, \\ &= \sqrt{1-\lambda}\,\mathrm{e}^{-\frac{\alpha^2+\beta^2}{2}} \mathrm{e}^{\sqrt{\lambda}\alpha\beta}, \quad \alpha, \beta \in \mathbb{R}. \end{aligned}$$

If one chooses

$$\begin{cases} \alpha = z - x\cosh(r) - y\sinh(r), \\ \beta = w - y\cosh(r) - x\sinh(r), \end{cases} \tag{5.37}$$

then

$$\begin{aligned} \alpha^2 + \beta^2 &= z^2 + (x\cosh(r) + y\sinh(r))^2 - 2z(x\cosh(r) + y\sinh(r)) \\ &\quad + w^2 + (y\cosh(r) + x\sinh(r))^2 - 2w(y\cosh(r) + x\sinh(r)), \end{aligned}$$

leading to

$$\ln\left(\frac{g_\lambda^{\mathrm{TMS}}(\mathbf{v})}{\sqrt{1-\lambda}}\right) = \frac{|x|^2+|y|^2+|z|^2+|w|^2}{2} - \frac{\alpha^2+\beta^2}{2} + \sqrt{\lambda}\alpha\beta, \tag{5.38}$$

where we set $\mathbf{v} = (x, y, z, w) \in \mathbb{R}^4$. After some calculations, one gets

$$\ln\left(\frac{g_\lambda^{\mathrm{TMS}}(\mathbf{v})}{\sqrt{1-\lambda}}\right) = \mathrm{sech}(r)(xz+yw) + \tanh(r)(zw-xy), \tag{5.39}$$

and consequently,

$$g_\lambda^{\mathrm{TMS}}(\mathbf{v}) = \sqrt{1-\lambda}\mathrm{e}^{\mathrm{sech}(r)(xz+yw)+\tanh(r)(zw-xy)}. \tag{5.40}$$

Finally, if we use the fact that $\lambda = \tanh^2(r)$, we get

$$g_\lambda^{\mathrm{TMS}}(\mathbf{v}) = \sqrt{1-\lambda}\mathrm{e}^{\sqrt{1-\lambda}(xz+yw)+\sqrt{\lambda}(zw-xy)}. \tag{5.41}$$

Consistency check: conservation of photon-number difference In the case of the two-mode squeezer, the conservation of the photon-number difference is reflected by

$$g_\lambda^{\mathrm{TMS}}(x, y, z, w) = g_\lambda^{\mathrm{TMS}}\left(tx, \frac{y}{t}, \frac{z}{t}, w\right), \quad \forall t. \tag{5.42}$$

From now on, we will define the transition amplitudes of a two-mode squeezer with parameter $\lambda = \tanh^2(r)$ as

$$a_n^{(i,k)} = \langle n, m|\, U_\lambda^{\mathrm{TMS}}\, |i, k\rangle\,, \tag{5.43}$$

noting that the index $m = n + k - i$ is redundant.

5.2.2 Partial time reversal

By comparing the above generating functions, it appears that the two-mode squeezer may be viewed as a beam splitter undergoing *partial time reversal* [100]. By interchanging variables y and w (which may be interpreted as reverting the arrow of time for mode $\hat{b}$) and taking $\eta = 1 - \lambda$, we see that the generating functions are connected by

$$g_\lambda^{\mathrm{TMS}}(x, y, z, w) = \sqrt{1-\lambda}\, g_{1-\lambda}^{\mathrm{BS}}(x, w, z, y). \tag{5.44}$$

This means that exchanging the roles of k and m in the matrix elements converts a two-mode squeezer into a beam splitter, namely

$$\langle n, m|\, U_\lambda^{\mathrm{TMS}}\, |i, k\rangle = \sqrt{1-\lambda}\ \langle n, k|\, U_{1-\lambda}^{\mathrm{BS}}\, |i, m\rangle\,. \tag{5.45}$$

Equation (5.44) can be rewritten as

$$g_\lambda^{\mathrm{TMS}}(x,y,z,w) = \frac{1}{\sqrt{G}}\, g_{1/G}^{\mathrm{BS}}(x,w,z,y), \tag{5.46}$$

where the gain G of the two-mode squeezer is defined such that $G = 1/(1-\lambda)$. This is reminiscent of the fact that the dual map of the quantum limited amplifier $\mathcal{A}_G$ is proportional to the pure-loss channel $\mathcal{B}_\eta$, where the transmittance of the latter is given by the inverse of the gain of the former, and the coefficient of proportionality is equal to the inverse of the gain. In other words,

$$\mathcal{A}_G^\dagger[\bullet] = \frac{1}{G}\mathcal{B}_{1/G}[\bullet]. \tag{5.47}$$

Indeed, in Equation (5.46), the two generating functions are proportional, and the transmittance η of g_η^{BS} is replaced by the inverse of the gain G.

5.2.3 Computing the transition amplitudes

We now exploit the generating functions in order to compute the transition amplitudes. An obvious, albeit interesting thing to notice is that

$$g_\eta^{\mathrm{BS}}(\mathbf{v}) = g_\eta^{\mathrm{BS}}(\mathbf{v})\Big|_{y=0}\, g_\eta^{\mathrm{BS}}(\mathbf{v})\Big|_{x=0}. \tag{5.48}$$

Since a product of generating functions corresponds to a convolution of their respective sequences, we get

$$b_n^{(i,k)} = \sum_{\tilde{n}=\max(0,n-k)}^{\min(n,i)} \sqrt{\binom{n}{\tilde{n}}\binom{i+k-n}{i-\tilde{n}}}\, b_{\tilde{n}}^{(i,0)} b_{n-\tilde{n}}^{(0,k)}, \tag{5.49}$$

for the transition amplitudes in a beam splitter. Similarly, one can express the transition amplitudes in a two-mode squeezer as

$$a_n^{(i,k)} = \sum_{\tilde{\imath}=\max(0,i-k)}^{\min(i,n)} \sqrt{\binom{i}{\tilde{\imath}}\binom{n-i+k}{n-\tilde{\imath}}}\, \frac{a_n^{(\tilde{\imath},0)} a_0^{(i-\tilde{\imath},k)}}{\sqrt{1-\lambda}}. \tag{5.50}$$

Compared to the direct calculation of $b_n^{(i,k)}$ as illustrated in Appendix D, the amplitudes $b_n^{(i,k)}$ as well as $a_n^{(i,k)}$ can be easily derived from Eqs. (5.49) and (5.50), using the relations

$$b_n^{(i,0)} = \binom{i}{n}^{1/2} \eta^{n/2}(1-\eta)^{(i-n)/2} \tag{5.51}$$

and

$$a_n^{(i,0)} = \binom{n}{i}^{1/2} (1-\lambda)^{(1+i)/2}\lambda^{(n-i)/2}. \tag{5.52}$$

The method based on generating functions we used here to obtain an expression for $b_n^{(i,k)}$ is to be compared with the approach developed in Appendix D. One can argue that the former technique is neater, as it exploits the symplectic formalism applied to Gaussian systems. One could claim that computing the generating function also takes time in itself. However, the generating functions computed here are standard objects in quantum optics. Furthermore, they need only be computed once, and can be used as a tool for several derivations and proofs (*e.g.* calculation of amplitudes, probabilities, proof of relations they verify, conservation of energy, ...).

5.2.4 Generating function for the transition probabilities

We are also interested in the generating functions of the transition probabilities $|\langle n,m|\, U\, |i,k\rangle|^2$, involving two-mode Gaussian unitaries. For any unitary U, the generating function is defined as

$$f(x,y,z,w) = \sum_{i,k,n,m} \langle n,m|\, U\, |i,k\rangle \langle i,k|\, U^\dagger\, |n,m\rangle\, x^i y^k z^n w^m. \tag{5.53}$$

with the conventions shown in Figure 5.2.1 The terms in the last quantity can be rearranged as

$$f(x,y,z,w) = \mathrm{Tr}\left[U\left(\sum_i x^i |i\rangle\langle i| \otimes \sum_k y^k |k\rangle\langle k|\right) \times U^\dagger\left(\sum_n z^n |n\rangle\langle n| \otimes \sum_m w^m |m\rangle\langle m|\right)\right],$$

so that thermal states naturally appear, leading to

$$f(\mathbf{v}) = \frac{\mathrm{Tr}\left[(\tau_z \otimes \tau_w) U (\tau_x \otimes \tau_y) U^\dagger\right]}{(1-x)(1-y)(1-z)(1-w)}, \tag{5.54}$$

where $\mathbf{v} = (x,y,z,w) \in \mathbb{R}^4$, τ_t being a Gaussian thermal state of parameter t such that $0 \leq t < 1$. By means of the sympletic formalism of Gaussian systems, the calculation of f consequently becomes rather easy for Gaussian unitaries.

5.2.4.1 Beam splitter

The generating function of the probability $|\langle n,m|\, U_\eta^{\mathrm{BS}}\, |i,k\rangle|^2$ is given by

$$f_\eta^{\mathrm{BS}}(x,y,z,w) = \sum_{i,k,n,m} |\langle n,m|\, U_\eta^{\mathrm{BS}}\, |i,k\rangle|^2 x^i y^k z^n w^m. \tag{5.55}$$

It can be rewritten in terms of thermal Gaussian states in the form of Equation (5.54). Now, the object $U_\eta^{\mathrm{BS}}\left(\tau_x \otimes \tau_y\right) U_\eta^{\mathrm{BS}\dagger}$ represents the effect of a beam-splitter unitary on the cross product of two Gaussian thermal states. Consequently, it represents a two-mode Gaussian state, which

we label by ϱ_1, i.e,

$$\varrho_1 = U_\eta^{\mathrm{BS}} \left(\tau_x \otimes \tau_y\right) U_\eta^{\mathrm{BS}\dagger}. \tag{5.56}$$

The object $\tau_z \otimes \tau_w$ is obviously a two-mode Gaussian state as well. We label it by $\varrho_2 = \tau_z \otimes \tau_w$. This means that $f_\eta^{\mathrm{BS}}(x, y, z, w)$ is proportional to the overlap $\mathrm{Tr}\left[\varrho_1 \varrho_2\right]$ between the two Gaussian states,

$$f_\eta^{\mathrm{BS}}(x, y, z, w) = \frac{1}{(1-x)(1-y)(1-z)(1-w)} \mathrm{Tr}\left[\varrho_1 \varrho_2\right]. \tag{5.57}$$

Since the first moments of each of the two Gaussian states ϱ_1 and ϱ_2 is null, their overlap can be computed using the formula [101]

$$\mathrm{Tr}\left[\varrho_1 \varrho_2\right] = \left(\det\left[\frac{\mathbf{V}_1 + \mathbf{V}_2}{2}\right]\right)^{-\frac{1}{2}} = \frac{4}{\sqrt{\det\left[\mathbf{V}_1 + \mathbf{V}_2\right]}}, \tag{5.58}$$

where $\mathbf{V}_1$ and $\mathbf{V}_2$ are the respective covariance matrices of ϱ_1 and ϱ_2. Now, the covariance matrix of $\tau_x \otimes \tau_y$ is diagonal, and is equal to $\left[(2n_x+1)\mathbb{1}_2\right] \oplus \left[(2n_y+1)\mathbb{1}_2\right]$, where $\mathbb{1}_2$ is the 2 by 2 identity matrix, and $n_t = t/(1-t)$ is the mean number of photons of the one-mode Gaussian thermal state τ_t of parameter t. The effect of the beam splitter on the covariance matrix of the product $\tau_x \otimes \tau_y$ in phase space is characterised by the symplectic matrix

$$S_\eta = \begin{pmatrix} \sqrt{\eta}\mathbb{1}_2 & \sqrt{1-\eta}\mathbb{1}_2 \\ -\sqrt{1-\eta}\mathbb{1}_2 & \sqrt{\eta}\mathbb{1}_2 \end{pmatrix}, \tag{5.59}$$

so that

$$\mathbf{V}_1 = \begin{pmatrix} c_{11} & 0 & c_{13} & 0 \\ 0 & c_{11} & 0 & c_{13} \\ c_{13} & 0 & c_{33} & 0 \\ 0 & c_{13} & 0 & c_{33} \end{pmatrix}, \tag{5.60}$$

where we defined $c_{11} = \eta(2n_x+1) + (1-\eta)(2n_y+1)$, $c_{33} = \eta(2n_y+1) + (1-\eta)(2n_x+1)$, and $c_{13} = 2\sqrt{\eta(1-\eta)}(n_y - n_x)$. The covariance matrix $\mathbf{V}_2$ is diagonal, and can be written $\mathbf{V}_2 = \left[(2n_z+1)\mathbb{1}_2\right] \oplus \left[(2n_w+1)\mathbb{1}_2\right]$, which means that $\mathbf{V} = \mathbf{V}_1 + \mathbf{V}_2$ is also of the form

$$\mathbf{V} = \begin{pmatrix} c_{11} + (2n_z+1) & 0 & c_{13} & 0 \\ 0 & c_{11} + (2n_z+1) & 0 & c_{13} \\ c_{13} & 0 & c_{33} + (2n_w+1) & 0 \\ 0 & c_{13} & 0 & c_{33} + (2n_w+1) \end{pmatrix}. \tag{5.61}$$

After some calculation, we get

$$\mathrm{Tr}\left[\varrho_1 \varrho_2\right] = \frac{1}{(n_x + n_w + 1)(n_y + n_z + 1) + \eta(n_x - n_y)(n_w - n_z)}. \tag{5.62}$$

Finally, by using the relation between the mean number of photon of a Gaussian thermal state and its parameter, we obtain

$$f_\eta^{\mathrm{BS}}(x,y,z,w) = \frac{1}{1-\eta xz-(1-\eta)xw-\eta yw-(1-\eta)yz+xyzw}. \tag{5.63}$$

As a consistency check, we note that

$$f_\eta^{\mathrm{BS}}(0) = |\langle 0,0|\, U_\eta^{\mathrm{BS}}\, |0,0\rangle|^2 = 1, \tag{5.64}$$

while normalisation

$$\sum_{n,m=0}^{\infty} |\langle n,m|\, U\, |i,k\rangle|^2 = 1, \quad \forall i,k \tag{5.65}$$

for the beam splitter translates into

$$f^{\mathrm{BS}}(x,y,1,1) = \frac{1}{1-x}\frac{1}{1-y}. \tag{5.66}$$

From now on, we will define the transition probabilities of a beam splitter of transmittance η as

$$B_n^{(i,k)} = |\langle n,m|\, U_\eta^{\mathrm{BS}}\, |i,k\rangle|^2, \tag{5.67}$$

with $m = i+k-n$.

5.2.4.2 Two-mode squeezer

The generating function of the probability $|\langle n,m|\, U_\lambda^{\mathrm{TMS}}\, |i,k\rangle|^2$ is given by

$$f_\lambda^{\mathrm{TMS}}(x,y,z,w) = \sum_{i,k,n,m} |\langle n,m|\, U_\lambda^{\mathrm{TMS}}\, |i,k\rangle|^2 x^i y^k z^n w^m. \tag{5.68}$$

One way to proceed would be to compute it from scratch, following the same procedure as in the case of the beam splitter. However, there is a much simpler way by taking advantage of the property of partial time reversal described by Equation (5.44). Using Property 25, as well as the fact that the amplitudes we consider take real values, we have

$$\begin{aligned} f_\lambda^{\mathrm{TMS}}(x,y,z,w) &= \frac{1}{(2\pi i)^4}\oint_{C^4} d\tilde{\mathbf{v}}\, \frac{1}{\tilde{x}\tilde{y}\tilde{z}\tilde{w}} g_\lambda^{\mathrm{TMS}}(\tilde{x},\tilde{y},\tilde{z},\tilde{w}) g_\lambda^{\mathrm{TMS}}\left(\frac{x}{\tilde{x}},\frac{y}{\tilde{y}},\frac{z}{\tilde{z}},\frac{w}{\tilde{w}}\right) \\ &= \frac{1}{(2\pi i)^4}\oint_{C^4} d\tilde{\mathbf{v}}\, \frac{1}{\tilde{x}\tilde{y}\tilde{z}\tilde{w}} \sqrt{1-\lambda} g_{1-\lambda}^{\mathrm{BS}}(\tilde{x},\tilde{w},\tilde{z},\tilde{y}) \sqrt{1-\lambda} g_{1-\lambda}^{\mathrm{BS}}\left(\frac{x}{\tilde{x}},\frac{w}{\tilde{w}},\frac{z}{\tilde{z}},\frac{y}{\tilde{y}}\right) \\ &= (1-\lambda) f_{1-\lambda}^{\mathrm{BS}}(x,w,z,y). \end{aligned}$$

Consequently,

$$f_\lambda^{\mathrm{TMS}}(x,y,z,w) = \frac{1-\lambda}{1-(1-\lambda)xz-\lambda xy-(1-\lambda)yw-\lambda wz+xyzw}. \tag{5.69}$$

As a consistency check, we note that

$$f_\lambda^{\mathrm{TMS}}(0) = |\langle 0,0|\, U_\lambda^{\mathrm{TMS}}\, |0,0\rangle|^2 = 1-\lambda, \tag{5.70}$$

while normalisation for the two-mode squeezer translates into

$$f^{\mathrm{TMS}}(x,y,1,1) = \frac{1}{1-x}\frac{1}{1-y}. \tag{5.71}$$

From now on, we will define the transition probabilities of a two-mode squeezer with parameter $\lambda = \tanh^2(r)$ as

$$A_n^{(i,k)} = |\langle n,m|\, U_\lambda^{\mathrm{TMS}}\, |i,k\rangle|^2, \tag{5.72}$$

with $m = n+k-i$.

5.2.5 Asymptotic analysis of the transition probabilities

Let us analyse the behaviour of the probability $B_n^{(i,i)}$ for a balanced beam splitter, meaning that we fix $\eta = 1/2$. We first need to find the generating function of $B_n^{(i,i)} = |\langle n,m|\, U_\eta^{\mathrm{BS}}\, |i,i\rangle|^2$ (with $m = 2i-n$) in i and n. We defined the generating function of the probabilities in the beam splitter as

$$f_\eta^{\mathrm{BS}}(x,y,z,w) = \sum_{i,k,n,m} |\langle n,m|\, U_\eta^{\mathrm{BS}}\, |i,k\rangle|^2 x^i y^k z^n w^m. \tag{5.73}$$

The sum over m was taken for the sake of of symmetry. Since the probability $|\langle n,m|\, U_\eta^{\mathrm{BS}}\, |i,k\rangle|^2$ will be zero if $m \neq i+k-n$, one readily understands that

$$B_n^{(i,k)} = |\langle n,i+k-n|\, U_\eta^{\mathrm{BS}}\, |i,k\rangle|^2 = \sum_{m=0}^{\infty} |\langle n,m|\, U_\eta^{\mathrm{BS}}\, |i,k\rangle|^2. \tag{5.74}$$

It actually means that if one computes the generating function of the sequence $B_n^{(i,k)}$ without considering the index m, one gets

$$\sum_{i,k,n} B_n^{(i,k)} x^i y^k z^n = \sum_{i,k,n} \left(\sum_m |\langle n,m|\, U_\eta^{\mathrm{BS}}\, |i,k\rangle|^2 w^m \right)\Bigg|_{w=1} x^i y^k z^n, \tag{5.75}$$

so that

$$f_\eta^{\mathrm{BS}}(x,y,z,1) = \sum_{i,k,n} B_n^{(i,k)} x^i y^k z^n. \tag{5.76}$$

5.2.5.1 Generating function of $B_n^{(i,i)}$

In order to derive the generating function of the diagonal elements $B_n^{(i,i)}$, we force the relation $k = i$ in the generating function of $B_n^{(i,k)}$ by only choosing the elements which verify it, *i.e.*,

$$\begin{aligned}\mathcal{T}_{i,n}\left[B_n^{(i,i)}\right](x,z) &= [y^0]\sum_{i,k,n} B_n^{(i,k)} x^i y^{k-i} z^n \\ &= [y^0]\sum_{i,k,n} B_n^{(i,k)} \left(\frac{x}{y}\right)^i y^k z^n \\ &= [y^0] f_\eta^{\mathrm{BS}}\left(\frac{x}{y}, y, z, 1\right).\end{aligned}$$

By Cauchy's integral formula for any function $g(z)$, one has

$$g(a) = \frac{1}{2\pi i}\oint \frac{g(z)}{z-a}\mathrm{d}z. \tag{5.77}$$

Applying this to our case, we get that, for some circle γ_x around $y = 0$,

$$\mathcal{T}_{i,n}\left[B_n^{(i,i)}\right](x,z) = \frac{1}{2\pi i}\int_{\gamma_x} \frac{f_\eta^{\mathrm{BS}}(x/y, y, z, 1)}{y}\mathrm{d}y. \tag{5.78}$$

Now, using the Residue Theorem, this can be transformed to

$$\mathcal{T}_{i,n}\left[B_n^{(i,i)}\right](x,z) = \sum_l \mathrm{Res}\left[\frac{f_\eta^{\mathrm{BS}}(x/y, y, z, 1)}{y}; y = s_l\right], \tag{5.79}$$

where the s_l are the singularities of $f_\eta^{\mathrm{BS}}(x/y, y, z, 1)/y$ satisfying

$$\lim_{x\to 0} s_l(x) = 0. \tag{5.80}$$

The singularities of $f_\eta^{\mathrm{BS}}(x/y, y, z, 1)/y$ can be computed to be

$$s_\mp(x,z,\eta) = \frac{1 + xz \mp \sqrt{(1+xz)^2 - 4(\eta + (1-\eta)z)(x(1-\eta) + \eta xz))}}{2(\eta + z(1-\eta))}. \tag{5.81}$$

If we take their limits for x going to zero, we obtain

$$\lim_{x\to 0} s_-(x,z,\eta) = 0, \quad \lim_{x\to 0} s_+(x,z,\eta) = \frac{1}{\eta + z(1-\eta)}. \tag{5.82}$$

The generating function we are interested in is equal to the residue at s_-, so that

$$\mathcal{T}_{i,n}\left[B_n^{(i,i)}\right](x,z) = \frac{1}{\sqrt{(1+xz)^2 - 4(\eta + (1-\eta)z)(x(1-\eta) + \eta xz))}}. \tag{5.83}$$

If we particularize this to a balanced beam splitter ($\eta = 1/2$), we simply get

$$\mathcal{T}_{i,n}\left[B_n^{(i,i)}\big|_{\eta=1/2}\right](x,z) = \frac{1}{\sqrt{(1-x)(1-z^2x)}}, \tag{5.84}$$

which is the generating function in i, n of the diagonal sequence $B_n^{(i,i)}$ for $\eta = 1/2$.

5.2.5.2 Asymptotical behaviour of $B_n^{(i,i)}$ for $\eta = 1/2$

Our aim is now to use Tauberian theorems in order to study the asymptotic behaviour of $B_n^{(i,i)}$ for $\eta = 1/2$. Theorem 26 can be generalised, and in case of multiple singularities, each one can be analyzed separately, and the different contributions can be combined in the end [97]. In our case, this must be done in two steps, since our sequence has two indices i and n. We begin by analyzing the behaviour of

$$[z^n]\mathcal{T}_{i,n}\left[B_n^{(i,i)}\big|_{\eta=1/2}\right](x,z) = \mathcal{T}_i\left[B_n^{(i,i)}\big|_{\eta=1/2}\right](x)\,, \tag{5.85}$$

the generating function in i, by studying the behaviour of

$$\mathcal{T}_{i,n}\left[B_n^{(i,i)}\big|_{\eta=1/2}\right](x,z)\,, \tag{5.86}$$

the generating function in i and n. We then investigate the resulting

$$\mathcal{T}_i\left[B_n^{(i,i)}\big|_{\eta=1/2}\right](x) \tag{5.87}$$

in order to conclude about

$$B_n^{(i,i)}\big|_{\eta=1/2}\,. \tag{5.88}$$

Behaviour of $\mathcal{T}_i\left[B_n^{(i,i)}\big|_{\eta=1/2}\right](x)$. The function

$$\mathcal{T}_{i,n}\left[B_n^{(i,i)}\big|_{\eta=1/2}\right](x,z) = \frac{1}{\sqrt{(1-x)(1-z^2x)}} \tag{5.89}$$

has two singularities,

$$z_1(x) = \frac{1}{\sqrt{x}} \quad \text{and} \quad z_2(x) = -\frac{1}{\sqrt{x}}. \tag{5.90}$$

On one hand,

$$\mathcal{T}_{i,n}\left[B_n^{(i,i)}\big|_{\eta=1/2}\right](x,z) \sim \frac{1}{\sqrt{2(1-x)(1-\sqrt{x}z)}}, \quad z \to z_1(x). \tag{5.91}$$

Define the sequence $\beta_{i,n}^{(1)}$ such that

$$\sum_{i,n} \beta_{i,n}^{(1)} x^i z^n = \frac{1}{\sqrt{2(1-x)(1-\sqrt{x}z)}}. \tag{5.92}$$

In other words,

$$\mathcal{T}_{i,n}\left[B_n^{(i,i)}\big|_{\eta=1/2}\right](x,z) \sim \sum_{i,n} \beta_{i,n}^{(1)} x^i z^n, \quad z \to z_1(x). \tag{5.93}$$

Equation (5.92) is the same as (x is positive)

$$\sum_{i,n} \beta_{i,n}^{(1)} x^{i-\frac{n}{2}} z^n = \frac{1}{\sqrt{2(1-x)(1-z)}}. \tag{5.94}$$

Now, for n increasing, according to Tauberian theorems,

$$[z^n]\frac{1}{\sqrt{2(1-x)(1-z)}} \sim \frac{1}{\sqrt{2(1-x)\pi n}}, \tag{5.95}$$

so that

$$[z^n]\sum_{i,n} \beta_{i,n}^{(1)} x^{i-\frac{n}{2}} z^n \sim \frac{1}{\sqrt{2(1-x)\pi n}}, \tag{5.96}$$

$$[z^n]\sum_{i,n} \beta_{i,n}^{(1)} x^i z^n \sim \frac{x^{\frac{n}{2}}}{\sqrt{2(1-x)\pi n}}. \tag{5.97}$$

Using Definition (5.92), we end up with

$$[z^n]\frac{1}{\sqrt{2(1-x)(1-\sqrt{x}z)}} \sim \frac{x^{\frac{n}{2}}}{\sqrt{2(1-x)\pi n}}. \tag{5.98}$$

On the other hand,

$$\mathcal{T}_{i,n}\left[B_n^{(i,i)}\big|_{\eta=1/2}\right](x,z) \sim \frac{1}{\sqrt{2(1-x)(1+\sqrt{x}z)}}, \quad z \to z_2(x). \tag{5.99}$$

We can do the same analysis, and get

$$[z^n]\frac{1}{\sqrt{2(1-x)(1+\sqrt{x}z)}} \sim \frac{(-1)^n x^{\frac{n}{2}}}{\sqrt{2(1-x)\pi n}}. \tag{5.100}$$

As we explained earlier, in the case of two singularities (with the same absolute value), the two asymptotic contributions can be added up [97], so that

$$[z^n]\mathcal{T}_{i,n}\left[B_n^{(i,i)}\big|_{\eta=1/2}\right](x,z) \sim \frac{x^{\frac{n}{2}}}{\sqrt{2(1-x)\pi n}} + \frac{(-1)^n x^{\frac{n}{2}}}{\sqrt{2(1-x)\pi n}}, \tag{5.101}$$

or,

$$\mathcal{T}_i\left[B_n^{(i,i)}\big|_{\eta=1/2}\right](x) \sim \frac{1+(-1)^n}{\sqrt{2\pi n}}\frac{x^{\frac{n}{2}}}{\sqrt{1-x}}. \tag{5.102}$$

The zero contribution for odd n is consistent with the fact that the total input photon number $2i$ is even.

Behaviour of $B_n^{(i,i)}\big|_{\eta=1/2}$. The function

$$\frac{1+(-1)^n}{\sqrt{2\pi n}}\frac{x^{\frac{n}{2}}}{\sqrt{1-x}} \tag{5.103}$$

has only one singularity, $x_0 = 1$. Since the dominant factor is $1/\sqrt{1-x}$ (compared to $x^{\frac{n}{2}}$) when $x \to x_0$, we can focus on it. We have [97]

$$[x^i]\frac{1}{\sqrt{1-x}} \sim \frac{1}{\sqrt{\pi i}}, \tag{5.104}$$

meaning that

$$[x^i]\left(x^{-\frac{n}{2}}\mathcal{T}_i\left[B_n^{(i,i)}\big|_{\eta=1/2}\right](x)\right) \sim \frac{1+(-1)^n}{\sqrt{2\pi n}}\frac{1}{\sqrt{\pi i}}. \tag{5.105}$$

Now,

$$\begin{aligned} x^{-\frac{n}{2}}\mathcal{T}_i\left[B_n^{(i,i)}\big|_{\eta=1/2}\right](x) &= \sum_{i=0}^{\infty} B_n^{(i,i)}\big|_{\eta=1/2}\, x^{i-\frac{n}{2}} \\ &= \sum_{j=-n/2}^{\infty} B_n^{(j+\frac{n}{2},j+\frac{n}{2})}\Big|_{\eta=1/2} x^j, \end{aligned}$$

and $B_n^{(i,i)} = 0$ if $n > 2i$, so that

$$x^{-\frac{n}{2}}\mathcal{T}_i\left[B_n^{(i,i)}\big|_{\eta=1/2}\right](x) = \sum_{j=0}^{\infty} B_n^{(j+\frac{n}{2},j+\frac{n}{2})}\Big|_{\eta=1/2} x^j, \tag{5.106}$$

$$\Rightarrow \quad [x^i]\left(x^{-\frac{n}{2}}\mathcal{T}_i\left[B_n^{(i,i)}\big|_{\eta=1/2}\right](x)\right) = B_n^{(i+\frac{n}{2},i+\frac{n}{2})}\Big|_{\eta=1/2}. \tag{5.107}$$

Equation (5.105),

$$B_n^{(i+\frac{n}{2},i+\frac{n}{2})}\Big|_{\eta=1/2} \sim \frac{1+(-1)^n}{\sqrt{2\pi n}}\frac{1}{\sqrt{\pi i}}, \tag{5.108}$$

or,

$$B_n^{(i,i)}\big|_{\eta=1/2} \sim \frac{1+(-1)^n}{\sqrt{2\pi n}}\frac{1}{\sqrt{\pi\left(i-\frac{n}{2}\right)}}. \tag{5.109}$$

After some simplification, we get

$$B_n^{(i,i)}\big|_{\eta=1/2} \sim \frac{1+(-1)^n}{\pi\sqrt{n\,(2i-n)}}. \tag{5.110}$$

which exactly coincides with the analysis in [102]. The output terms around $n \sim i$ are maximally suppressed, which is reminiscent of the Hong-Ou-Mandel effect (defined later). Interestingly, we can again exploit partial time reversal and extend this analysis to a TMS with $\lambda = 1/2$, giving

$$A_k^{(i,k)} \sim \frac{1+(-1)^i}{2\pi\sqrt{i(2k-i)}}, \quad k, i \to \infty. \tag{5.111}$$

5.3 Generating functions for N-mode passive Gaussian unitaries

In Chapter 6, the expressions of the generating functions for the transition probabilities will prove very valuable for the derivation of relations characterising the effect of quantum interferences. The generating functions we have computed until now always involved two modes. It seems natural to ask the question whether this analysis can easily be extended to an arbitrary number of modes N. This is the purpose of this section, in which we choose to focus on N-mode passive Gaussian unitaries; that is, Gaussian unitaries which preserve the energy, as they are of particular interest, since they can actually be combined with Gaussian unitaries acting on a smaller number of modes, in order to generate general N-mode Gaussian unitaries. As we are going to demonstrate in Chapter 6, the generating function we compute here will be crucial in the derivation of an equation describing the effect of interferences in N-mode passive Gaussian unitaries.

Consider an N-mode passive interferometer whose effect on the bosonic field operators in phase space is characterised by an orthogonal matrix $\mathbf{U}$ of dimension N, i.e,

$$\hat{\mathbf{a}} \to \mathbf{U}\hat{\mathbf{a}}, \quad \hat{\mathbf{a}} = (\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_N)\,. \tag{5.112}$$

Define the transition probabilities

$$B_{\mathbf{n}}^{(\mathbf{i})} = \left| \left(\prod_{r=1}^{N} \langle n_r| \right) U^{\mathrm{PI}} \left(\prod_{s=1}^{N} |i_s\rangle \right) \right|^2 = \left| \langle n_1, n_2, \ldots, n_N|\, U^{\mathrm{PI}}\, |i_1, i_2, \ldots, i_N\rangle \right|^2, \tag{5.113}$$

where $\mathbf{i} = (i_1, i_2, \ldots, i_N) \in \mathbb{N}_0^N$, $\mathbf{n} = (n_1, n_2, \ldots, n_N) \in \mathbb{N}_0^N$, $\mathbb{N}_0$ being the set of all natural numbers (including zero) and U^{PI} is the unitary which characterises the effect of the interferometer in state space. The $2N$-variate generating function of the transition probabilities is given

by

$$f^{\mathrm{PI}}(\mathbf{x},\mathbf{z}) = \mathcal{T}_{\mathbf{i},\mathbf{n}}\left[B_{\mathbf{n}}^{(\mathbf{i})}\right](\mathbf{x},\mathbf{z}) = \sum_{\mathbf{i}\in\mathbb{N}_0^N}\sum_{\mathbf{n}\in\mathbb{N}_0^N} B_{\mathbf{n}}^{(\mathbf{i})}\left(\prod_{s=1}^{N} x_s^{i_s}\right)\left(\prod_{r=1}^{N} z_r^{n_r}\right), \tag{5.114}$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_N) \in \{[0,1)\}^N$ and $\mathbf{z} = (z_1, z_2, \ldots, z_N) \in \{[0,1)\}^N$. If we use the definition of the probabilities $B_{\mathbf{n}}^{(\mathbf{i})}$, we get

$$f^{\mathrm{PI}}(\mathbf{x},\mathbf{z}) = \sum_{\mathbf{i}\in\mathbb{N}_0^N}\sum_{\mathbf{n}\in\mathbb{N}_0^N}\left|\left(\prod_{r=1}^{N}\langle n_r|\right) U^{\mathrm{PI}}\left(\prod_{s=1}^{N}|i_s\rangle\right)\right|^2\left(\prod_{s=1}^{N} x_s^{i_s}\right)\left(\prod_{r=1}^{N} z_r^{n_r}\right). \tag{5.115}$$

Like we did in the previous section, we introduce thermal states into the picture by rewriting the generating function as

$$f^{\mathrm{PI}}(\mathbf{x},\mathbf{z}) = \left(\prod_{s=1}^{N}(1-x_s)^{-1}\right)\left(\prod_{r=1}^{N}(1-z_r)^{-1}\right)\mathrm{Tr}\left[U^{\mathrm{PI}}\Gamma_{\mathbf{x}}U^{\mathrm{PI}\dagger}\Gamma_{\mathbf{z}}\right], \tag{5.116}$$

where

$$\Gamma_{\mathbf{x}} = \bigotimes_{s=1}^{N}\tau_{x_s}, \tag{5.117}$$

and τ_x is a Gaussian thermal state of mean number of photon μ_x such that

$$\tau_x = (1-x)\sum_{i=0}^{\infty} x^i\,|i\rangle\langle i|, \quad \mu_x = \frac{x}{1-x}, \quad 0 \le x < 1. \tag{5.118}$$

Since the calculations of the generating function are quite involved in the N-mode case, we choose to include them in the appendix rather than in the main text. The interested reader is referred to Appendix E.1, in which we prove that f^{PI} has the following form.

Lemma 7. *The generating function $f^{\mathrm{PI}}(\mathbf{x},\mathbf{z})$ of the probabilities $B_{\mathbf{n}}^{(\mathbf{i})}$ satisfies*

$$f^{\mathrm{PI}}(\mathbf{x},\mathbf{z}) = \left(\sum_{m=0}^{N}(-1)^m\sum_{\alpha\in\mathcal{R}_m^{(N)}}\sum_{\beta\in\mathcal{R}_m^{(N)}}\left(\det\left[\mathbf{U}(\beta,\alpha)\right]\right)^2\det\left[\mathbf{X}(\alpha)\right]\det\left[\mathbf{Z}(\beta)\right]\right)^{-1}, \tag{5.119}$$

where $\mathbf{X}$ and $\mathbf{Z}$ are diagonal matrices whose diagonals are given by the vectors $\mathbf{x}$ and $\mathbf{z}$, respectively, $\mathbf{M}(\beta,\alpha)$ is the sub-matrix of $\mathbf{M}$ corresponding to the rows, columns whose indices belong in β,α, respectively (furthermore, we write $\mathbf{M}(\alpha)$ for $\mathbf{M}(\alpha,\alpha)$), and $\mathcal{R}_m^{(N)}$ is the set of all subsets of $\{1,2,\ldots,N\}$ of cardinality m.

The expression of the $2N$-variate generating function in Equation (5.119) is to be compared with Equation (5.63) which was computed for a beam-splitter unitary. Obviously, Equation (5.63) can be obtained starting from Equation (5.119) by setting $N = 2$. A first thing to notice about Equation (5.119) is that in each term $\left(\det\left[\mathbf{U}(\beta,\alpha)\right]\right)^2\det\left[\mathbf{X}(\alpha)\right]\det\left[\mathbf{Z}(\beta)\right]$ of the de-

nominator of the generating function, the sets α and β will always have the same cardinalities, so that the matrices $\mathbf{X}(\alpha)$ and $\mathbf{Z}(\beta)$ will always have the same dimensions. This means that the number of variables z_r characterising the outputs will always be accompanied by the same number of variables x_s characterising the inputs. The second thing to understand, is that these factors $\det[\mathbf{X}(\alpha)]\det[\mathbf{Z}(\beta)]$ will be multiplied by the right squared minor $(\det[\mathbf{U}(\beta,\alpha)])^2$ of the orthogonal matrix $\mathbf{U}$ which governs the evolution of the annihilation operators in phase space. As long as one understands what each term $(\det[\mathbf{U}(\beta,\alpha)])^2\det[\mathbf{X}(\alpha)]\det[\mathbf{Z}(\beta)]$ actually looks like and represents, one can readily acknowledge that Equation (5.119) is actually quite elegant and simple for such a complex object. As we already mentioned, it will prove very useful in the proof of some relation characterising quantum interferences for N-mode passive Gaussian unitaries.

6

Multi-photon interference effects in Gaussian transformations

In this chapter, we address multi-photon interference by exploiting the framework developped in Chapter 5, which relies on the generating function of Gaussian matrix elements in Fock basis. The latter can be expressed in a simple form while it enables accessing intrinsically non-Gaussian features such as multi-photon transition probabilities. In particular, it exhibits a suppression term that generalises the Hong-Ou-Mandel effect to many photons. Remarkably, applying this tool to active transformations, we find an analogous effect that had been left unnoticed in an optical amplification medium of gain 2. In Section 6.1, we prove a recurrence relation for the transition probabilities in the case of the most important 2-mode Gaussian unitaries, namely the beam splitter and the two-mode squeezer. The relation clearly illustrates the effect of quantum interference as a consequence of the non-distinguishability of bosons. This will lead us to a generalisation of the Hong-Ou-Mandel effect for both cases in Section 6.2. Finally, in Section 6.3, we prove the existence of a similar relation for the transition probabilities involving N-mode passive Gaussian transformations.

6.1 Transition probabilities of two-mode Gaussian unitaries

6.1.1 Recurrence for the transition probabilities in a beam splitter

In Chapter 5, we illustrated the power of the generating functions in a framework involving Gaussian unitaries by easily computing expressions for the transition amplitudes. As we are now going to see, the advantage of generating functions becomes even more evident when turning to transition probabilities $B_n^{(i,k)} = |b_n^{(i,k)}|^2$. Taking the square modulus of Equation 5.49 can be quite cumbersome, as shown in Appendix D. Yet, the generating functions can be exploited to prove the following recurrence equation.

Theorem 27. *Let $B_n^{(i,k)} = |\langle n,m| U_\eta^{\mathrm{BS}} |i,k\rangle|^2$ be the transition probabilities of a beam splitter of transmittance η with $m = i+k-n$, then*

$$B_n^{(i,k)} = \tilde{B}_n^{(i,k,j)} - \tilde{B}_{n-1}^{(i-1,k-1,j-1)}, \tag{6.1}$$

for all j such that $0 \leq j \leq i+k$, where

$$\tilde{B}_n^{(i,k,j)} = \sum_{l=\max(0,j-i)}^{\min(j,k)} \left\{ B_\bullet^{(j-l,l)} * B_\bullet^{(i-j+l,k-l)} \right\}_n . \tag{6.2}$$

Here, the convolution is noted as $\{B_\bullet^{(i,k)} * B_\bullet^{(j,l)}\}_n = \sum_{m=0}^{n} B_m^{(i,k)} B_{n-m}^{(j,l)}$. This recurrence is obvious for $j = 0$ given that $B_n^{(0,0)} = \delta_{n,0}$, and can otherwise be proven easily for either of the indices (i,k,n) equal to zero by using the property that the convolution of two binomials stays a binomial.

Proof. Consider the probability

$$B_n^{(i,k)} = \langle n| \mathrm{Tr}_2 \left[U_\eta |i,k\rangle \langle i,k| U_\eta^\dagger \right] |n\rangle . \tag{6.3}$$

We need to show that

$$B_{n+1}^{(i+1,k+1)} \stackrel{?}{=} \tilde{B}_{n+1}^{(i+1,k+1,j+1)} - \tilde{B}_n^{(i,k,j)}, \quad 0 \leq j \leq i+k+1, \tag{6.4}$$

since we know it is obviously true for $j = -1$. Furthermore, notice that

$$\tilde{B}_n^{(i,k,j)} = 0 \quad \text{if } j < 0 \text{ or } j > i+k. \tag{6.5}$$

Using this information, we need to show that

$$\sum_{j=0}^{i+k+1} B_{n+1}^{(i+1,k+1)} s^j \stackrel{?}{=} \mathcal{T}_j\left[\tilde{B}_{n+1}^{(i+1,k+1,j+1)}\right](s) - \mathcal{T}_j\left[\tilde{B}_n^{(i,k,j)}\right](s), \quad \forall s, \tag{6.6}$$

where, for instance,

$$\mathcal{T}_j\left[\tilde{B}_n^{(i,k,j)}\right](s) = \sum_{j=0}^{\infty} \tilde{B}_n^{(i,k,j)} s^j. \tag{6.7}$$

The shifting property of generating functions leads to

$$\begin{aligned} &\sum_{j=0}^{i+k+1} B_{n+1}^{(i+1,k+1)} s^j \stackrel{?}{=} \frac{1}{s}\left(\mathcal{T}_j\left[\tilde{B}_{n+1}^{(i+1,k+1,j)}\right](s) - \tilde{B}_{n+1}^{(i+1,k+1,0)}\right) - \mathcal{T}_j\left[\tilde{B}_n^{(i,k,j)}\right](s), \\ \Leftrightarrow \quad &\sum_{j=0}^{i+k+1} B_{n+1}^{(i+1,k+1)} s^j \stackrel{?}{=} \frac{1}{s}\left(\mathcal{T}_j\left[\tilde{B}_{n+1}^{(i+1,k+1,j)}\right](s) - B_{n+1}^{(i+1,k+1)}\right) - \mathcal{T}_j\left[\tilde{B}_n^{(i,k,j)}\right](s). \end{aligned}$$

Now,

$$\sum_{j=0}^{i+k+1} B_{n+1}^{(i+1,k+1)} s^j = \frac{1}{s}\sum_{j=0}^{i+k+1} B_{n+1}^{(i+1,k+1)} s^{j+1} = \frac{1}{s}\sum_{j=1}^{i+k+2} B_{n+1}^{(i+1,k+1)} s^j, \tag{6.8}$$

meaning that what we are trying to show can be written

$$\sum_{j=0}^{i+k+2} B_{n+1}^{(i+1,k+1)} s^j \stackrel{?}{=} \mathcal{T}_j\left[\tilde{B}_{n+1}^{(i+1,k+1,j)}\right](s) - s\mathcal{T}_j\left[\tilde{B}_n^{(i,k,j)}\right](s), \tag{6.9}$$

at least for $s \neq 0$ (which is sufficient). Next, we take the generating functions in i, getting

$$\mathcal{T}_i\left[\sum_{j=0}^{i+k+2} B_{n+1}^{(i+1,k+1)} s^j\right](x) \stackrel{?}{=} \mathcal{T}_{i,j}\left[\tilde{B}_{n+1}^{(i+1,k+1,j)}\right](x,s) - s\mathcal{T}_{i,j}\left[\tilde{B}_n^{(i,k,j)}\right](x,s). \tag{6.10}$$

Using the shifting property again for the index i, we have

$$\begin{aligned} &\mathcal{T}_i\left[\sum_{j=0}^{i+k+1} B_{n+1}^{(i,k+1)} s^j\right](x) - \sum_{j=0}^{k+1} B_{n+1}^{(0,k+1)} s^j \\ &\quad \stackrel{?}{=} \mathcal{T}_{i,j}\left[\tilde{B}_{n+1}^{(i,k+1,j)}\right](x,s) - \mathcal{T}_j\left[\tilde{B}_{n+1}^{(0,k+1,j)}\right](s) - xs\mathcal{T}_{i,j}\left[\tilde{B}_n^{(i,k,j)}\right](x,s), \end{aligned} \tag{6.11}$$

for $x \neq 0$. Now, it was shown in [103] that the recurrence relation (6.1) is true for $i = 0$ and $j = 0$, but it can easily be extended for $i = 0$ to any j such that $0 \leq j \leq k$, by simply using the fact that the convolution of two binomials gives another binomial. This means that

$$B_n^{(0,k)} = \tilde{B}_n^{(0,k,j)}, \quad 0 \leq j \leq k, \tag{6.12}$$

or,

$$B_{n+1}^{(0,k+1)} = \tilde{B}_{n+1}^{(0,k+1,j)}, \quad 0 \leq j \leq k+1. \tag{6.13}$$

Taking the sum over j going from 0 to $k+1$ after multiplying by s^j, we get

$$\sum_{j=0}^{k+1} B_{n+1}^{(0,k+1)} s^j = \mathcal{T}_j \left[\tilde{B}_{n+1}^{(0,k+1,j)} \right](s), \tag{6.14}$$

which is actually kind of an initial condition for i. This last relation allows us to simplify Equation (6.11) into

$$\mathcal{T}_i \left[\sum_{j=0}^{i+k+1} B_{n+1}^{(i,k+1)} s^j \right](x) \stackrel{?}{=} \mathcal{T}_{i,j} \left[\tilde{B}_{n+1}^{(i,k+1,j)} \right](x,s) - xs\mathcal{T}_{i,j} \left[\tilde{B}_n^{(i,k,j)} \right](x,s). \tag{6.15}$$

Obviously, the same can be done with indices k and n, and using what is known for $k = 0$ or $n = 0$, we finally end up with

$$\mathcal{T}_{i,k,n} \left[\sum_{j=0}^{i+k} B_n^{(i,k)} s^j \right](x,y,z) \stackrel{?}{=} (1-xyzs)\mathcal{T}_{i,k,n,j} \left[\tilde{B}_n^{(i,k,j)} \right](x,y,z,s). \tag{6.16}$$

Now, it can be easily shown that

$$\mathcal{T}_j \left[\tilde{B}_n^{(i,k,j)} \right](s) = \sum_{l=0}^{k} \sum_{j=0}^{i} \left\{ B_\bullet^{(j,l)} * B_\bullet^{(i-j,k-l)} \right\}_n s^{j+l}, \tag{6.17}$$

which is a triple convolution over the indices i, k and n, so that [96]

$$\mathcal{T}_{i,k,n,j} \left[\tilde{B}_n^{(i,k,j)} \right](x,y,z,s) = f_\eta^{\mathrm{BS}}(x,y,z,1) f_\eta^{\mathrm{BS}}(sx,sy,z,1). \tag{6.18}$$

Having in mind that

$$\sum_{j=0}^{i+k} s^j = \frac{1-s^{i+k+1}}{1-s}, \tag{6.19}$$

the left-hand side of Equation (6.16) can be computed to be

$$\mathcal{T}_{i,k,n} \left[\sum_{j=0}^{i+k} B_n^{(i,k)} s^j \right](x,y,z) = \frac{1}{1-s} \left(f_\eta^{\mathrm{BS}}(x,y,z,1) - s f_\eta^{\mathrm{BS}}(sx,sy,z,1) \right). \tag{6.20}$$

Finally, the recurrence relation will be true if and only if

$$f_\eta^{\mathrm{BS}}(x,y,z,1) - s f_\eta^{\mathrm{BS}}(sx,sy,z,1) \stackrel{?}{=} (1-xyzs)(1-s) f_\eta^{\mathrm{BS}}(x,y,z,1) f_\eta^{\mathrm{BS}}(sx,sy,z,1). \tag{6.21}$$

This will be the case if and only if

$$\frac{1}{f_\eta^{\mathrm{BS}}(sx, sy, z, 1)} - s\frac{1}{f_\eta^{\mathrm{BS}}(x, y, z, 1)} \stackrel{?}{=} (1 - xyzs)(1 - s), \tag{6.22}$$

which is readily checked using the definition of $f_\eta^{\mathrm{BS}}(sx, sy, z, 1)$. □

The right-hand side of Equation (6.1) involves two terms $\tilde{B}_n^{(i,k,j)}$ and $\tilde{B}_{n-1}^{(i-1,k-1,j-1)}$. Notice that the second term is multiplied by a negative coefficient. It can actually be identified as interferences which are specific to the quantum realm, in the sense that the bosons entering the beam splitter are indistinguishable. This can be made clearer by analysing the "classical" case, which involves distinguishable photons.

6.1.2 Comparison between distinguishable and indistinguishable photons

In order to identify the effect of quantum interferences, we express a recurrence relation for the probabilities associated with a classical description of the beam splitter. By this, we mean that the photons are distinguishable, and can be treated like balls of, say, different colours, which take different paths in the beam splitter. In this case, the probability $p(n|i,k)$ to find n photons in one output, given i photons on one input and k on the other, can be computed using standard probability theory, and can be written as a convolution,

$$p(n|i,k) = \{p(\bullet|i,0) * p(\bullet|0,k)\}_n = \sum_{\tilde{n}=0}^{n} p(\tilde{n}|i,0)p(n-\tilde{n}|0,k). \tag{6.23}$$

In fact, one understands that this relation can be generalised to

$$p(n|i,k) = \sum_{\tilde{n}=0}^{n} p(\tilde{n}|j,l)p(n-\tilde{n}|i-j,k-l), \quad j \leq i \text{ and } l \leq k. \tag{6.24}$$

One easy way to show this is to take the generating functions of both Equations (6.23) and (6.24), and check that they are consistent. If one defines the generating function of the classical probability $p(n|i,k)$ as

$$\tilde{f}(x,y,z) = \mathcal{T}_{i,k,n}\left[p(n|i,k)\right](x,y,z) = \sum_{i,k,n} p(n|i,k)x^i y^k z^n, \tag{6.25}$$

the generating function of Equation (6.23) can be found to be

$$\tilde{f}(x,y,z) = \tilde{f}(x,0,z)\tilde{f}(0,y,z), \tag{6.26}$$

where we used the fact that the generating function of a convolution is given by a product. Now, classically (in the sense defined above), the probability $p(n|i,0)$ is given by a simple binomial

in our setup, i.e,

$$p(n|i,0) = \binom{i}{n}\eta^n(1-\eta)^{i-n}, \tag{6.27}$$

like in the quantum case. Obviously, the same can be said about the probability $p(n|0,k)$. Using this, we have that

$$\tilde{f}(x,0,z) = \frac{1}{1-\eta xz-(1-\eta)x}, \tag{6.28}$$

while

$$\tilde{f}(0,y,z) = \frac{1}{1-\eta y-(1-\eta)yz}, \tag{6.29}$$

implying that

$$\tilde{f}(x,y,z) = \frac{1}{(1-\eta xz-(1-\eta)x)(1-\eta y-(1-\eta)yz)}. \tag{6.30}$$

This expression for the generating function of a beam splitter is to be compared with Equation (5.63) [taking $w=1$], the difference being due to whether photons are distinguishable or not. Now, in order to show that Equation (6.24) is consistent with Equation (6.23), one only needs to compute the generating function of the former. We begin by multiplying both its sides by s_1^j and s_2^l, before summing them over both indices j and l, going from 0 to i and k, respectively, i.e,

$$\sum_{j,l=0}^{i,k} p(n|i,k)s_1^j s_2^l = \sum_{j,l=0}^{i,k}\sum_{\tilde{n}=0}^{n} p(\tilde{n}|j,l)p(n-\tilde{n}|i-j,k-l)s_1^j s_2^l, \tag{6.31}$$

which will be true for all s_1 and s_2 if and only if Equation (6.24) is true for all i,k such that $0 \leq j \leq i$ and $0 \leq l \leq k$. Now, the left-hand side of the last equation can trivially be computed to be

$$\sum_{j,l=0}^{i,k} p(n|i,k)s_1^j s_2^l = \frac{(1-s_1^{i+1})(1-s_2^{k+1})}{(1-s_1)(1-s_2)}p(n|i,k), \tag{6.32}$$

meaning that its generating function in i,k,n is given by

$$\mathcal{T}_{i,k,n}\left[\sum_{j,l=0}^{i,k} p(n|i,k)s_1^j s_2^l\right](x,y,z) = \frac{\tilde{f}(x,y,z)-s_1\tilde{f}(xs_1,y,z)-s_2\tilde{f}(x,ys_2,z)+s_1s_2\tilde{f}(xs_1,ys_2,z)}{(1-s_1)(1-s_2)}. \tag{6.33}$$

On another hand, the right-hand side of Equation (6.31) happens to be given by a triple convolution in i,k,n of the two objects $p(n|i,k)s_1^i s_2^k$ and $p(n|i,k)$. Consequently, its generating function can be calculated to be $\tilde{f}(xs_1,ys_2,z)\tilde{f}(x,y,z)$. This means that Equation (6.24) will be consistent with Equation (6.23) if and only if

$$\begin{aligned}&\tilde{f}(x,y,z)-s_1\tilde{f}(xs_1,y,z)\\ &\quad-s_2\tilde{f}(x,ys_2,z)+s_1s_2\tilde{f}(xs_1,ys_2,z) = (1-s_1)(1-s_2)\tilde{f}(xs_1,ys_2,z)\tilde{f}(x,y,z),\end{aligned}$$

which is easily verified.

Now that we are convinced that Equation (6.24) properly describes a beam splitter which would involve distinguishable particles, we choose to modify it by replacing j by $j-l$, getting

$$p(n|i,k) = \sum_{\tilde{n}=0}^{n} p(\tilde{n}|j-l,l)p(n-\tilde{n}|i-j+l,k-l), \tag{6.34}$$

where j and l should verify $j-l \leq i$ and $l \leq k$ (which implies $j \leq i+k$) for the relation to be true. We can consequently sum it over l, taking the right limits for the sum so that the relation is true, ending up with

$$\sum_{l=\max(0,j-i)}^{\min(j,k)} p(n|i,k) = \sum_{l=\max(0,j-i)}^{\min(j,k)} \sum_{\tilde{n}=0}^{n} p(\tilde{n}|j-l,l)p(n-\tilde{n}|i-j+l,k-l), \quad 0 \leq j \leq i+k, \tag{6.35}$$

or,

$$p(n|i,k) = \frac{1}{c(i,k,j)} \sum_{l=\max(0,j-i)}^{\min(j,k)} \sum_{\tilde{n}=0}^{n} p(\tilde{n}|j-l,l)p(n-\tilde{n}|i-j+l,k-l), \quad 0 \leq j \leq i+k, \tag{6.36}$$

where the coefficient c is such that

$$c(i,k,j) = \begin{cases} 1-j+i+k & \text{if } j \geq i \text{ and } j \geq k, \\ 1+\min(i,k,j) & \text{else.} \end{cases} \tag{6.37}$$

We see that Equation (6.36) is almost identical to Equation (6.1), if the latter only had its first term $\tilde{B}_n^{(i,k,j)}$ on the right-hand side, with a normalisation coefficient c. The second term $\tilde{B}_{n-1}^{(i-1,k-1,j-1)}$ of the right-hand side of Equation (6.1) is actually suppressed in what could be interpreted as a classical description of the beam splitter. As a consequence, it seems natural to associate $\tilde{B}_{n-1}^{(i-1,k-1,j-1)}$ with quantum interferences. This will be made even clearer when taking the particular case of $j=1$, as we will do in Section 6.2, in which we will show how these interferences can lead to the Hong-Ou-Mandel effect.

6.1.3 Recurrence for the transition probabilities in a two-mode squeezer

An even more appealing application of generating functions is to describe multi-photon interferences in a two-mode squeezer. Given that f_η^{BS} and f_λ^{TMS} are linked by a partial time reversal, their corresponding transition probabilities satisfy

$$A_n^{(i,k)} = (1-\lambda)B_i^{(n,k)}, \tag{6.38}$$

It is then easy to prove the following recurrence.

Theorem 28. *Let* $A_n^{(i,k)} = |\langle n,m|\, U_\lambda^{\mathrm{TMS}}\, |i,k\rangle|^2$ *be the transition probabilities of a two-mode squeezer with parameter* $\lambda = \tanh^2(r)$ *and* $m = n+k-i$*, then*

$$(1-\lambda)\, A_n^{(i,k)} = \tilde{A}_{n,j}^{(i,k)} - \tilde{A}_{n-1,j-1}^{(i-1,k-1)}, \tag{6.39}$$

for all j *such that* $0 \leq j \leq n+k$*, where*

$$\tilde{A}_{n,j}^{(i,k)} = \sum_{l=\max(0,j-k)}^{\min(j,n)} \left\{ A_l^{(\bullet,j-l)} * A_{n-l}^{(\bullet,k-j+l)} \right\}_i . \tag{6.40}$$

Here, the convolution acts on variable i.

Proof. The recurrence relation can be derived in the case of a two-mode squeezer, by taking advantage of the property of partial time reversal. Indeed, since

$$f_\lambda^{\mathrm{TMS}}(x,y,z,w) = (1-\lambda) f_{1-\lambda}^{\mathrm{BS}}(z,y,x,w), \tag{6.41}$$

one readily understands that the corresponding probabilities verify Equation (6.38), or,

$$A_i^{(n,k)} = (1-\lambda) B_n^{(i,k)}. \tag{6.42}$$

Multiplying the recurrence (6.1) by $(1-\lambda)^2$ gives

$$\begin{aligned}(1-\lambda)^2 B_n^{(i,k)} = &\sum_{l=\max(0,j-i)}^{\min(j,k)} \sum_{m=0}^{n} (1-\lambda) B_m^{(j-l,l)} (1-\lambda) B_{n-m}^{(i-j+l,k-l)} \\ &- \sum_{l=\max(0,j-i)}^{\min(j,k)-1} \sum_{m=0}^{n-1} (1-\lambda) B_m^{(j-1-l,l)} (1-\lambda) B_{n-1-m}^{(i-j+l,k-1-l)}, \quad 0 \leq j \leq i+k.\end{aligned}$$

Using Equation (6.42), it can be rewritten as

$$\begin{aligned}(1-\lambda) A_i^{(n,k)} = &\sum_{l=\max(0,j-i)}^{\min(j,k)} \sum_{m=0}^{n} A_{j-l}^{(m,l)} A_{i-j+l}^{(n-m,k-l)} \\ &- \sum_{l=\max(0,j-i)}^{\min(j,k)-1} \sum_{m=0}^{n-1} A_{j-1-l}^{(m,l)} A_{i-j+l}^{(n-1-m,k-1-l)}, \quad 0 \leq j \leq i+k.\end{aligned}$$

Exchanging i and n and performing the change of variables $l = j - r$ in the sums over l (before finally replacing r by l), we end up with

$$(1-\lambda) A_n^{(i,k)} = \tilde{A}_{n,j}^{(i,k)} - \tilde{A}_{n-1,j-1}^{(i-1,k-1)}, \quad 0 \leq j \leq n+k, \tag{6.43}$$

where

$$\tilde{A}^{(i,k)}_{n,j} = \sum_{l=\max(0,j-k)}^{\min(j,n)} \sum_{m=0}^{i} A^{(m,j-l)}_{l} A^{(i-m,k-j+l)}_{n-l} = \sum_{l=\max(0,j-k)}^{\min(j,n)} \left\{ A^{(\bullet,j-l)}_{l} * A^{(\bullet,k-j+l)}_{n-l} \right\}_{i} . \quad (6.44)$$

□

It is rather interesting to see that the same kind of interference effects can be witnessed in a two-mode squeezer. This was, of course, to be expected, since the two-mode squeezer is itself a quantum object based on the beam splitter. Still, the way interferences enter the picture is very similar to what happens in a simple beam splitter, mainly because of the property of partial time reversal. Under the circumstances, we are tempted to look for a phenomenon similar to the Hong-Ou-Mandel effect in an active transformation that is the two-mode squeezer. This is the purpose of the next section.

6.2 Generalised Hong-Ou-Mandel effects

The Hong-Ou-Mandel effect was first highlighted in 1987 [17]. It basically describes a phenomenon which is primarily the consequence of the indistinguishability of bosons. As such, the Hong-Ou-Mandel effect can be considered as a purely quantum effect. It can be witnessed in a balanced beam splitter (transmittance $\eta = 1/2$), in which two identical single photons automatically follow the same path. One can understand the Hong-Ou-Mandel as a consequence of the destructive interference between two possible situations, one in which both photons cross the beam splitter and the other in which both are reflected, as depicted in Figure 6.2.1.

In this section, we generalise the mathematical framework which puts forth the Hong-Ou-Mandel effect in a beam splitter. More interestingly, we show the existence of a similar phenomenon in an active Gaussian transformation, specifically the two-mode squeezer. To our knowledge, this effect has remained unnoticed until now.

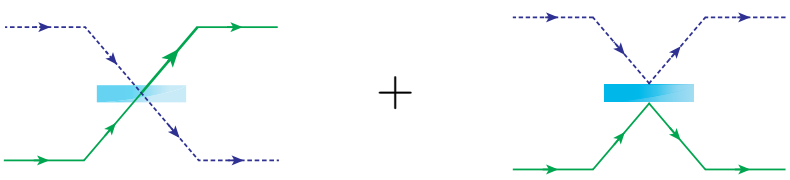

Figure 6.2.1: Hong-Ou-Mandel effect as a consequence of the indistinguishability between two photons impinging on a beam splitter of transmittance $\eta = 1/2$. It originates from the destructive interference between both photons crossing the beam splitter with amplitude $\sqrt{\eta} \times \sqrt{\eta}$ and being reflected with amplitude $-\sqrt{1-\eta} \times \sqrt{1-\eta}$.

6.2.1 Multi-photon Hong-Ou-Mandel effect in a beam splitter

The recurrence (6.1) can be nicely interpreted in the context of the Hong-Ou-Mandel effect by taking the special case of $j = 1$ for $i, k > 0$ and replacing the probabilities $B_n^{(1,0)}$ or $B_n^{(0,1)}$ by their values, namely η or $1 - \eta$ (or 0 for $n > 1$).

Corollary 4. *The recurrence for $B_n^{(i,k)}$ when $j = 1$ is*

$$B_n^{(i,k)} = \eta\, B_{n-1}^{(i-1,k)} + (1-\eta)\, B_n^{(i-1,k)} + \eta\, B_n^{(i,k-1)} + (1-\eta)\, B_{n-1}^{(i,k-1)} - B_{n-1}^{(i-1,k-1)}. \tag{6.45}$$

Proof. If we choose $j = 1$, $i > 0$ and $k > 0$ in Equation (6.1), we end up with

$$B_n^{(i,k)} = B_1^{(1,0)} B_{n-1}^{(i-1,k)} + B_0^{(1,0)} B_n^{(i-1,k)} + B_0^{(0,1)} B_n^{(i,k-1)} + B_1^{(0,1)} B_{n-1}^{(i,k-1)} - B_0^{(0,0)} B_{n-1}^{(i-1,k-1)}. \tag{6.46}$$

Using the initial conditions

$$B_0^{(0,0)} = 1, \quad \begin{cases} B_1^{(1,0)} = \eta, \\ B_0^{(1,0)} = 1 - \eta, \end{cases} \quad \text{and} \quad \begin{cases} B_0^{(0,1)} = \eta, \\ B_1^{(0,1)} = 1 - \eta, \end{cases} \tag{6.47}$$

we get (6.45). □

As illustrated in Figure 6.2.2, the first four terms of the right-hand side of Equation (6.45) corroborate the classical intuition one may have about $B_n^{(i,k)}$: one should add the probabilities corresponding to the different scenarios in which the nth photon has not reached the beam

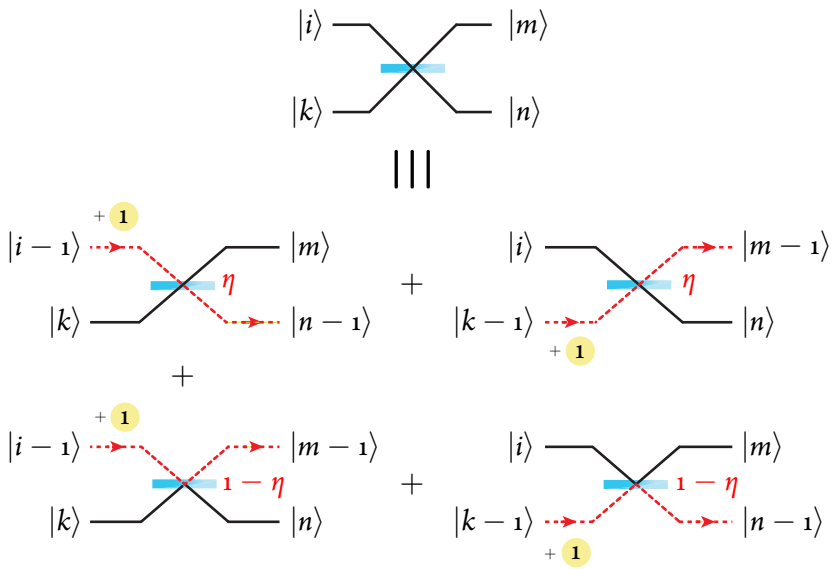

Figure 6.2.2: Classical component of the recurrence formula (6.45) for the transition probabilities $B_n^{(i,k)}$ in a beam splitter.

splitter yet, multiplied by the right coefficient depending on which path it takes. For example, $B_{n-1}^{(i-1,k)}$ must be multiplied by η since the extra photon must be injected on the input mode $\hat{a}$ and exit on the output mode $\hat{a}$. Crucially, as a consequence of the bosonic statistics, there exists a fifth term in Equation (6.45) that accounts for quantum interferences and may be viewed as a generalised interference suppression term. In the special case when $i = k = 1$ and $\eta = 1/2$, it gives rise to the standard Hong-Ou-Mandel effect, see Figure 6.2.1. Note that if $k = 0$ and $i > 0$, the interference term disappears and one gets the recurrence $B_n^{(i,0)} = \eta\, B_{n-1}^{(i-1,0)} + (1-\eta)\, B_n^{(i-1,0)}$ that was derived in the context of majorization theory applied to bosonic transformations [103].

Note that if we set $j = 1$ in Equation (6.36), and choose $i \geq 1$ and $k \geq 1$, we end up with

$$\begin{aligned} p(n|i,k) = \frac{1}{2}\Big(&p(1|1,0)p(n-1|i-1,k) + p(0|1,0)p(n|i-1,k) \\ &+ p(0|0,1)p(n|i,k-1) + p(1|0,1)p(n-1|i,k-1)\Big). \end{aligned} \tag{6.48}$$

Replacing the probability $p(1|1,0)$ and the like by their values in this expression, we get

$$\begin{aligned} p(n|i,k) = \frac{1}{2}\Big(&\eta\, p(n-1|i-1,k) + (1-\eta)\, p(n|i-1,k) \\ &+ \eta\, p(n|i,k-1) + (1-\eta)\, p(n-1|i,k-1)\Big). \end{aligned}$$

Up to the normalisation constant $1/2$, this is exactly Equation (6.45) without the fifth term (associated with quantum interferences) in the right-hand side.

6.2.2 Hong-Ou-Mandel effect in a two-mode squeezer

By taking $j = 1$ and $k, n > 0$ in the recurrence relation of Equation (6.39), one can deduce the following corollary.

Corollary 5. *The recurrence for $A_n^{(i,k)}$ when $j = 1$ is*

$$A_n^{(i,k)} = \lambda\, A_n^{(i-1,k-1)} + (1-\lambda)\, A_{n-1}^{(i-1,k)} + \lambda\, A_{n-1}^{(i,k)} + (1-\lambda)\, A_n^{(i,k-1)} - A_{n-1}^{(i-1,k-1)}. \tag{6.49}$$

Proof. If we choose $j = 1$, $n > 0$ and $k > 0$ in Equation (6.39), we get

$$(1-\lambda)A_n^{(i,k)} = A_0^{(0,1)}A_n^{(i,k-1)} + A_0^{(1,1)}A_n^{(i-1,k-1)} + A_1^{(0,0)}A_{n-1}^{(i-0,k)} + A_1^{(1,0)}A_{n-1}^{(i-1,k)} - A_{n-1}^{(i-1,k-1)}. \tag{6.50}$$

In the case of the two-mode squeezer, if one uses Equation (6.38) again, where $\eta = 1-\lambda$ in the probabilities of the beam splitter, the initial conditions can be found to be

$$A_0^{(0,0)} = 1-\lambda, \quad \begin{cases} A_0^{(0,1)} = (1-\lambda)^2, \\ A_0^{(1,1)} = (1-\lambda)\lambda, \end{cases} \quad \text{and} \quad \begin{cases} A_1^{(0,0)} = (1-\lambda)\lambda, \\ A_1^{(1,0)} = (1-\lambda)^2, \end{cases} \tag{6.51}$$

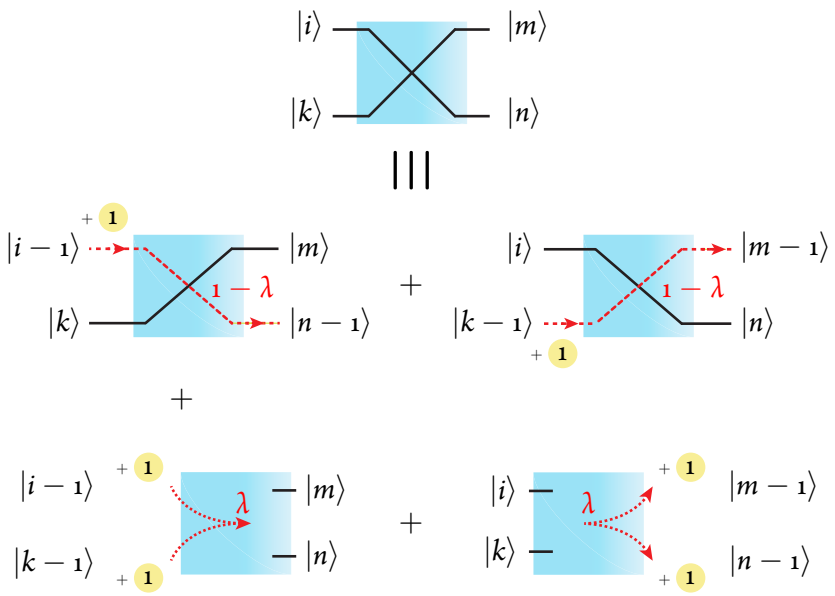

Figure 6.2.3: Classical component of the recurrence formula (6.49) for the transition probabilities $A_n^{(i,k)}$ in a two-mode squeezer.

resulting in (6.49). □

As illustrated in Figure 6.2.3, the first term of Equation (6.49) (bottom left drawing in the figure) corresponds to the stimulated annihilation of an input photon pair, while the third term (bottom right drawing in the figure) corresponds to the stimulated emission of an output photon pair (both with probability proportional to λ). The second and fourth terms correspond to the photon crossing the nonlinear medium without stimulating pair emission nor absorption (both with probability proportional to $1-\lambda$). Again, quantum interferences are responsible for a fifth term, which give rise to a suppression effect akin to the HOM effect. Note that Equation (6.49) reduces to the recurrence $A_n^{(i,0)} = (1-\lambda)\, A_{n-1}^{(i-1,0)} + \lambda\, A_{n-1}^{(i,0)}$ obtained for $k = 0$ in relation with majorization in an amplifier channel [81].

Remarkably, when $i = k = 1$ and $\lambda = 1/2$, we observe a complete extinction of the out-

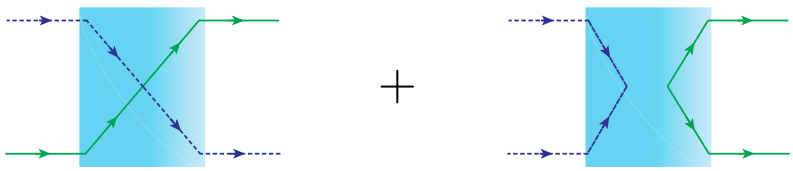

Figure 6.2.4: Parametric amplification with gain 2 ($\lambda = 1/2$) exhibits an analogous Hong-Ou-Mandel destructive interference effect between both photons crossing a nonlinear medium with amplitude $\sqrt{1-\lambda} \times \sqrt{1-\lambda}$ and the stimulated annihilation of the input photon pair accompanied by the stimulated emission of a distinct output pair with amplitude $-\sqrt{\lambda} \times \sqrt{\lambda}$. The indistinguishability between the input and output photon pairs is responsible for the suppression effect $\langle 1,1|\, U_{1/2}^{\mathrm{TMS}}\, |1,1\rangle = 0$.

put state $|1\rangle\,|1\rangle$ due to destructive interference, as explained in Figure 6.2.4. This heretofore unknown effect is a direct consequence of quantum indistinguishability. The probability amplitude that a photon goes through is $a_1^{(1,0)} = a_0^{(0,1)} \propto \sqrt{1-\lambda}$ for each of the two photons. In contrast, the probability amplitude that an input photon pair is annihilated is $a_0^{(1,1)} \propto -\sqrt{\lambda}$, while a new photon pair is created with probability amplitude $a_1^{(0,0)} \propto \sqrt{\lambda}$. If the output photons were distinguishable from the input photons, the probabilities of the two scenarios would add up, but quantum indistinguishability requires us to add amplitudes, leading to cancellation when $\lambda = 1/2$.

6.3 Transition probabilities of N-mode passive Gaussian unitaries

As already hinted at in Chapter 5, we are now going to exploit the generating function computed for the transition probabilities in the case of an N-mode passive Gaussian unitary in order to characterise quantum interferences taking place in such a scheme. We could try to generalise Theorem 27, which provides quite a broad relation in the case of 2-mode passive unitaries. However, we are going to turn to Corollary 4, since it is easier to handle, while already providing us with an elegant description of photon interferences in a beam splitter. We therefore generalise it hereafter.

Consider an N-mode passive interferometer whose effect on the bosonic field operators in phase space is characterised by the orthogonal matrix $\mathbf{U}$ of dimension N, i.e,

$$\hat{\mathbf{a}} \to \mathbf{U}\hat{\mathbf{a}}, \quad \hat{\mathbf{a}} = (\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_N)\,. \tag{6.52}$$

We prove the existence of a recurrence relation satisfied by the transition probabilities

$$B_{\mathbf{n}}^{(\mathbf{i})} = \left| \left(\prod_{r=1}^{N} \langle n_r| \right) U^{\mathrm{PI}} \left(\prod_{s=1}^{N} |i_s\rangle \right) \right|^2 . \tag{6.53}$$

Since the proof is quite involved, we include it in the appendix. The interested reader is referred to Appendix E.2. Define $1_N^{(a)}$ to be the N dimensional vector with ones at positions $j \in a$ and zeros everywhere else. The relation is encompassed in the following theorem.

Theorem 29. *The probabilities $B_{\mathbf{n}}^{(\mathbf{i})}, \mathbf{i} \in \mathbb{N}_+^N, \mathbf{n} \in \mathbb{N}_+^N$, obey the recurrence relation*

$$B_{\mathbf{n}}^{(\mathbf{i})} = \sum_{m=1}^{N} (-1)^{m-1} \sum_{\substack{a \in \mathcal{R}_m^{(N)} \\ \mathbf{i}_s \neq 0, s \in a}} \sum_{\substack{\beta \in \mathcal{R}_m^{(N)} \\ \mathbf{n}_r \neq 0, r \in \beta}} \left(\det\left[\mathbf{U}\left(\beta, a\right)\right]\right)^2 B_{\mathbf{n}-1_N^{(\beta)}}^{\left(\mathbf{i}-1_N^{(a)}\right)}, \tag{6.54}$$

where $\mathcal{R}_m^{(N)}$ is the set of all subsets of $\{1, 2, \ldots, N\}$ of cardinality m.

As an illustration, we particularise the previous relation for $N = 2$ in the following example,

showing that it is indeed consistent with what we did for 2 modes in the previous section.

Example 1. For $N = 2$,

$$B_{n_1,n_2}^{(i_1,i_2)} = \sum_{\substack{a\in\mathcal{R}_1^{(2)}\\ i_{a_1}\neq 0, i_{a_2}\neq 0}} \sum_{\substack{\beta\in\mathcal{R}_1^{(2)}\\ n_{\beta_1}\neq 0, n_{\beta_2}\neq 0}} \left(\det\left[\mathbf{U}\left(\beta,a\right)\right]\right)^2 B_{\mathbf{n}-1_N^{(\beta)}}^{\left(\mathbf{i}-1_N^{(a)}\right)} - \sum_{\substack{a\in\mathcal{R}_2^{(2)}\\ i_{a_1}\neq 0, i_{a_2}\neq 0}} \sum_{\substack{\beta\in\mathcal{R}_2^{(2)}\\ n_{\beta_1}\neq 0, n_{\beta_2}\neq 0}} \left(\det\left[\mathbf{U}\left(\beta,a\right)\right]\right)^2 B_{\mathbf{n}-1_N^{(\beta)}}^{\left(\mathbf{i}-1_N^{(a)}\right)}.$$

Suppose i_1, i_2, n_1, n_2 are all strictly positive. In this case,

$$B_{n_1,n_2}^{(i_1,i_2)} = \sum_{a\in\mathcal{R}_1^{(2)}} \sum_{\beta\in\mathcal{R}_1^{(2)}} \left(\det\left[\mathbf{U}\left(\beta,a\right)\right]\right)^2 B_{\mathbf{n}-1_N^{(\beta)}}^{\left(\mathbf{i}-1_N^{(a)}\right)} - \sum_{a\in\mathcal{R}_2^{(2)}} \sum_{\beta\in\mathcal{R}_2^{(2)}} \left(\det\left[\mathbf{U}\left(\beta,a\right)\right]\right)^2 B_{\mathbf{n}-1_N^{(\beta)}}^{\left(\mathbf{i}-1_N^{(a)}\right)},$$

$$\begin{aligned} B_{n_1,n_2}^{(i_1,i_2)} = {} & \left(\left(\det\left[\mathbf{U}\left(\beta,a\right)\right]\right)^2 B_{\mathbf{n}-1_N^{(\beta)}}^{\left(\mathbf{i}-1_N^{(a)}\right)}\right)\Big|_{a=\{1\},\beta=\{1\}} \\ & + \left(\left(\det\left[\mathbf{U}\left(\beta,a\right)\right]\right)^2 B_{\mathbf{n}-1_N^{(\beta)}}^{\left(\mathbf{i}-1_N^{(a)}\right)}\right)\Big|_{a=\{1\},\beta=\{2\}} \\ & + \left(\left(\det\left[\mathbf{U}\left(\beta,a\right)\right]\right)^2 B_{\mathbf{n}-1_N^{(\beta)}}^{\left(\mathbf{i}-1_N^{(a)}\right)}\right)\Big|_{a=\{2\},\beta=\{1\}} \\ & + \left(\left(\det\left[\mathbf{U}\left(\beta,a\right)\right]\right)^2 B_{\mathbf{n}-1_N^{(\beta)}}^{\left(\mathbf{i}-1_N^{(a)}\right)}\right)\Big|_{a=\{2\},\beta=\{2\}} \\ & - \left(\left(\det\left[\mathbf{U}\left(\beta,a\right)\right]\right)^2 B_{\mathbf{n}-1_N^{(\beta)}}^{\left(\mathbf{i}-1_N^{(a)}\right)}\right)\Big|_{a=\{1,2\},\beta=\{1,2\}}, \end{aligned}$$

$$\begin{aligned} B_{n_1,n_2}^{(i_1,i_2)} = {} & u_{11}^2\, B_{n_1-1,n_2}^{(i_1-1,i_2)} + u_{12}^2\, B_{n_1-1,n_2}^{(i_1,i_2-1)} \\ & + u_{21}^2\, B_{n_1,n_2-1}^{(i_1-1,i_2)} + u_{22}^2\, B_{n_1,n_2-1}^{(i_1,i_2-1)} \\ & - \left(u_{11}u_{22} - u_{12}u_{21}\right)^2\, B_{n_1-1,n_2-1}^{(i_1-1,i_2-1)}, \end{aligned}$$

$$\begin{aligned} B_{n_1,n_2}^{(i_1,i_2)} = {} & u_{11}^2\, B_{n_1-1,n_2}^{(i_1-1,i_2)} + u_{12}^2\, B_{n_1-1,n_2}^{(i_1,i_2-1)} \\ & + u_{21}^2\, B_{n_1,n_2-1}^{(i_1-1,i_2)} + u_{22}^2\, B_{n_1,n_2-1}^{(i_1,i_2-1)} \\ & - B_{n_1-1,n_2-1}^{(i_1-1,i_2-1)}. \end{aligned}$$

If we define the beam-splitter matrix as

$$\mathbf{U} = \begin{pmatrix} \sqrt{\eta} & \sqrt{1-\eta} \\ -\sqrt{1-\eta} & \sqrt{\eta} \end{pmatrix}, \tag{6.55}$$

we end up with

$$B_{n_1,n_2}^{(i_1,i_2)} = \eta\, B_{n_1-1,n_2}^{(i_1-1,i_2)} + (1-\eta)\, B_{n_1-1,n_2}^{(i_1,i_2-1)} + (1-\eta)\, B_{n_1,n_2-1}^{(i_1-1,i_2)} + \eta\, B_{n_1,n_2-1}^{(i_1,i_2-1)} - B_{n_1-1,n_2-1}^{(i_1-1,i_2-1)}, \tag{6.56}$$

which is exactly Equation (6.45).

In each of the terms $\left(\det\left[\mathbf{U}\left(\beta,a\right)\right]\right)^2 B_{\mathbf{n}-1_N^{(\beta)}}^{\left(\mathbf{i}-1_N^{(a)}\right)}$ appearing in the right-hand side of Equation (6.54), the sets a and β have the same cardinalities. As a result, the vectors $1_N^{(\beta)}$ and $1_N^{(a)}$ have the same number of zero entries, and the same number of one entries. This means that each probability $B_{\mathbf{n}-1_N^{(\beta)}}^{\left(\mathbf{i}-1_N^{(a)}\right)}$ is obtained starting from the probability $B_{\mathbf{n}}^{(\mathbf{i})}$ by removing one photon on the same amount of modes at the output of the interferometer as its input. Furthermore, a maximum of one photon only is removed from each input and output mode. Each of the corresponding possible probabilities is then multiplied by the right squared minor $\left(\det\left[\mathbf{U}\left(\beta,a\right)\right]\right)^2$ in the right-hand side of Equation (6.54), as well as the right coefficient $+1$ or -1. Like in the 2-mode case, this can be viewed as an effect of the interferences due to the indistinguishability of bosons.

The probabilities $B_{\mathbf{n}}^{(\mathbf{i})}$ involved in the N-mode passive interferometer can be understood to be quite complex objects. As a consequence, the fact that a relation as simple as Equation (6.54) exists is actually a pleasant surprise.

III

Gaussian dilatable bosonic channels

7

Gaussian-dilatable channels with passive environment

In Chapter 6, we exploited the formalism of the generating function introduced in Chapter 5 in order to investigate the transition probabilities characterising Gaussian unitary transformations, such as the probabilities $B_n^{(i,k)}$ in a beam splitter, and the probabilities $A_n^{(i,k)}$ in a two-mode squeezer. This analysis proved to be fruitful, as it allowed us to bring forth the effect of quantum interferences in such Gaussian operations. It also provided us with a first situation in which we exploited the symplectic formalism in phase space in order to describe the effect of Gaussian transformations on non-Gaussian objects, specifically the Fock states of the harmonic oscillator. In recent years, it has become clear that in order to perform many crucial quantum informations tasks, not only is it necessary to employ non-Gaussian resources in the form of non-Gaussian states, one actually often needs to consider non-Gaussian operations in general. The latter are for instance required in order to perform entanglement distillation or swapping [8–10]. One then understands the need to build a framework for the characterisation of non-Gaussian channels.

In this Chapter, we illustrate how the description of Gaussian channels introduced in Chapter 3 can be coupled with the formalism of generating functions of Chapter 5 in order to investigate specific non-Gaussian bosonic channels. We focus on maps constructed using Gaussian unitaries and specific non-Gaussian states such a Fock-passive states, which we introduce hereafter. This provides us with a description of an important class of so-called Gaussian-dilatable

channels based on Fock-passive states.

7.1 Passive-environment bosonic channels

7.1.1 Bosonic passive states

Passive quantum states often arise when studying quantum systems from a thermodynamical point of view. They can be seen as quantum states from which no work can be extracted under Hamiltonian processes, making them the most stable states among all reachable states through a unitary transformation [104]. If one defines the ergotropy $W_{\max}$, or maximal extractable work under Hamiltonian processes, of a state ρ as [105, 106]

$$W_{\max}(\rho) = \max_{U} \operatorname{Tr}\left[\hat{H}\left(\rho - U\rho U^{\dagger}\right)\right], \tag{7.1}$$

where U is unitary and $\hat{H}$ is the Hamiltonian of the system, then the passive states will be those for which the ergotropy is zero. They can be shown to be diagonal in the eigenbasis of the Hamiltonian $\hat{H}$ of the system, with non-increasing eigenvalues when the energy of the eigenvector increases [107]. Consequently, we will denote a passive state using an arrow facing down as a superscript. Since there is only one passive state for a fixed spectrum, $\rho^{\downarrow}$ will denote the passive state having the same spectrum as ρ. In the context of bosonic quantum systems, a passive state admits a spectral decomposition in the Fock basis, and can be written as

$$\rho^{\downarrow} = \sum_{i} \lambda_i^{\downarrow} |i\rangle\langle i|, \tag{7.2}$$

with $\lambda_i^{\downarrow} \geq \lambda_{i+1}^{\downarrow}$, for all $i \geq 0$.

The thermal state defined in Equation (3.29) is the most fundamental passive state, as it usually characterises a system in thermodynamical equilibrium. As such, it is the passive state with the minimal energy for a fixed entropy. Thermal states are also the only completely passive states [104]. By definition, a state ρ is completely passive if $\rho^{\otimes n}$ is passive for all $n \geq 1$. Interestingly, this implies the idea that one can always "activate" the extraction of work from a non-thermal passive state by jointly acting on it and an ancillary system, using joint unitary processes [104, 107]. As an example, suppose one has access to at least two copies of the passive state $\rho^{\downarrow}$. A system comprised of two of them (meaning one copy of the state was chosen as the ancilla described before) is not passive any more in general, allowing one to extract energy from the joint system.

It is in our interest to introduce the set of linearly independent states $\left\{P_k^{\downarrow}\right\}_{k\in\mathbb{N}_0}$, whose elements are defined as

$$P_k^{\downarrow} = \frac{1}{k+1} \sum_{i=0}^{k} |i\rangle\langle i|. \tag{7.3}$$

We label the $P_k^{\downarrow}$ as extremal-(Fock-)passive state. Any one-mode passive bosonic state $\rho^{\downarrow}$ can be written as a convex combination of such states,

$$\rho^{\downarrow} = \sum_k d_k P_k^{\downarrow}, \tag{7.4}$$

with $d_k \geq 0$ for all $k \geq 0$ and $\sum_k d_k = 1$. As a consequence, the set of passive states defines a convex polytope, whose vertices are given by the $P_k^{\downarrow}$ [106].

As a consequence of the close relationship between the concepts of energy, entropy and passive states, the latter will play a crucial role in the statement of conjectures related to the evolution of disorder in quantum bosonic channels (see Section 9.1).

7.1.2 Non-Gaussian bosonic channels

A non-Gaussian bosonic channel $\mathcal{C}$ is simply one which does not necessarily output a Gaussian state for any Gaussian input state. Gaussian channels can mathematically be characterised in an elegant way, and with a finite number of parameters. As a consequence, they are often studied using the symplectic formalism of phase space. The same cannot be said about non-Gaussian channels. In general, the phase space formalism does not introduce any simplification for the investigation of the latter. One way to define a non-Gaussian channel would be by taking either a non-Gaussian state of the environment, or a non-Gaussian unitary, or both, in the dilation of the channel. When only the state of the environment is taken to be non-Gaussian, one obtains a so-called Gaussian-dilatable channel. An example of such a map is readily found in the photon-added Gaussian channels introduced in [108]. These channels are obtained by considering Fock states in the environment of the dilation of the channel, while still taking a Gaussian unitary. The channel is shown in Figure 7.1.1. The Fock state $|k\rangle$ in the environment is produced by acting on the vacuum with a photon addition, which we represented in the figure using the operator $\left(\hat{a}^{\dagger}\right)^k$. This is obviously just a notation, since the resulting object is not a state (it is not normalised).

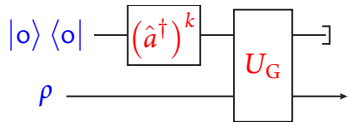

Figure 7.1.1: Representation of a photon-added channel using its dilation. The unitary U_G is Gaussian, while the environment is obtained by acting with a photon-addition on a vacuum state.

7.1.3 Definition of passive-environment bosonic channels

Passive states often arise in the context of modelling the dynamics of quantum thermodynamical systems, where some specific passive states are usually chosen as free *resources*. When constructing a resource theory, one needs to define the set of allowed state transformations [109]. This can be done by combining the following operations: composing the state with a fixed environment (in other words, a bath), acting on the joint resulting state with a unitary, which is usually chosen to conserve the energy, and finally, discarding the environment. The latter is usually chosen to be thermal, as it is a reasonable physical assumption about its state. Still, one could choose to construct a simplest, less realistic model, by choosing some maximally mixed state as an environment. When doing so, one ends up with the resource theory of so-called noisy operations, which have the form

$$\mathcal{C}_{\mathrm{NO}}\left[\rho_{\mathrm{S}}\right] = \mathrm{Tr}_{\mathrm{E}}\left[U_{\mathrm{SE}}\left(\rho_{\mathrm{S}} \otimes \frac{\mathbb{I}_{\mathrm{E}}}{n_{\mathrm{E}}}\right) U_{\mathrm{SE}}^{\dagger}\right], \tag{7.5}$$

where $\mathbb{I}_{\mathrm{E}}$ is the identity defined on the environment of dimension n_{E} and U_{AE} is an energy conserving unitary acting on both the system and the environment. As already mentioned, a more realistic model can be obtained by choosing a thermal state τ_{E} as an environment, constructing the so-called thermal operations,

$$\mathcal{C}_{\mathrm{TO}}\left[\rho_{\mathrm{S}}\right] = \mathrm{Tr}_{\mathrm{E}}\left[U_{\mathrm{SE}}\left(\rho_{\mathrm{S}} \otimes \tau_{\mathrm{E}}\right) U_{\mathrm{SE}}^{\dagger}\right]. \tag{7.6}$$

In each of these two situations, the environment is a passive state. An interesting intermediate case consists in choosing a general passive state in the environment of the dilation of the channel. This actually generalises the two resource theories we introduced. Since we are interested in bosonic systems, we obviously choose the state to be passive in the Fock basis. Since the beam splitter is the realisation of the Gaussian energy-conserving unitary acting on bosonic systems, we make use of it in order to make our system interact with the passive environment. As depicted in Figure 7.1.2, we end up with an operation we label as passive-environment channel, of the form

$$\mathcal{B}_{\eta}^{\downarrow}\left[\rho_{\mathrm{S}}\right] = \mathrm{Tr}_{\mathrm{E}}\left[U_{\eta}^{\mathrm{BS}}\left(\rho_{\mathrm{S}} \otimes \sigma_{\mathrm{E}}^{\downarrow}\right) U_{\eta}^{\mathrm{BS}\dagger}\right], \tag{7.7}$$

where $\sigma_{\mathrm{E}}^{\downarrow}$ is some passive state. As we have already seen, the two-mode squeezing unitary plays a dominant role in the study of bosonic systems, for instance in the definition of Gaussian chan-

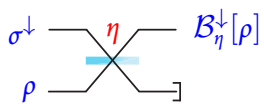

Figure 7.1.2: Passive-environment channel involving a beam splitter.

nels which cannot be obtained by means of a simple beam splitter, *e.g.* amplifier channels. As a consequence, we choose to generalise the map of equation (7.7) by considering a two-mode squeezer as well, *i.e.*,

$$\mathcal{C}^{\downarrow}\left[\rho_{\mathrm{S}}\right] = \mathrm{Tr}_{\mathrm{E}}\left[U_{\mathrm{G}}\left(\rho_{\mathrm{S}} \otimes \sigma_{\mathrm{E}}^{\downarrow}\right) U_{\mathrm{G}}^{\dagger}\right], \tag{7.8}$$

where U_{G} is either a beam-splitting unitary or a two-mode squeezing unitary. Since any passive state $\sigma_{\mathrm{E}}^{\downarrow}$ can be decomposed by means of extremal-passive states through Equation (7.4), we also define the so called extremal passive-environment channels (or simply extremal-passive channels), as

$$\mathcal{B}_{\eta}^{[k]}\left[\bullet\right] = \mathrm{Tr}_{2}\left[U_{\eta}^{\mathrm{BS}}\left(\bullet \otimes P_{k}^{\downarrow}\right) U_{\eta}^{\mathrm{BS}\dagger}\right], \tag{7.9}$$

in the case of a beam splitter, and

$$\mathcal{A}_{G}^{[k]}\left[\bullet\right] = \mathrm{Tr}_{2}\left[U_{\lambda}^{\mathrm{TMS}}\left(\bullet \otimes P_{k}^{\downarrow}\right) U_{\lambda}^{\mathrm{TMS}\dagger}\right], \quad \lambda = \frac{G-1}{G} \tag{7.10}$$

in the case of a two-mode squeezer. Note that the gain G which appears in the definition of $\mathcal{A}_{G}^{[k]}$ has nothing to do with G in Equation (7.8) which characterises the fact that the unitary is Gaussian. Compared to the notations $\mathcal{B}_{\eta}^{(\varepsilon)}$ of a lossy channel and $\mathcal{A}_{G}^{(\varepsilon)}$ of an amplifier channel (introduced in Chapter 3), the superscripts are now surrounded by square brackets, as a reminder of the fact that it is an integer number characterising the extremal-passive state of the environment. Any passive-environment channel of the form of Equation (7.8) can be written as a convex mixture of channels $\mathcal{B}_{\eta}^{[k]}$ or $\mathcal{A}_{G}^{[k]}$. As a consequence, it is often sufficient to investigate the latter when wanting to study some property of general passive-environment channels.

7.2 Gaussian decomposition of passive-environment bosonic channels

A surprising application of the generating function resides in the characterisation of passive-environment channels. The extremal-passive channels $\mathcal{B}_{\eta}^{[k]}$ and $\mathcal{A}_{G}^{[k]}$ both depend on an index k which characterises the rank of the extremal-passive state of their environments. Since k is a non-negative integer, taking the generating function of either of $\mathcal{B}_{\eta}^{[k]}$ and $\mathcal{A}_{G}^{[k]}$ is a well-defined operation. Let us denote by $\mathcal{C}_{\kappa}^{[k]}$ any of the two types of extremal passive channels, such that

$$\mathcal{C}_{\kappa}^{[k]} = \mathcal{B}_{\kappa}^{[k]}, \quad \text{if } \kappa < 1, \tag{7.11}$$

and

$$\mathcal{C}_{\kappa}^{[k]} = \mathcal{A}_{\kappa}^{[k]}, \quad \text{if } \kappa > 1, \tag{7.12}$$

while $\kappa = 1$ corresponds to an identity operation. Lets us compute the generating function of

the map $\tilde{\mathcal{C}}_\kappa^{[k]}[\rho] = (k+1)\mathcal{C}_\kappa^{[k]}[\rho]$, which is simply given by

$$\sum_k \tilde{\mathcal{C}}_\kappa^{[k]}[\rho] y^k, \quad y \in \mathbb{R}. \tag{7.13}$$

Now, remember that Gaussian channels can be shown to possess a semi-group structure, and can be represented using a map of the form of Equation (3.112). As such, they can be characterised by a time $t \geq 0$ related to the parameter κ which represents the transmittance $\eta = e^{-t}$ of beam splitter when it is less than one, or the gain $G = e^t$ in a two-mode squeezer if it is greater than one. Since this representation is quite elegant, we choose to introduce a time t in the present characterisation of passive-environment channels, setting $t = |\ln \kappa|$. As a consequence, we parametrise the generating function using the time t, and denote it as

$$\mathcal{N}(t, y)[\rho] = \sum_k \tilde{\mathcal{C}}_\kappa^{[k]}[\rho] y^k, \quad t = |\ln \kappa|, \; y \in \mathbb{R}. \tag{7.14}$$

In this context, the extremal-passive channels can simply be retrieved through

$$\tilde{\mathcal{C}}_\kappa^{[k]}[\rho] = \frac{1}{k!} \frac{\partial^k}{\partial y^k} \mathcal{N}(t, y)[\rho] \Big|_{y=0}, \quad t = |\ln \kappa|, \tag{7.15}$$

the derivative being well-defined in this case. Using the definition of extremal-passive channels, the map $\mathcal{N}(t, y)$ can be rewritten

$$\mathcal{N}(t, y)[\rho] = \sum_k \mathrm{Tr}_2 \left[U_\kappa^G (\rho \otimes (k+1) P_k^\downarrow) U_\kappa^{G\dagger} \right] y^k, \quad t = |\ln \kappa|, \tag{7.16}$$

$$\mathcal{N}(t, y)[\rho] = \mathrm{Tr}_2 \left[U_\kappa^G (\rho \otimes \sum_k (k+1) P_k^\downarrow y^k) U_\kappa^{G\dagger} \right], \quad t = |\ln \kappa|, \tag{7.17}$$

where U_κ^G is a beam splitter of transmittance κ if $\kappa < 1$ and a two-mode squeezer of gain κ if $\kappa > 1$. Now, if $y \in [0, 1)$, it can easily be shown that

$$\sum_k (k+1) P_k^\downarrow y^k = (1-y)^{-2} \tau_y, \quad y \in [0, 1), \tag{7.18}$$

where τ_y is a Gaussian thermal state of parameter y. Using this, we can rewrite the map $\mathcal{N}(t, y)$ as

$$\mathcal{N}(t, y)[\rho] = (1-y)^{-2} \mathrm{Tr}_2 \left[U_\kappa^G (\rho \otimes \tau_y) U_\kappa^{G\dagger} \right], \quad t = |\ln \kappa|. \tag{7.19}$$

We introduce the Gaussian channel

$$\mathcal{G}_\kappa^{(y)}[\bullet] = \begin{cases} \mathcal{B}_\kappa^{(y)}[\bullet], & \kappa < 1, \\ \mathcal{A}_\kappa^{(y)}[\bullet], & \kappa > 1, \end{cases} \tag{7.20}$$

so that we can rewrite $\mathcal{N}(t,y)$ as

$$\mathcal{N}(t,y)[\rho] = (1-y)^{-2}\mathcal{G}_\kappa^{(y)}[\rho], \quad t = |\ln\kappa|. \tag{7.21}$$

The thermal Gaussian state τ_y can be obtained starting from the vacuum state by applying a classical-noise channel with the right amount of noise as

$$\tau_y = e^{N_y\mathcal{L}_0}\left[|0\rangle\langle 0|\right], \tag{7.22}$$

where $N_y = y/(1-y)$, and $\mathcal{L}_0$ is the Lindbladian defined in Equation (3.120), so that

$$\mathcal{G}_\kappa^{(y)}[\rho] = \mathrm{Tr}_2\left[U_\kappa^G(\rho\otimes e^{N_y\mathcal{L}_0}\left[|0\rangle\langle 0|\right])U_\kappa^{G\dagger}\right]. \tag{7.23}$$

It can be shown that the classical-noise channel acting before the beam splitter can actually be replaced by another classical-noise channel acting after the beam splitter [91], with the right amount of noise,

$$\mathcal{G}_\kappa^{(y)}[\rho] = e^{M_y(t)\mathcal{L}_0}\left[\mathrm{Tr}_2\left[U_\kappa^G(\rho\otimes|0\rangle\langle 0|)U_\kappa^{G\dagger}\right]\right], \tag{7.24}$$

where $M_y(t) = |\kappa-1|N_y$, with $t = |\ln\kappa|$, so that

$$\mathcal{G}_\kappa^{(y)}[\rho] = e^{M_y(t)\mathcal{L}_0}\left[\mathcal{G}_\kappa[\rho]\right]. \tag{7.25}$$

where $\mathcal{G}_\kappa[\bullet] := \mathcal{G}_\kappa^{(0)}[\bullet]$. We now develop the exponential as a series,

$$\mathcal{G}_\kappa^{(y)}[\rho] = \sum_{n=0}^{\infty}\frac{1}{n!}\left(M_y(t)\right)^n\mathcal{L}_0^n\left[\mathcal{G}_\kappa[\rho]\right], \tag{7.26}$$

meaning that the map $\mathcal{N}(t,y)$ can be written

$$\mathcal{N}(t,y)[\rho] = (1-y)^{-2}\sum_{n=0}^{\infty}\frac{1}{n!}\left(M_y(t)\right)^n\mathcal{L}_0^n\left[\mathcal{G}_\kappa[\rho]\right], \quad t = |\ln\kappa|. \tag{7.27}$$

Our goal is now to derive an interesting form for the extremal-passive channels, by using Equation (7.15). In order to do so, we compute the derivatives in y explicitly. We have

$$\frac{\partial^k}{\partial y^k}\mathcal{N}(t,y)[\rho] = \sum_{n=0}^{\infty}\frac{1}{n!}\frac{\partial^k}{\partial y^k}\left[(1-y)^{-2}M_y^n(t)\right]\mathcal{L}_0^n\left[\mathcal{G}_\kappa[\rho]\right], \quad t = |\ln\kappa|, \tag{7.28}$$

meaning that we need to compute

$$\frac{\partial^k}{\partial y^k}\left[(1-y)^{-2}M_y^n(t)\right] = \sum_{m=0}^{k}\binom{k}{m}\frac{\partial^{K-m}}{\partial y^{K-m}}\left[(1-y)^{-2}\right]\frac{\partial^m}{\partial y^m}\left[M_y^n(t)\right], \quad t = |\ln\kappa|,$$

$$\frac{\partial^k}{\partial y^k}\left[(1-y)^{-2}M_y^n(t)\right] = \sum_{m=0}^{k}\binom{k}{m}(k-m+1)!(1-y)^{-(K-m+2)}\frac{\partial^m}{\partial y^m}\left[M_y^n(t)\right], \quad t = |\ln\kappa|.$$

Since we only need the derivatives at $y = 0$, we need only calculate the object

$$\frac{\partial^k}{\partial y^k}\left[(1-y)^{-2}M_y^n(t)\right]\Big|_{y=0} = \sum_{m=0}^{k}\binom{k}{m}(k-m+1)!\,\frac{\partial^m}{\partial y^m}\left[M_y^n(t)\right]\Big|_{y=0}, \quad t = |\ln\kappa|. \tag{7.29}$$

The function M_y is defined as

$$M_y(t) = |\kappa-1|N_y = |\kappa-1|\frac{y}{1-y}, \quad t = |\ln\kappa|. \tag{7.30}$$

so that

$$\frac{\partial^m}{\partial y^m}M_y^n(t) = |\kappa-1|^n\frac{\partial^m}{\partial y^m}\left(\frac{y}{1-y}\right)^n. \tag{7.31}$$

Some simple mathematical steps lead to

$$\begin{aligned}\frac{\partial^m}{\partial y^m}M_y^n(t) &= |\kappa-1|^n\frac{\partial^m}{\partial y^m}\left[y^n(1-y)^{-n}\right]\\ &= |\kappa-1|^n\sum_{l=0}^{m}\binom{m}{l}\frac{\partial^{m-l}}{\partial y^{m-l}}y^n\frac{\partial^l}{\partial y^l}(1-y)^{-n}\\ &= |\kappa-1|^n\sum_{l=0}^{m}\binom{m}{l}\frac{n!}{(n-m+l)!}y^{n-m+l}\frac{(n+l-1)!}{(n-1)!}(1-y)^{-n-l},\end{aligned}$$

so that

$$\frac{\partial^m}{\partial y^m}M_y^n(t)\Big|_{y=0} = |\kappa-1|^n\binom{m}{m-n}n!\frac{(m-1)!}{(n-1)!} = |\kappa-1|^n m!\binom{m-1}{n-1}, \tag{7.32}$$

and

$$\frac{\partial^k}{\partial y^k}\left[(1-y)^{-2}M_y^n(t)\right]\Big|_{y=0} = |\kappa-1|^n\sum_{m=n}^{k}\binom{k}{m}(k-m+1)!m!\binom{m-1}{n-1}. \tag{7.33}$$

The derivatives in y of the map $\mathcal{N}(t,y)$ can now be computed to be

$$\begin{aligned}\frac{\partial^k}{\partial y^k}\mathcal{N}(t,y)[\rho]\Big|_{y=0} &= \sum_{n=0}^{\infty}\frac{1}{n!}\frac{\partial^k}{\partial y^k}\left[(1-y)^{-2}M_y^n(t)\right]\Big|_{y=0}\mathcal{L}_0^n\left[\mathcal{G}_\kappa[\rho]\right]\\ &= \sum_{n=0}^{\infty}\frac{1}{n!}|\kappa-1|^n\sum_{m=n}^{k}\binom{k}{m}(k-m+1)!m!\binom{m-1}{n-1}\mathcal{L}_0^n\left[\mathcal{G}_\kappa[\rho]\right],\end{aligned}$$

for $t = |\ln \kappa|$, meaning that the maps $\tilde{\mathcal{C}}_\kappa^{[k]}$ are given by

$$\tilde{\mathcal{C}}_\kappa^{[k]}[\rho] = \frac{1}{k!}\sum_{n=0}^{\infty}\frac{1}{n!}|\kappa-1|^n\sum_{m=n}^{k}\binom{k}{m}(k-m+1)!m!\binom{m-1}{n-1}\mathcal{L}_0^n\left[\mathcal{G}_\kappa[\rho]\right], \tag{7.34}$$

so that the extremal-passive channels can be written as

$$\mathcal{C}_\kappa^{[k]}[\rho] = \frac{1}{k+1}\sum_{n=0}^{\infty}\frac{1}{n!}|\kappa-1|^n\sum_{m=n}^{k}(k-m+1)\binom{m-1}{n-1}\mathcal{L}_0^n\left[\mathcal{G}_\kappa[\rho]\right]. \tag{7.35}$$

In the last equation, we define the quantity

$$\begin{aligned} c_n^{(k)} &= \sum_{m=n}^{k}\frac{(k-m+1)}{n!(k+1)}\binom{m-1}{n-1} \\ &= \frac{(k-n+1)}{n(n+1)!}\binom{k}{n-1}, \end{aligned}$$

$$\begin{aligned} c_n^{(k)} &= \frac{(k-n+1)}{n(n+1)!}\frac{k!}{(n-1)!(k-n+1)!} \\ &= \frac{k!}{n!(n+1)!}\frac{1}{(k-n)!}, \end{aligned}$$

and end up with

$$c_n^{(k)} = \frac{1}{(n+1)!}\binom{k}{n}. \tag{7.36}$$

Since we have

$$\mathcal{C}_\kappa^{[k]}[\rho] = \sum_{n=0}^{k}|\kappa-1|^n c_n^{(k)}\mathcal{L}_0^n\left[\mathcal{G}_\kappa[\rho]\right], \tag{7.37}$$

we can finally define the zero-trace map

$$\mathfrak{C}_\kappa^{[k]} = \sum_{n=0}^{k}|\kappa-1|^n c_n^{(k)}\mathcal{L}_0^n, \tag{7.38}$$

which allows us to write

$$\mathcal{C}_\kappa^{[k]}[\rho] = \mathfrak{C}_\kappa^{[k]}\circ\mathcal{G}_\kappa[\rho]. \tag{7.39}$$

As an example, the maps of Equation (7.38) for the few first values of k can readily be found to satisfy

$$\begin{cases} \mathfrak{C}_\eta^{[0]} = \mathbb{1}, \\ \mathfrak{C}_\eta^{[1]} = \mathbb{1} + \frac{1}{2}|\kappa-1|\mathcal{L}_0, \\ \mathfrak{C}_\eta^{[2]} = \mathbb{1} + |\kappa-1|\mathcal{L}_0 + \frac{1}{6}|\kappa-1|^2\mathcal{L}_0^2. \end{cases} \tag{7.40}$$

For completeness, we give the following two lemmas in order to stress the difference between the passive lossy channels and the passive amplifying channels.

Lemma 8. *In the case of a beam-splitter unitary, the corresponding extremal passive channel can be decomposed as*

$$\mathcal{B}_{\eta}^{[k]}[\bullet] = \mathfrak{B}_{\eta}^{[k]} \circ \mathcal{B}_{\eta}[\bullet], \tag{7.41}$$

where

$$\mathfrak{B}_{\eta}^{[k]}[\bullet] = \sum_{n=0}^{k} (1-\eta)^n c_n^{(k)} \mathcal{L}_0^n[\bullet], \tag{7.42}$$

and the Lindbladian is defined as

$$\mathcal{L}_0[\bullet] = \hat{a}^\dagger \bullet a - \frac{1}{2} a\hat{a}^\dagger \bullet - \frac{1}{2} \bullet a\hat{a}^\dagger + a \bullet \hat{a}^\dagger - \frac{1}{2}\hat{a}^\dagger a \bullet - \frac{1}{2} \bullet \hat{a}^\dagger a. \tag{7.43}$$

Lemma 9. *In the case of a two-mode squeezer unitary, the corresponding extremal passive channel can be decomposed as*

$$\mathcal{A}_{G}^{[k]}[\bullet] = \mathfrak{A}_{G}^{[k]} \circ \mathcal{A}_{G}[\bullet], \tag{7.44}$$

where

$$\mathfrak{A}_{G}^{[k]}[\bullet] = \sum_{n=0}^{k} (G-1)^n c_n^{(k)} \mathcal{L}_0^n[\bullet], \tag{7.45}$$

and the Lindbladian is defined as

$$\mathcal{L}_0[\bullet] = \hat{a}^\dagger \bullet a - \frac{1}{2} a\hat{a}^\dagger \bullet - \frac{1}{2} \bullet a\hat{a}^\dagger + a \bullet \hat{a}^\dagger - \frac{1}{2}\hat{a}^\dagger a \bullet - \frac{1}{2} \bullet \hat{a}^\dagger a. \tag{7.46}$$

Note that this decomposition can be derived in the case of photon-added Gaussian-dilatable channels as well.

7.3 Dual map of passive-environment bosonic channels

In order to illustrate the usefulness of the decomposition of extremal-passive channels derived in Section 7.2, we show some interesting properties of these channels, namely that they verify the same duality relation as the one exhibited by Gaussian channels. We will actually exploit this relation later on, in the context of the theory of Fock-majorization, which we introduce in Chapter 8. We begin by proving the following lemma concerning the map $\mathfrak{C}_{\kappa}^{[k]}$.

Lemma 10. *The map $\mathfrak{C}_{\kappa}^{[k]}$ is self-adjoint for all $k \in \mathbb{N}$, in the sense that $\mathfrak{C}_{\kappa}^{[k]\dagger} = \mathfrak{C}_{\kappa}^{[k]}$, where $\mathfrak{C}_{\kappa}^{[k]\dagger}$ is the dual (adjoint) map of $\mathfrak{C}_{\kappa}^{[k]}$.*

Proof. Using the cyclic property of the trace, we have that

$$\begin{aligned}\mathrm{Tr}\left[Y\mathcal{L}_-[X]\right] &= \mathrm{Tr}\left[Y\hat{a}X\hat{a}^\dagger - \frac{1}{2}Y\hat{a}^\dagger\hat{a}X - \frac{1}{2}YX\hat{a}^\dagger\hat{a}\right]\\ &= \mathrm{Tr}\left[\hat{a}^\dagger Y\hat{a}X - \frac{1}{2}Y\hat{a}^\dagger\hat{a}X - \frac{1}{2}\hat{a}^\dagger\hat{a}YX\right]\\ &= \mathrm{Tr}\left[\hat{a}^\dagger Y\hat{a}X - \frac{1}{2}\hat{a}\hat{a}^\dagger YX - \frac{1}{2}Y\hat{a}\hat{a}^\dagger X + XY\right]\\ &= \mathrm{Tr}\left[(\mathcal{L}_+[Y] + \mathbb{1})X\right],\end{aligned}$$

where we used the bosonic commutation relation, so that

$$\mathcal{L}_o^\dagger = \mathcal{L}_+^\dagger + \mathcal{L}_+ + \mathbb{1} = (\mathcal{L}_+ + \mathbb{1})^\dagger + \mathcal{L}_+ = \mathcal{L}_- + \mathcal{L}_+ = \mathcal{L}_o. \tag{7.47}$$

As a consequence, $\mathcal{L}_o^n$ is self-adjoint for any $n \in \mathbb{N}$. Since $\mathfrak{C}_\kappa^{[k]}$ is given by a linear combination of these maps, it it self-adjoint as well for all $\kappa > 0$ and all $k \in \mathbb{N}$. □

In order to find the dual of the extremal-passive channel, we also show the following lemma concerning the Linbladians of bosonic systems.

Lemma 11. *For any $n \in \mathbb{N}$,*

$$\mathcal{L}_-^n\mathcal{L}_o = \mathcal{L}_o\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}\mathcal{L}_-^k. \tag{7.48}$$

Proof. Equation (7.48) is obviously true for $n = 0$. Suppose it is true for some arbitrary n,

$$\mathcal{L}_-^{n+1}\mathcal{L}_o = \mathcal{L}_-\mathcal{L}_-^n\mathcal{L}_o = \mathcal{L}_-\mathcal{L}_o\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}\mathcal{L}_-^k. \tag{7.49}$$

It can easily be shown that $[\mathcal{L}_o, \mathcal{L}_-] = \mathcal{L}_o$, so that $\mathcal{L}_-\mathcal{L}_o = \mathcal{L}_o\mathcal{L}_- - \mathcal{L}_o$. As a consequence,

$$\begin{aligned}\mathcal{L}_-^{n+1}\mathcal{L}_o &= \mathcal{L}_o\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}\mathcal{L}_-^{k+1} + \mathcal{L}_o\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}\mathcal{L}_-^k\\ &= \mathcal{L}_o\sum_{k=1}^{n+1}(-1)^{n-k+1}\binom{n}{k-1}\mathcal{L}_-^k + \mathcal{L}_o\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}\mathcal{L}_-^k\\ &= \mathcal{L}_o\mathcal{L}_-^{n+1} + \mathcal{L}_o\sum_{k=1}^{n}(-1)^{n-k+1}\left[\binom{n}{k-1} + \binom{n}{k}\right]\mathcal{L}_-^k + (-1)^{n+1}\mathcal{L}_o.\end{aligned}$$

Using the recurrence relation satisfied by binomial coefficients, we get

$$\begin{aligned}\mathcal{L}_-^{n+1}\mathcal{L}_0 &= \mathcal{L}_0\mathcal{L}_-^{n+1} + \mathcal{L}_0\sum_{k=1}^{n}(-1)^{n-k+1}\binom{n+1}{k}\mathcal{L}_-^k + (-1)^{n+1}\mathcal{L}_0 \\ &= \mathcal{L}_0\sum_{k=0}^{n+1}(-1)^{n-k+1}\binom{n+1}{k}\mathcal{L}_-^k.\end{aligned}$$

A recursive argument concludes the proof. □

We are now in position to prove the following theorem concerning the duality of extremal-passive channels.

Theorem 30. *The extremal-passive channels are connected through the duality relation*

$$\mathcal{A}_G^{[k]\dagger} = \eta\mathcal{B}_\eta^{[k]}, \quad \eta = \frac{1}{G}, \tag{7.50}$$

for all $k \in \mathbb{N}$ and $G > 1$.

Proof. Consider any map $\mathcal{C} = \mathcal{C}_2 \circ \mathcal{C}_1$. We have that

$$\begin{aligned}\mathrm{Tr}\left[\mathcal{C}[X]Y\right] &= \mathrm{Tr}\left[\left(\mathcal{C}_2 \circ \mathcal{C}_1\right)[X]Y\right] \\ &= \mathrm{Tr}\left[\mathcal{C}_1[X]\mathcal{C}_2^\dagger[Y]\right] \\ &= \mathrm{Tr}\left[X\left(\mathcal{C}_1^\dagger \circ \mathcal{C}_2^\dagger\right)[Y]\right],\end{aligned}$$

so that $\mathcal{C}^\dagger = \mathcal{C}_1^\dagger \circ \mathcal{C}_2^\dagger$. As a consequence,

$$\mathcal{A}_G^{[k]\dagger} = \mathcal{A}_G^\dagger \circ \mathfrak{A}_G^{[k]\dagger} = \frac{1}{G}\mathcal{B}_{1/G} \circ \mathfrak{A}_G^{[k]}, \tag{7.51}$$

where we used Lemma 10, as well as the fact that Theorem 30 is already known to be true for $k = 0$ (Gaussian case). Now, $\mathcal{B}_{1/G} = \mathrm{e}^{t\mathcal{L}_-}$, with $G = \mathrm{e}^t$, so that

$$\mathcal{B}_{1/G} \circ \mathcal{L}_0 = \sum_{n=0}^{\infty}\frac{1}{n!}t^n\mathcal{L}_-^n \circ \mathcal{L}_0. \tag{7.52}$$

Using Lemma 11, we get

$$\mathcal{B}_{1/G} \circ \mathcal{L}_0 = \sum_{n=0}^{\infty}\frac{1}{n!}t^n\mathcal{L}_0\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}\mathcal{L}_-^k = \mathcal{L}_0\sum_{k=0}^{\infty}(-1)^k\left(\sum_{n=k}^{\infty}\frac{1}{n!}t^n(-1)^n\binom{n}{k}\right)\mathcal{L}_-^k. \tag{7.53}$$

Since

$$\sum_{n=k}^{\infty}(-1)^n\binom{n}{k}\frac{t^n}{n!} = (-1)^k\mathrm{e}^{-t}\frac{t^k}{k!}, \tag{7.54}$$

we end up with

$$\mathcal{B}_{1/G} \circ \mathcal{L}_0 = \mathrm{e}^{-t} \mathcal{L}_0 \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathcal{L}_-^k = \frac{1}{G} \mathcal{L}_0 \mathrm{e}^{t\mathcal{L}_-} = \frac{1}{G} \mathcal{L}_0 \mathcal{B}_{1/G}, \tag{7.55}$$

and, more generally,

$$\mathcal{B}_{1/G} \circ \mathcal{L}_0^n = \frac{1}{G^n} \mathcal{L}_0^n \mathcal{B}_{1/G}. \tag{7.56}$$

As a consequence,

$$\frac{1}{G} \mathcal{B}_{1/G} \circ \mathfrak{A}_G^{[k]} = \frac{1}{G} \sum_{n=0}^{k} \left(1 - \frac{1}{G}\right)^n c_n^{(k)} \mathcal{L}_0^n \circ \mathcal{B}_{1/G} = \frac{1}{G} \mathfrak{B}_{1/G}^{[k]} \circ \mathcal{B}_{1/G}, \tag{7.57}$$

so that

$$\mathcal{A}_G^{[k]\dagger} = \frac{1}{G} \mathcal{B}_{1/G}^{[k]}. \tag{7.58}$$

This concludes the proof. □

This result generalises the fact that the adjoint of a pure-loss channel is proportional to a quantum-limited amplifier. This property happens to be very useful in the case of Gaussian channels, as it allows one to focus on either of the two quantum-limited Gaussian channels when proving specific properties, before extending the latter using the duality. As we will show later, it will also prove useful in the context of this work, in the study of majorization relations in non-Gaussian passive-environment channels.

8

Fock-majorization relations

The algebraic theory of majorization has long been known to play a prominent role in quantum information theory. Its relationship with the von Neumann entropy, *i.e.*,

$$\rho \succ \sigma \quad \Rightarrow \quad S(\rho) \leq S(\sigma), \tag{8.1}$$

makes it a powerful tool for the study of quantum systems. Indeed, there are situations in which the derivation of properties of von Neumann's functional becomes quite involved, specifically as a result of the logarithm appearing in its expression. The proof of the entropy photon-number inequality for a fixed Gaussian environment, *i.e.* Equation (4.40), gives rise to such a situation in which one has to compute the entropy at the output of a Gaussian channel conditional on the input. Majorization theory enters the picture whenever one can prove the existence of a majorization relation, implying the von Neumann inequality in question. In the case of the EPnI with a fixed Gaussian environment, one can precisely find such a majorization relation (see Section 9.1 for the details). Obviously, this kind of method cannot always be exploited, since the existence of an entropic inequality does not necessarily guaranty the existence of an underlying majorization criterion. Nevertheless, majorization theory proves to be quite effective in the remaining (various) cases.

The role of majorization in quantum mechanics goes beyond the establishment of proofs of entropic inequalities, specifically in the framework of entanglement theory. This can be witnessed through the interconversion of two pure entangled states using local operations and

classical communication, which is only possible whenever the subsystems of the pure states in question satisfy some majorization relation (see Theorem 16), *i.e.*,

$$|\varphi\rangle_{AB} \xrightarrow{\text{LOCC}} |\psi\rangle_{AB} \quad \Leftrightarrow \quad \rho_\psi \succ \rho_\varphi, \quad \rho_\psi = \text{Tr}_B \left[|\psi\rangle_{AB} \langle\psi| \right]. \tag{8.2}$$

The results provides us with a paradigm in which the investigation of majorization criteria turns out to be crucial for the study of a purely quantum resource such as entanglement.

In this chapter, we introduce the notion of Fock-majorization, denoted as $\succ_F$, which induces a novel (pre)order relation between quantum states and is closely connected to the theory of majorization, as one may have anticipated. As a result of its particular relation with the eigenstates of the Hamiltonian of the quantum harmonic oscillator, Fock-majorization plays an important role in the framework of bosonic systems. We begin by defining the notion of Fock-majorization in Section 8.1, explaining along the way that it holds an interesting connection with the concept of energy. In Section 8.2, we prove several properties of Fock-majorization. By doing so, we first show that the latter can be interpreted as a relation indicating the existence of an *amplifying* map between two quantum states. We then analyse the behaviour of a Fock-majorization relation between two quantum states when they evolve through a general quantum channel, providing a criterion which assert the preservation of the relation after the effect of the channel.

8.1 Definition of the Fock-majorization relation

In analogy with the definition of majorization given in Lemma 3, we introduce the concept of Fock-majorization, which compares two states in terms of their diagonal elements in the eigenbasis of the Hamiltonian of the harmonic oscillator.

Definition 31 (Fock-majorization)**.** *Let ρ and σ be two density matrices. We say that ρ Fock-majorizes σ, denoted as $\rho \succ_F \sigma$, whenever*

$$\text{Tr}\left[Q_n^{\downarrow} \rho\right] \geq \text{Tr}\left[Q_n^{\downarrow} \sigma\right], \quad \forall\, n \geq 0, \tag{8.3}$$

where $Q_n^{\downarrow} = \sum_{i=0}^{n} |i\rangle \langle i|$.

This yields a distinct (pre)order relation in state space, which only depends on the diagonal elements of ρ and σ (or their eigenvalues if the states are Fock-diagonal). In contrast with regular majorization, the diagonal elements are not ordered by decreasing values, but instead by increasing photon number. Such a relaxed definition of majorization without prior sorting is sometimes called "unordered majorization" [5]; it is useful only when there exists a natural way of ordering the elements (in the present case, it is the energy). To our knowledge, such a notion of Fock-majorization has never been defined nor exploited in the context of Gaussian bosonic channels or more generally continuous-variable quantum information.

An interesting feature of Fock-majorization is that if $\rho \succ_{\mathrm{F}} \sigma$ and $\sigma \succ_{\mathrm{F}} \rho$ both hold, then $\mathrm{diag}(\rho) = \mathrm{diag}(\sigma)$. By comparison, for regular majorization, if $\rho \succ \sigma$ and $\sigma \succ \rho$ both hold, then the states are equivalent (isospectral). The first clear connection between Fock-majorization and energy can be witnessed in the fact that any two Fock states $|n\rangle$ and $|m\rangle$ satisfy the Fock-majorization relation $|n\rangle\langle n| \succ_{\mathrm{F}} |m\rangle\langle m|$ only if $n \leq m$, whereas they are always equivalent (isospectral) in terms of usual majorization. Interestingly, Fock-majorization implies a more general energy order relation between comparable states, namely

$$\rho \succ_{\mathrm{F}} \sigma \quad \Rightarrow \quad \mathrm{Tr}(\rho\,\hat{n}) \leq \mathrm{Tr}(\sigma\,\hat{n}), \tag{8.4}$$

where $\hat{n} = \hat{a}^\dagger \hat{a}$ is the number operator. Although Equation (8.4) holds in general, only diagonal elements of states in the Fock basis are essential for the calculation, meaning that we can consider Fock-diagonal states. Take two such states

$$\rho = \sum_{i=0}^{N} r_i\,|i\rangle\langle i| \quad \text{and} \quad \sigma = \sum_{i=0}^{N} s_i\,|i\rangle\langle i|\,, \tag{8.5}$$

whose support is the space spanned by $\{|0\rangle, \ldots |N\rangle\}$ (if their support have unequal sizes, we take the largest size for N.) We assume that $\rho \succ_{\mathrm{F}} \sigma$, that is

$$\sum_{i=0}^{n} r_i \geq \sum_{i=0}^{n} s_i \quad \Rightarrow \quad \sum_{i=n}^{N} r_i \leq \sum_{i=n}^{N} s_i, \qquad \forall n \text{ s.t. } 0 \leq n \leq N. \tag{8.6}$$

Summing this expression over n and interchanging the two summations gives

$$\sum_{i=1}^{N}\sum_{n=1}^{i} r_i \leq \sum_{i=1}^{N}\sum_{n=1}^{i} s_i, \tag{8.7}$$

$$\sum_{i=1}^{N} i\,r_i \leq \sum_{i=1}^{N} i\,s_i. \tag{8.8}$$

By taking the limit $N \to \infty$, we conclude that the mean energy of ρ is lower than that of σ, which proves Equation (8.4). Note that the converse of Equation (8.4) is not true.

Finally, it is straightforward to see that Fock-majorization $\rho \succ_{\mathrm{F}} \sigma$ coincides with regular majorization $\rho \succ \sigma$ over the set of passive states. Otherwise, $\rho \succ_{\mathrm{F}} \sigma$ and $\rho \succ \sigma$ are distinct order relations.

8.2 Properties of the Fock-majorization relation

Now that we introduced the notion of Fock-majorization, our goal is to prove properties that will be useful in the study of such a notion. In particular, we explain how the latter can be un-

derstood as a witness of an amplifying map connecting two quantum states.

8.2.1 Fock-majorization as an amplifying map

The connection between the two concepts of Fock-majorization and regular majorization can be further strengthened through the following properties, which can easily be shown to hold.

Property 26.

$$\rho \succ \sigma \quad \Leftrightarrow \quad \rho^{\downarrow} \succ_{\mathrm{F}} \sigma^{\downarrow}, \tag{8.9}$$

where $\rho^{\downarrow}$ ($\sigma^{\downarrow}$) is the passive state with the same spectrum as ρ (σ).

Furthermore, we obviously have $\rho^{\downarrow} \equiv \rho$ in terms of regular majorization, while $\rho^{\downarrow} \succ_{\mathrm{F}} \rho$ in terms of Fock-majorization. Using $\rho^{\downarrow} \equiv \rho$ and $\sigma^{\downarrow} \succ_{\mathrm{F}} \sigma$, this yields the following implication from regular to Fock-majorization.

Property 27.

$$\rho^{\downarrow} \succ \sigma \quad \Rightarrow \quad \rho^{\downarrow} \succ_{\mathrm{F}} \sigma. \tag{8.10}$$

Conversely, using $\rho^{\downarrow} \succ_{\mathrm{F}} \rho$ and $\sigma^{\downarrow} \equiv \sigma$, we have the following implication.

Property 28.

$$\rho \succ_{\mathrm{F}} \sigma^{\downarrow} \quad \Rightarrow \quad \rho \succ \sigma^{\downarrow}. \tag{8.11}$$

In the context of quantum thermodynamics, Fock-majorization bears some similarity to the relation called *upper-triangular majorization* introduced in [110]. There, the authors show that two states obeying such a relation can be related by a so-called *cooling map*, which happens to be a special case of the thermal operations (7.6) when the environment is set at zero temperature (it is in the vacuum state). Instead, we show that Fock-majorization can be interpreted as a relation indicating the existence of a *heating* or *amplifying* map between the two states, corresponding to a *lower-triangular majorization*, as exhibited by the following theorem.

Theorem 31. *Two states ρ and σ whose diagonal elements in the Fock basis are given by the respective vectors $\mathbf{r}$ and $\mathbf{s}$ obey $\rho \succ_{\mathrm{F}} \sigma$ if and only if there exists a stochastic lower-triangular matrix $\mathbf{L}$ such that $\mathbf{s} = \mathbf{L}\mathbf{r}$, with $L_{ij} \geq 0, \forall\, i \geq j \geq 1$, and $\sum_{i=j}^{d} L_{ij} = 1, \forall\, j \geq 1$.*

Note that the indices range from 1 to d, corresponding to Fock states ranging from $|0\rangle$ to $|d-1\rangle$. At the end of the proof, we must take the limit $d \to \infty$ resulting in the full Fock space. Interestingly, Theorem 31 is reminiscent of the property that two probability distributions related by a majorization relation can be connected through a bistochastic matrix (it is replaced here by a stochastic lower-triangular matrix).

Proof. The proof we give here is slightly simpler than the corresponding one given in [110] for the *cooling maps*. First, suppose there exists a matrix $\mathbf{L}$ satisfying the conditions of Theorem 31.

In this case, we have

$$\sum_{i=1}^{m} s_i = \sum_{i=1}^{m}\sum_{j=1}^{i} L_{ij} r_j = \sum_{j=1}^{m} r_j \sum_{i=j}^{m} L_{ij}, \qquad \forall\, m \geq 1. \tag{8.12}$$

Since

$$\sum_{i=1}^{d} L_{ij} = 1, \quad \forall\, j \geq 1, \tag{8.13}$$

we have that

$$\sum_{i=1}^{m} L_{ij} \leq 1, \quad \forall\, j \geq 1 \quad \text{and} \quad \forall\, m \geq 1, \tag{8.14}$$

(with the condition that $L_{ij} \geq 0, \forall\, i \geq j \geq 1$). This yields the relation

$$\sum_{i=1}^{m} s_i \leq \sum_{j=1}^{m} r_j, \forall\, m \geq 1, \tag{8.15}$$

which concludes the first part of the proof.

Now, suppose that $\rho \succ_{\mathrm{F}} \sigma$. We are going to construct $\mathbf{s}$ step by step starting from $\mathbf{r}$, using a succession of lower-triangular matrices. Starting with the vector $\mathbf{r} = (r_1, r_2, \ldots r_d)^{\mathrm{T}}$, we first define

$$\mathbf{w}^{(1)} = (s_1, (r_2 + r_1 - s_1), r_3, \ldots r_d)^{\mathrm{T}}. \tag{8.16}$$

Since $\mathbf{r} \succ_{\mathrm{F}} \mathbf{s}$, we have that $r_2 + r_1 - s_1 \geq s_2 \geq 0$, which means that $\mathbf{w}^{(1)}$ is a well-defined vector of probability distribution, its elements being non-negative and summing to one. Similarly, we construct

$$\mathbf{w}^{(2)} = (s_1, s_2, (r_3 + r_2 + r_1 - s_1 - s_2), r_4, \ldots r_d)^{\mathrm{T}}, \tag{8.17}$$

which also represents a well-defined probability distribution for the same reasons. More generally, we define

$$\mathbf{w}^{(k)} = \left(s_1, s_2, \ldots s_k, \left(\sum_{j=1}^{k+1} r_j - \sum_{j=1}^{k} s_j \right), r_{k+2}, \ldots r_d \right)^{\mathrm{T}}, \tag{8.18}$$

each of the $\mathbf{w}^{(k)}$ representing a well-defined probability distribution, for $k \leq d$. Furthermore, we end up with $\mathbf{w}^{(d)} = \mathbf{s}$, which we wanted to reach starting from $\mathbf{r}$. Now, we show that each $\mathbf{w}^{(k)}$ is related to the corresponding $\mathbf{w}^{(k-1)}$ through a lower-triangular matrix, which has all its diagonal elements equal to one, apart from the one on column k. In order to do this, write

$$\begin{cases} w_k^{(k)} &= \mu_1 w_k^{(k-1)}, \\ w_{k+1}^{(k)} &= \mu_2 w_k^{(k-1)} + \mu_3 w_{k+1}^{(k-1)}. \end{cases}$$

which correspond to

$$\begin{cases} s_k & = \mu_1 \left(\sum_{j=1}^{k} r_j - \sum_{j=1}^{k-1} s_j \right), \\ \sum_{j=1}^{k+1} r_j - \sum_{j=1}^{k} s_j & = \mu_2 \left(\sum_{j=1}^{k} r_j - \sum_{j=1}^{k-1} s_j \right) + \mu_3 r_{k+1}. \end{cases}$$

If we want the matrix which relates $\mathbf{w}^{(k-1)}$ to $\mathbf{w}^{(k)}$ to be column stochastic (as well as lower-triangular), we need $\mu_3 = 1$. This is also consistent with the fact that the diagonal element of column $k+1$ should be equal to one, as we chose earlier. We still need to check if both our equations are compatible with the fact that $\mu_1 \geq 0$, $\mu_2 \geq 0$, and $\mu_1 + \mu_2 = 1$. According to our first equation,

$$\mu_1 = \frac{s_k}{\sum_{j=1}^{k} r_j - \sum_{j=1}^{k-1} s_j}. \tag{8.19}$$

Since $\sum_{j=1}^{k} r_j - \sum_{j=1}^{k-1} s_j \geq s_k$, we indeed have that μ_1 is non-negative and smaller than one. The second equation tells us that

$$\mu_2 = \frac{\sum_{j=1}^{k+1} r_j - \sum_{j=1}^{k} s_j - r_{k+1}}{\sum_{j=1}^{k} r_j - \sum_{j=1}^{k-1} s_j} = \frac{\sum_{j=1}^{k} r_j - \sum_{j=1}^{k} s_j}{\sum_{j=1}^{k} r_j - \sum_{j=1}^{k-1} s_j}, \tag{8.20}$$

which is non-negative and smaller than one for the same reasons. Now, it is also trivial to see that $\mu_1 + \mu_2 = 1$, which means that the matrix relating $\mathbf{w}^{(k-1)}$ and $\mathbf{w}^{(k)}$ has indeed non-negative elements, is column stochastic, and is lower-triangular. This also means that $\mathbf{r}$ can be related to $\mathbf{s}$ using a product of lower-triangular matrices, which is also lower-triangular (and which is column-stochastic and has non-negative elements in this case, as needed). Taking the limit $d \to \infty$ ends the proof. □

Equation (8.4) exhibits a connection between Fock-majorization and energy. We can actually go a step further by generalising this property to functions of H and making it an equivalence.

Theorem 32. *Two states ρ and σ obey $\rho \succ_F \sigma$ if and only if $\mathrm{Tr}\left[f(H)\rho\right] \leq \mathrm{Tr}\left[f(H)\sigma\right]$ for any function $f : \mathbb{R} \to \mathbb{R}$ which is continuous and increasing.*

Again, this property of Fock-majorization should be compared with the one relating regular majorization and sums of convex functions of the state.

Proof. First, suppose $\rho \succ_F \sigma$. Again, denote by $\mathbf{r}$ and $\mathbf{s}$ the vectors of diagonal elements of ρ and σ in the Fock-basis, and fix their dimension to be d (at the end of the proof, we take the limit $d \to \infty$.) We need to show that, for any function $f : \mathbb{R} \to \mathbb{R}$ which is continuous and increasing,

$$\sum_{i=1}^{d} f(i) r_i - \sum_{i=1}^{d} f(i) s_i \leq 0. \tag{8.21}$$

According to Theorem 31, there exists a lower-triangular matrix $\mathbf{L}$ with non-negative elements, which is column-stochastic, and such that $\mathbf{s} = \mathbf{L}\mathbf{r}$. Thus,

$$\sum_{j=1}^{d} f(j)s_j = \sum_{j=1}^{d} f(j) \sum_{i=1}^{j} L_{ji} r_i = \sum_{i=1}^{d} r_i \sum_{j=i}^{d} f(j) L_{ji}, \tag{8.22}$$

meaning that

$$\sum_{i=1}^{d} f(i)r_i - \sum_{i=1}^{d} f(i)s_i = \sum_{i=1}^{d} r_i \left[f(i) - \sum_{j=i}^{d} f(j)L_{ji} \right]. \tag{8.23}$$

Now,

$$f(i) - \sum_{j=i}^{d} f(j)L_{ji} = \sum_{j=i}^{d} L_{ji} f(i) - \sum_{j=i}^{d} f(j)L_{ji} = \sum_{j=i}^{d} L_{ji} \left[f(i) - f(j) \right]. \tag{8.24}$$

Since f is increasing, we have that $f(i) - f(j) \leq 0$ when $j \geq i$. Furthermore, all the elements of $\mathbf{L}$ are non-negative, meaning that the left-hand side of Equation (8.24) is negative or equal to zero. Consequently, the left-hand side of Equation (8.23) is also negative or equal to zero. This concludes the first part of the proof.

Now, suppose that

$$\sum_{i=1}^{d} f(i)r_i \leq \sum_{i=1}^{d} f(i)s_i, \tag{8.25}$$

for any function $f : \mathbb{R} \to \mathbb{R}$ which is continuous and increasing. Choose the series of functions $f_k : \mathbb{R} \to \mathbb{R}$ which verify

$$f_k(x) = \begin{cases} -1 & \text{if } x \leq k, \\ 0 & \text{else.} \end{cases}$$

We can always find continuous and increasing functions which verify these properties. This means that

$$\sum_{i=1}^{d} f_k(i)r_i \leq \sum_{i=1}^{d} f_k(i)s_i, \quad \forall\, k, \tag{8.26}$$

so that

$$\sum_{i=1}^{k} r_i \geq \sum_{i=1}^{k} s_i, \quad \forall\, k, \tag{8.27}$$

which essentially means that $\rho \succ_{\mathrm{F}} \sigma$. This concludes the second part of the proof. □

As already hinted at, the fact that a Fock-majorization relation between two quantum states ρ and σ guaranties the existence of a lower-triangular matrix connecting their diagonal values in the Fock basis can be interpreted as the fact that it guaranties the existence of a heating map connecting the two states. This is strengthened by the idea that the expectation values $\langle f(H) \rangle_\rho = \mathrm{Tr}\left[f(H)\rho\right]$ and $\langle f(H) \rangle_\sigma = \mathrm{Tr}\left[f(H)\sigma\right]$ we verify the same inequality $\langle f(H) \rangle_\rho \leq \langle f(H) \rangle_\sigma$ for all increasing functions f. This hints to the fact that Fock-majorization may be seen as a concept

that generalises energy.

8.2.2 Behaviour of Fock-majorization in quantum channels

We know investigate the behaviour of a Fock-majorization relation in general CPTP maps. As a result of the role played by the energy basis in the context of our new relation, we begin by proving a criterion which dictates the way a passive state evolves in a quantum channel. We will call a channel $\mathcal{C}$ Fock preserving when it is such that if ρ is a Fock-diagonal state, then $\mathcal{C}[\rho]$ is also a Fock-diagonal state, *i.e.*,

$$\langle n| \left(\mathcal{C}\left[|i\rangle\langle i|\right]\right)|m\rangle = 0, \quad \forall n \neq m. \tag{8.28}$$

Phase-insensitive Gaussian bosonic channels are well known to be Fock-preserving channels since they map Fock states onto mixtures of Fock states [111]. A stronger condition is that a Fock-preserving channel $\mathcal{C}$ is passive-preserving, *i.e.,* it maps passive states onto passive states. The following theorem provides a key to determine whether any channel $\mathcal{C}$ is passive-preserving.

Theorem 33. *A bosonic quantum channel $\mathcal{C}$ satisfying*

$$\langle n| \left(\mathcal{C}\left[|i\rangle\langle i|\right]\right)|m\rangle = 0, \quad \forall n \neq m, \tag{8.29}$$

is passive-preserving if and only if its adjoint $\mathcal{C}^\dagger$ obeys the ladder of Fock-majorization relations

$$\mathcal{C}^\dagger\left[\,|k\rangle\langle k|\,\right] \succ_{\mathrm{F}} \mathcal{C}^\dagger\left[\,|k+1\rangle\langle k+1|\,\right], \quad \forall k \geq 0. \tag{8.30}$$

Proof. Form Equation (8.29), we already know that the channel is Fock preserving. Equation (8.30) is equivalent to

$$\mathrm{Tr}\left[Q_n^{\downarrow}\mathcal{C}^\dagger\left[\,|k\rangle\langle k|\,\right]\right] \geq \mathrm{Tr}\left[Q_n^{\downarrow}\mathcal{C}^\dagger\left[\,|k+1\rangle\langle k+1|\,\right]\right], \quad \forall n \geq 0, \tag{8.31}$$

where $Q_n^{\downarrow} = \sum_{i=0}^{n} |i\rangle\langle i|$. Using the definition of the adjoint of a channel, we get

$$\mathrm{Tr}\left[|k\rangle\langle k|\,\mathcal{C}[Q_n^{\downarrow}]\right] \geq \mathrm{Tr}\left[|k+1\rangle\langle k+1|\,\mathcal{C}[Q_n^{\downarrow}]\right], \quad \forall n \geq 0. \tag{8.32}$$

Now, assume that the input of channel $\mathcal{C}$ is a Fock-passive state

$$\rho = \sum_{n=0}^{\infty} r_n |n\rangle\langle n|, \quad \text{with } r_n \geq r_{n+1},\ \forall n \geq 0. \tag{8.33}$$

It can also be rewritten as

$$\rho = \sum_{n=0}^{\infty} e_n Q_n^{\downarrow}, \qquad \text{with } e_n = r_n - r_{n+1}. \tag{8.34}$$

where $e_n \geq 0$, $\forall n \geq 0$, since ρ is passive. Then, we may take the convex combination of inequalities (8.32) with weights e_n and n going from 0 to ∞, resulting in

$$\mathrm{Tr}\left[|k\rangle\langle k|\,\mathcal{C}[\rho]\right] \geq \mathrm{Tr}\left[|k+1\rangle\langle k+1|\,\mathcal{C}[\rho]\right]. \tag{8.35}$$

Hence, the output state $\mathcal{C}[\rho]$ is passive, so that channel $\mathcal{C}$ is indeed passive-preserving. Conversely, it is trivial to see that $\mathcal{C}$ being passive-preserving implies Equation (8.32) since $Q_n^{\downarrow}$ is (proportional to) a passive state, hence it implies Equation (8.30). □

We now turn to the notion of Fock-majorization, in the context of CPTP maps. We will call a channel $\mathcal{C}$ Fock-majorization preserving provided it is such that if $\rho \succ_{\mathrm{F}} \sigma$, then $\mathcal{C}[\rho] \succ_{\mathrm{F}} \mathcal{C}[\sigma]$. Fock-majorization preserving channels can be characterised through the following theorem.

Theorem 34. *A bosonic quantum channel $\mathcal{C}$ satisfying*

$$\langle n|\left(\mathcal{C}\left[|i\rangle\langle j|\right]\right)|n\rangle = 0, \quad \forall i \neq j, \tag{8.36}$$

is Fock-majorization preserving if and only if it obeys the ladder of Fock-majorization relations

$$\mathcal{C}\left[\,|k\rangle\langle k|\,\right] \succ_{\mathrm{F}} \mathcal{C}\left[\,|k+1\rangle\langle k+1|\,\right], \quad \forall k \geq 0. \tag{8.37}$$

The condition stated in Equation (8.36) is specific to passive-environment channels, as we will show later.

Proof. From Equation (8.36), one understands that the non-diagonal elements of any input state do not affect the diagonal elements of the state at the output of the channels $\mathcal{C}$. Consequently, it is sufficient to prove Theorem 34 for Fock-diagonal states. With this in mind, we start with two Fock-diagonal states

$$\rho = \sum_{i=0}^{N} r_i\,|i\rangle\langle i|, \qquad \sigma = \sum_{i=0}^{N} s_i\,|i\rangle\langle i|, \tag{8.38}$$

whose supports is the space spanned by $\{|0\rangle, \ldots |N\rangle\}$ (if their supports have unequal sizes, we take the largest size for N.) We assume that we have a Fock-majorization relation between two

states at the input of the channel, that is

$$\rho \succ_{\mathrm{F}} \sigma \qquad \Leftrightarrow \qquad R_n \geq S_n, \forall n \geq 0, \tag{8.39}$$

where

$$R_n = \mathrm{Tr}\left[Q_n^{\downarrow}\rho\right] = \sum_{i=0}^{n} r_i, \quad S_n = \mathrm{Tr}\left[Q_n^{\downarrow}\sigma\right] = \sum_{i=0}^{n} s_i. \tag{8.40}$$

We want to prove that the same Fock-majorization relation holds at the output,

$$\mathcal{C}[\rho] \succ_{\mathrm{F}} \mathcal{C}[\sigma] \qquad \Leftrightarrow \qquad R'_n \geq S'_n, \forall n \geq 0 \tag{8.41}$$

where

$$R'_n = \mathrm{Tr}\left[Q_n^{\downarrow}\mathcal{C}[\rho]\right] = \sum_{i=0}^{N} r_i \Theta_n^{(i)}, \tag{8.42}$$

and

$$S'_n = \mathrm{Tr}\left[Q_n^{\downarrow}\mathcal{C}[\sigma]\right] = \sum_{i=0}^{N} s_i \Theta_n^{(i)}, \tag{8.43}$$

with

$$\Theta_n^{(i)} = \mathrm{Tr}\left[Q_n^{\downarrow}\mathcal{C}\left[\,|i\rangle\langle i|\,\right]\right]. \tag{8.44}$$

Now, we define the quantities

$$a_n^{(k)} = R_k \Theta_n^{(k)} + \sum_{i=k+1}^{N} r_i \Theta_n^{(i)}, \qquad k = 0, \ldots N, \tag{8.45}$$

where the second term in the right-hand side is taken equal to zero when $k = N$, so that $a_n^{(N)} = R_N \Theta_n^{(N)}$. Similarly, we define

$$\beta_n^{(k)} = S_k \Theta_n^{(k)} + \sum_{i=k+1}^{N} s_i \Theta_n^{(i)}, \qquad k = 0, \ldots N, \tag{8.46}$$

with the convention $\beta_n^{(N)} = S_N \Theta_n^{(N)}$. The Fock-majorization relation we need to prove, Equation (8.41), is equivalent to

$$a_n^{(0)} \geq \beta_n^{(0)}, \qquad \forall n \geq 0 \tag{8.47}$$

corresponding to $k = 0$. We will now prove

$$a_n^{(k)} \geq \beta_n^{(k)}, \qquad \forall n \geq 0 \tag{8.48}$$

by recurrence in k, starting from $k = N$ and ending at $k = 0$. We have trivially $a_n^{(N)} \geq \beta_n^{(N)}$,

$\forall n \geq 0$, since $R_N = S_N = 1$. Now, we assume that

$$a_n^{(k+1)} \geq \beta_n^{(k+1)}, \qquad \forall n \geq 0 \tag{8.49}$$

which can be rewritten as

$$R_{k+1}\Theta_n^{(k+1)} + \sum_{i=k+2}^{N} r_i \Theta_n^{(i)} \geq S_{k+1}\Theta_n^{(k+1)} + \sum_{i=k+2}^{N} s_i \Theta_n^{(i)}.$$

Using $R_{k+1} = R_k + r_{k+1}$ and $S_{k+1} = S_k + s_{k+1}$, we reexpress it as

$$R_k\Theta_n^{(k+1)} + \sum_{i=k+1}^{N} r_i \Theta_n^{(i)} \geq S_k\Theta_n^{(k+1)} + \sum_{i=k+1}^{N} s_i \Theta_n^{(i)}, \tag{8.50}$$

which is equivalent to

$$R_k \left(\Theta_n^{(k+1)} - \Theta_n^{(k)}\right) + a_n^{(k)} \geq S_k \left(\Theta_n^{(k+1)} - \Theta_n^{(k)}\right) + \beta_n^{(k)}, \tag{8.51}$$

or simply

$$a_n^{(k)} - \beta_n^{(k)} \geq (R_k - S_k) \left(\Theta_n^{(k)} - \Theta_n^{(k+1)}\right). \tag{8.52}$$

Since ρ Fock-majorizes σ by hypothesis (Equation (8.39)), we have $R_k - S_k \geq 0, \forall k \geq 0$. If $\mathcal{C}\left[\, |k\rangle \langle k| \,\right]$ Fock-majorizes $\mathcal{C}\left[\, |k+1\rangle \langle k+1| \,\right]$, which means that $\Theta_n^{(k)} - \Theta_n^{(k+1)} \geq 0, \forall n \geq 0$, then the right-hand side of Equation (8.52) is greater than zero. Thus, Equation (8.49) implies Equation (8.48), which concludes the recurrence in k and proves Equation (8.47), hence Equation (8.41). Conversely, it is trivial to see that Fock-majorization preservation for channel $\mathcal{C}$ implies the ladder of Fock-majorization relations since individual Fock states satisfy the Fock-majorization relation $|n\rangle \langle n| \succ_{\text{F}} |n+1\rangle \langle n+1|, \forall n \geq 0$. □

Finally, one notices that there is a duality between the concepts of passive preservation and Fock-majorization preservation. This can be seen from the fact that the statements of Theorems 33 and 34 are very similar. In order to be able to express this, we generalise the notion of passive-preserving channel to passive-preserving map $\mathcal{M}$, which is such that if a state $\rho^{\downarrow}$ is passive, then $\mathcal{M}[\rho]^{\downarrow}$ is such that

$$\langle n| \left(\mathcal{M}\left[\rho^{\downarrow}\right]\right) |n\rangle \geq \langle n+1| \left(\mathcal{M}\left[\rho^{\downarrow}\right]\right) |n+1\rangle, \quad \forall n \geq 0. \tag{8.53}$$

The reason we do this is because the adjoint map of a channel is not necessarily a channel as well, since it does not have to be trace preserving. This is for instance the case for extremal-passive channels, as implied by Theorem 30. We are now in position to state the following Theorem.

Theorem 35. *A bosonic quantum channel $\mathcal{C}$ satisfying*

$$\langle n| \left(\mathcal{C}\left[|i\rangle\langle j|\right]\right)|n\rangle = 0, \quad \forall i \neq j, \tag{8.54}$$

is Fock-majorization preserving if and only if its adjoint $\mathcal{C}^\dagger$ is passive-preserving.

Proof. This is simply a result from the fact that Equation (8.54) is equivalent to

$$\langle j| \left(\mathcal{C}^\dagger\left[|n\rangle\langle n|\right]\right)|i\rangle = 0, \quad \forall i \neq j, \tag{8.55}$$

and from Theorems 33 and 34. □

As already mentioned, Fock-majorization bears some analogy with the "upper-triangular majorization" of Reference [110]. It may thus be quite fruitful to investigate the thermodynamical consequences of the existence of Fock-majorization, just as it was done for upper-triangular majorization in the context of cooling maps. The latter maps happen to be a special case of the so-called "thermal maps", which result from the coupling with a finite-temperature heat bath and are linked to another type of majorization relation, called "thermo-majorization" [112]. Since these various thermal operations provide a suitable model in the study of thermodynamical processes for microscopic systems, we anticipate that our results on Fock-majorization will find interesting applications in the field of quantum thermodynamics. In order to conclude, let us mention that since the results we proved in the context of our new order relation do not depend on the bosonic nature of the system, the definition of the relation can be generalised to non-bosonic systems. One may then denote it as energy-majorization, since it would maintain its connection with the eigenstates of the Hamiltonian of the system.

9

Fock-majorization in Gaussian-dilatable channels

The importance of majorization theory in continuous-variable quantum information theory was first suggested by Guha in [7], specifically in the context of Gaussian bosonic channels. Guha was concerned with the classical capacity of these channels (see [113]), which was known to require proving a Gaussian minimum entropy conjecture [114]. Denoting an arbitrary phase-insensitive Gaussian bosonic channel by $\mathcal{G}[\bullet]$, the conjecture was that

$$S\left(\mathcal{G}[|\psi\rangle\langle\psi|]\right) \geq S\left(\mathcal{G}[|0\rangle\langle 0|]\right), \tag{9.1}$$

for any input pure state $|\psi\rangle$, where $|0\rangle$ is the vacuum state. The intuition was that a majorization relation

$$\mathcal{G}\left[|0\rangle\langle 0|\right] \succ \mathcal{G}\left[|\psi\rangle\langle\psi|\right] \tag{9.2}$$

might be responsible for the conjectured entropic inequality. Equation (9.1) was later generalised as stated in Equation (4.40) of Chapter 4.

The existence of majorization relations in Gaussian bosonic channels was first proven in [81], where the quantum-limited amplifier $\mathcal{A}_G[\bullet]$ was proven to obey an infinite ladder of majorization relations when the input state is an individual Fock state, namely

$$\mathcal{A}_G\left[|k\rangle\langle k|\right] \succ \mathcal{A}_G\left[|k+1\rangle\langle k+1|\right], \quad \forall k \geq 0. \tag{9.3}$$

A parametric majorization relation was also proven for varying gain G, namely

$$\mathcal{A}_G\left[|k\rangle\langle k|\right] \succ \mathcal{A}_{G+\delta G}\left[|k\rangle\langle k|\right] \quad \text{if } \delta G \geq 0. \tag{9.4}$$

Then, in Reference [103], a similar ladder of majorization relations was proven for a pure-loss channel (see Equation (9.17)). Finally, the interconversion between pure Gaussian states was also investigated based on majorization theory [75, 115], which revealed the existence of surprising situations where a non-Gaussian LOCC is required although the states considered are Gaussian.

In this chapter, drawing inspiration from the many applications of majorization in Gaussian bosonic channels, we study the concept of Fock-majorization in the more general framework of Gaussian-dilatable channels with a passive environment, which were introduced in Chapter 7. Specifically, we begin by conjecturing a generalisation of the entropy photon-number inequality in terms of majorization in Section 9.1; we call it the precursor of the EPnI. We also conjecture an extension of the ladder of majorization relations of Equation (9.3) to passive-environment channels. Motivated by the precursor of the EPnI, we show how the Fock-majorization relation happens to be the fundamental relation that is conserved in a passive-environment channel, unlike regular majorization. We begin by showing this in the context of Gaussian bosonic channels in Section 9.2, discussing the implications in terms of the recently solved minimum output entropy conjecture [92, 116] in the process. In Section 9.3, we turn to general passive-environment channels, for which this kind of entropic inequalities (related to the EPnI) remain to be proven. Still, we show that Fock-majorization relations are preserved in any passive-environment channel, before discussing open problems regarding regular majorization.

9.1 Motivation: the precursor of the EPnI

In Chapter 4, we introduced the entropy photon-number inequality proposed by Guha in the form of Equation 4.36. To state the inequality, one defines the map $\Phi_\eta\left[\bullet,\bullet\right]$ acting on two states ρ_a and ρ_b as

$$\Phi_\eta\left[\rho_a,\rho_b\right] = \mathrm{Tr}_2\left[U_\eta^{\mathrm{BS}}\left(\rho_a\otimes\rho_b\right)U_\eta^{\mathrm{BS}\dagger}\right], \tag{9.5}$$

where U_η^{BS} is a beam splitter, as depicted in Figure 9.1.1 and $\mathrm{Tr}_2\left[\bullet\right]$ represents the partial trace on the second mode of the system. The entropy photon-number inequality then asserts that

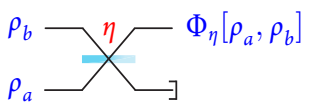

Figure 9.1.1: Set-up considered for the EPnI. The states ρ_a and ρ_b evolve in a beam splitter of transmittance η. The second mode is then discarded.

$$S\left(\Phi_\eta\left[\rho_a,\rho_b\right]\right)\geq S\left(\Phi_\eta\left[\zeta_{\bar{n}_a},\zeta_{\bar{n}_b}\right]\right), \tag{9.6}$$

where $\zeta_{\bar{n}_a}$ and $\zeta_{\bar{n}_b}$ are two thermal Gaussian states having the same entropies as ρ_a and ρ_b, respectively, *i.e.*,

$$\begin{cases} S\left(\zeta_{\bar{n}_a}\right)=S\left(\rho_a\right), \\ S\left(\zeta_{\bar{n}_b}\right)=S\left(\rho_b\right). \end{cases} \tag{9.7}$$

As we explained in Chapter 4, the entropy photon-number inequality is inspired from the entropy power inequality introduced by Shannon. We also showed how the latter can be generalised through the concept of majorization for continuous probability densities, in which case the rearrangements defined in Equation (2.47) happen to be optimal in the sense that they majorize all equivalent distributions after being transformed by a convolution (see Section 4.1.4, Theorem 25). We conjecture that the entropy photon-number inequality can be generalised following the same intuition. The conjecture is encompassed in the following statement.

Conjecture 2 (Precursor of the entropy photon-number inequality)**.** *Consider two quantum states ρ_a and ρ_b, and the two corresponding Fock-passive states $\rho_a^{\downarrow}\equiv\rho_a$ and $\rho_b^{\downarrow}\equiv\rho_b$. In this case,*

$$\Phi_\eta\left[\rho_a^{\downarrow},\rho_b^{\downarrow}\right]\succ\Phi_\eta\left[\rho_a,\rho_b\right], \tag{9.8}$$

where

$$\Phi_\eta\left[\rho_a,\rho_b\right]=\mathrm{Tr}_2\left[U_\eta^{\mathrm{BS}}\left(\rho_a\otimes\rho_b\right)U_\eta^{\mathrm{BS}\dagger}\right], \tag{9.9}$$

U_η^{BS} being a beam splitter of transmittance η.

For completeness, let us mention that Equation (9.8) can be generalised as follows.

Lemma 12. *The precursor of the entropy photon-number inequality as stated in Conjecture 2 is equivalent to the relation*

$$\Phi_\eta\left[U_{\mathrm{G}}\rho_a^{\downarrow}U_{\mathrm{G}}^{\dagger},U_{\mathrm{G}}\rho_b^{\downarrow}U_{\mathrm{G}}^{\dagger}\right]\succ\Phi_\eta\left[\rho_a,\rho_b\right], \tag{9.10}$$

for any one-mode Gaussian unitary U_{G}.

Proof. Any one-mode Gaussian unitary U_{G} can be expressed as

$$U_{\mathrm{G}}=D_\alpha U_\theta^{\mathrm{R}}U_r^{\mathrm{S}}U_\varphi^{\mathrm{R}}, \tag{9.11}$$

where D_α is a displacement, U_r^{S} is a squeezer unitary and U_θ^{R} and U_φ^{R} are phase rotation unitaries. Now, Equation (5.22) implies that

$$U_\eta^{\mathrm{BS}}\left(D_\alpha\otimes D_\alpha\right)=\left(D_{\left(\sqrt{\eta}+\sqrt{1-\eta}\right)\alpha}\otimes D_{\left(\sqrt{\eta}-\sqrt{1-\eta}\right)\alpha}\right)U_\eta^{\mathrm{BS}}. \tag{9.12}$$

Using the same kind of arguments, it can be shown that

$$U_\eta^{\mathrm{BS}} \left(U_{\mathrm{G}} \otimes U_{\mathrm{G}}\right) = \left(U_{\mathrm{G}}^{(1)} \otimes U_{\mathrm{G}}^{(2)}\right) U_\eta^{\mathrm{BS}}, \tag{9.13}$$

for some one-mode unitaries $U_{\mathrm{G}}^{(1)}$ and $U_{\mathrm{G}}^{(2)}$. Since one traces over the second mode in order to obtain the effect of Φ_η, and since any two states which can be connected through a unitary are equivalent in terms of majorization, we have that

$$\Phi_\eta \left[U_{\mathrm{G}} \rho_a U_{\mathrm{G}}^\dagger, U_{\mathrm{G}} \rho_b U_{\mathrm{G}}^\dagger\right] \equiv \Phi_\eta \left[\rho_a, \rho_b\right]. \tag{9.14}$$

This ends the proof. □

Conjecture 2 was shown to be true in the specific case in which ρ_b is fixed to be in a thermal Gaussian state τ_{ε_b}, which is passive in the Fock basis. In this case, one actually deals with a lossy Gaussian channel acting on ρ_a,

$$\Phi[\rho_a, \tau_b] = \mathcal{B}_\eta^{\varepsilon_b}[\rho_a], \tag{9.15}$$

and Equation (9.8) becomes

$$\mathcal{B}_\eta^{\varepsilon_b}[\rho_a^{\downarrow}] \succ \mathcal{B}_\eta^{\varepsilon_b}[\rho_a]. \tag{9.16}$$

Equation (9.16) was first proven for all pure input state ρ_a [117], before being extended to all input mixed states [116]. Furthermore, in the same context, the pure-loss channel $\mathcal{B}_\eta$ was shown to obey some ladder majorization relations for Fock input states [103], in the form of

$$\mathcal{B}_\eta \left[|i\rangle \langle i|\right] \succ \mathcal{B}_\eta \left[|i+1\rangle \langle i+1|\right], \quad \forall\, i \in \mathbb{N}_0. \tag{9.17}$$

This implied in particular that among all Fock states, the vacuum majorizes all other Fock states having evolved through the pure-loss channel. In order to prove Equation (9.17), the authors considered the recurrence relation of Equation (6.1), for $k = 0$ and $j = 1$, *i.e.*,

$$B_{n+1}^{(i+1,0)} = \eta\, B_n^{(i,0)} + (1-\eta)\, B_{n+1}^{(i,0)}, \tag{9.18}$$

which is connected to the present problem through the relation $\langle n| \mathcal{B}_\eta \left[|i\rangle \langle i|\right] |n\rangle = B_n^{(i,0)}$. Equation (9.18) can actually be translated into a bistochastic transformation connecting the two sides of Equation (9.17), so that a majorization relation can be obtained. Again, we conjecture the same ladder of majorization relations for extremal-passive channels.

Conjecture 3. *The extremal passive-channels $\mathcal{B}_\eta^{[k]}$ obey the infinite ladder of majorization relations*

$$\mathcal{B}_\eta^{[k]} \left[|i\rangle \langle i|\right] \succ \mathcal{B}_\eta^{[k]} \left[|i+1\rangle \langle i+1|\right], \quad \forall\, k \in \mathbb{N}_0, \quad \forall\, i \in \mathbb{N}_0. \tag{9.19}$$

In order to prove Conjecture 3, a natural idea would be to use the recurrence relation of Equation (6.1) for $j = 1$, but this time for an arbitrary value of k, *i.e.*,

$$B_{n+1}^{(i+1,k+1)} = \eta\, B_n^{(i,k+1)} + (1-\eta)\, B_{n+1}^{(i,k+1)} + \eta\, B_{n+1}^{(i+1,k)} + (1-\eta)\, B_n^{(i+1,k)} - B_n^{(i,k)}, \tag{9.20}$$

by applying the right summation on k. As already explained in details in Chapter 6, the last equation involves an interference term multiplied by a negative factor, which does not appear if $k = -1$ (Gaussian case). Unfortunately, this makes it difficult to develop a method which would provide us with a bistochastic matrix implying Equation (9.19) for any value of k in $\mathcal{B}_\eta^{[k]}$, as a bistochastic matrix should only have non-negative entries.

The precursor of the entropy photon-number inequality encompassed in Conjecture 2 basically states that if one fixes a spectrum on each of the two input ports of a beam splitter, the optimal states will be given by two passive states. Here, they will be optimal in the sense that they will generate a state which will majorize all others at each output port of the beam splitter. One may wonder whether this can be generalised further by imposing majorization relations at the two input port of the beam splitter. In other words, we ask the question whether majorization relation are preserved through a beam splitter. This can mathematically be stated as

$$\begin{cases} \rho_a^{(1)} \succ \rho_a^{(2)} \\ \rho_b^{(1)} \succ \rho_b^{(2)} \end{cases} \quad \overset{?}{\Rightarrow} \quad \Phi_\eta\left[\rho_a^{(1)}, \rho_b^{(1)}\right] \succ \Phi_\eta\left[\rho_a^{(2)}, \rho_b^{(2)}\right]. \tag{9.21}$$

One can actually already see that Equation (9.21) is not consistent with Equation (9.8). Indeed, since $\rho_a \equiv \rho_a^{\downarrow}$ and $\rho_b \equiv \rho_b^{\downarrow}$, we also have that $\rho_a \succ \rho_a^{\downarrow}$ and $\rho_b \succ \rho_b^{\downarrow}$. In this case, Equation (9.21) implies that $\Phi_\eta\left[\rho_a, \rho_b\right] \succ \Phi_\eta\left[\rho_a^{\downarrow}, \rho_b^{\downarrow}\right]$, while Equation (9.8) states $\Phi_\eta\left[\rho_a^{\downarrow}, \rho_b^{\downarrow}\right] \succ \Phi_\eta\left[\rho_a, \rho_b\right]$. This is possible if and only if

$$\Phi_\eta\left[\rho_a^{\downarrow}, \rho_b^{\downarrow}\right] \equiv \Phi_\eta\left[\rho_a, \rho_b\right], \tag{9.22}$$

which is obviously not the case in general. Furthermore, one can easily find counter-examples for the statement of Equation (9.21). With this in mind, we are now motivated to look for the fundamental relation which is preserved through a beam splitter. In order to do this, we turn to passive-environment channels, which are of primary importance in this thesis. If one particularises the precursor of the entropy photon-number inequality by fixing one of the input states to already be optimal, one ends up with the following less general conjecture.

Conjecture 4. *Consider a quantum state ρ, and the corresponding Fock-passive state $\rho^{\downarrow} \equiv \rho$. In this case,*

$$\mathcal{B}_\eta^{\downarrow}\left[\rho^{\downarrow}\right] \succ \mathcal{B}_\eta^{\downarrow}\left[\rho\right], \tag{9.23}$$

where

$$\mathcal{B}_\eta^{\downarrow}[\rho] = \mathrm{Tr}_2\left[U_\eta^{\mathrm{BS}}\left(\rho \otimes \sigma^{\downarrow}\right) U_\eta^{\mathrm{BS}\dagger}\right], \tag{9.24}$$

U_η^{BS} *being a beam splitter of transmittance* η *and* $\sigma^{\downarrow}$ *being any Fock-passive state.*

As one may have anticipated, inspired by Conjecture 4, we find that the fundamental relation which is conserved through any passive-environment channel is exactly the Fock-majorization relation introduced in Chapter 8. This fact is investigated in details in the next sections of this chapter. Let us mention that Conjectures 2, 3 and 4 can all be restated with the beam splitter U_η^{BS} being replaced by a two-mode squeezer U_λ^{TMS} in each of them.

9.2 Preservation of a majorization relation in Gaussian channels

In this section, we prove a new type of intrinsic majorization property in Gaussian bosonic channels, namely the conservation across any channel $\mathcal{G}$ of a Fock-majorization relation between any two comparable states. This implies in turn that Gaussian bosonic channels preserve regular majorization over the set of passive states of the harmonic oscillator. We then discuss the connection of this result with the entropy photon-number inequality particularised to Gaussian channels.

We begin by showing that the most basic Fock-majorization relation is preserved in Gaussian channels. We state and prove the relation for the quantum-limited Gaussian channels, since they constitute the building blocks which allow one to decompose any Gaussian channel. The results are contained in the following lemmas. We begin with the pure-loss channel.

Lemma 13. *The pure-loss channel* $\mathcal{B}_\eta$ *exhibits a ladder of Fock-majorization relations*

$$\mathcal{B}_\eta\left[\,|i\rangle\langle i|\,\right] \succ_{\mathrm{F}} \mathcal{B}_\eta\left[\,|i+1\rangle\langle i+1|\,\right], \quad \forall i \geq 0. \tag{9.25}$$

It is known that a similar relation holds when replacing Fock-majorization with majorization [103]. Here, we will adapt this proof in order to derive a Fock-majorization relation.

Proof. We have

$$\rho^{(i)} := \mathcal{B}_\eta\left[\,|i\rangle\langle i|\,\right] = \sum_{n=0}^{i} B_n^{(i,0)}\,|n\rangle\langle n|\,, \tag{9.26}$$

where

$$B_n^{(i,0)} = \binom{i}{n}\eta^n(1-\eta)^{i-n}, \tag{9.27}$$

and η is the transmittance of channel $\mathcal{B}_\eta$. Majorization was proven in [103] based on the recur-

rence relation of Equation (6.1), for $k = 0$ and $j = 1$, *i.e.*,

$$B_n^{(i+1,0)} = \eta\, B_{n-1}^{(i,0)} + (1-\eta)\, B_n^{(i,0)}, \quad \forall i \geq 0, \forall n \geq 0, \tag{9.28}$$

where the first term in the right-hand side is taken equal to zero for $n = 0$. We can rewrite it as

$$B_n^{(i,0)} - B_n^{(i+1,0)} = \eta \left(B_n^{(i,0)} - B_{n-1}^{(i,0)} \right), \tag{9.29}$$

Hence,

$$\sum_{n=0}^{j} B_n^{(i,0)} - \sum_{n=0}^{j} B_n^{(i+1,0)} = \eta\, B_j^{(i,0)} \geq 0, \quad \forall j \geq 0, \tag{9.30}$$

which gives the Fock-majorization relation $\rho^{(i)} \succ_F \rho^{(i+1)}$ in addition to the majorization relation $\rho^{(i)} \succ \rho^{(i+1)}$ of Reference [103]. □

We now turn to the quantum-limited amplifier.

Lemma 14. *The quantum-limited amplifier $\mathcal{A}_G$ exhibits a ladder of Fock-majorization relations*

$$\mathcal{A}_G\big[\, |i\rangle\langle i| \,\big] \succ_F \mathcal{A}_G\big[\, |i+1\rangle\langle i+1| \,\big], \quad \forall i \geq 0. \tag{9.31}$$

We also use the related majorization property for an amplifier as proven in Reference [81].

Proof. We have

$$\sigma^{(i)} := \mathcal{A}_G\big[\, |i\rangle\langle i| \,\big] = \sum_{n=0}^{\infty} A_{i+n}^{(i,0)}\, |n+i\rangle\langle n+i| = \sum_{n=i}^{\infty} A_n^{(i,0)}\, |n\rangle\langle n|, \tag{9.32}$$

where

$$A_{i+n}^{(i,0)} = \binom{n+i}{n} \lambda^n (1-\lambda)^{i+1}, \tag{9.33}$$

and $\lambda = \tanh^2(r)$ is related to the gain $G = 1/(1-t)$ of the amplifier $\mathcal{A}_G$, with r being the squeezing parameter. Majorization was proven in [81] by using the recurrence relation of Equation (6.39), for $k = 0$ and $j = 1$, *i.e.*,

$$A_{i+1+n}^{(i+1,0)} = \lambda\, A_{i+n}^{(i+1,0)} + (1-\lambda)\, A_{i+n}^{(i,0)}, \quad \forall i \geq 0, \forall n \geq 0, \tag{9.34}$$

where the first term in the right-hand side is taken equal to zero for $n = 0$. We can rewrite it as

$$A_{i+n}^{(i,0)} - A_{i+1+n}^{(i+1,0)} = (G-1) \left(A_{i+1+n}^{(i+1,0)} - A_{i+n}^{(i+1,0)} \right). \tag{9.35}$$

The differences between the cumulated sums of eigenvalues are given by

$$\sum_{n=i}^{j} A_n^{(i,0)} - \sum_{n=i+1}^{j} A_n^{(i+1,0)} \geq \sum_{n=i}^{j} (A_n^{(i,0)} - A_{n+1}^{(i+1,0)}). \tag{9.36}$$

Using Equation (9.35), we have

$$\sum_{n=i}^{j} A_n^{(i,0)} - \sum_{n=i+1}^{j} A_n^{(i+1,0)} \geq (G-1) A_{j+1}^{(i+1,0)} \geq 0, \quad \forall j \geq 0, \tag{9.37}$$

giving the Fock-majorization relation $\sigma^{(i)} \succ_{\mathrm{F}} \sigma^{(i+1)}$ in addition to the majorization relation $\sigma^{(i)} \succ \sigma^{(i+1)}$ [81]. □

We are now in position to prove the Fock-majorization preservation of Gaussian channels. Since we plan to use the fact that majorization and Fock-majorization are completely equivalent for passive states, we begin by showing the preservation of passivity, which is the result of the following theorem.

Theorem 36. *Phase-insensitive Gaussian bosonic channels are passive preserving.*

Proof. Using Lemma 13 and 14 together with Theorem 33, we obtain that the pure-loss channel $\mathcal{B}_\eta$, whose adjoint is $1/\eta$ times the quantum-limited amplifier $\mathcal{A}_{1/\eta}$, as well as the quantum-limited amplifier $\mathcal{A}_G$, whose adjoint is $1/G$ times the pure-loss channel $\mathcal{B}_{1/G}$, are both passive preserving. Then, the corollary follows from the fact that any phase-insensitive Gaussian bosonic channel $\mathcal{G}$ can be expressed as the concatenation of a pure-loss channel $\mathcal{B}$ and a quantum-limited amplifier $\mathcal{A}$, i.e., $\mathcal{G} = \mathcal{A} \circ \mathcal{B}$ (see Equations (3.106) and (3.108)), and that passive-preservation is transitive over channel composition. □

Theorem 37. *Phase-insensitive Gaussian bosonic channels are Fock-majorization-preserving.*

Proof. We use again the fact that any phase-insensitive Gaussian bosonic channel $\mathcal{G}$ can be expressed as the concatenation $\mathcal{G} = \mathcal{A} \circ \mathcal{B}$ and that Fock-majorization preservation is transitive over channel composition. Since any phase-insensitive Gaussian channel $\mathcal{G}$ satisfies

$$\langle n | (\mathcal{G}[|i\rangle\langle j|]) | n \rangle = 0, \quad \forall i \neq j, \tag{9.38}$$

we can make use of Theorem 34 to end the proof. □

Corollary 6. *Phase-insensitive Gaussian bosonic channels are majorization-preserving over the set of passive states.*

Proof. As a consequence of the equivalence between Fock-majorization and regular majorization over the set of passive states, a Fock-majorization preserving channel is necessarily also

majorization preserving over the set of passive states provided it is passive preserving. Since phase-insensitive Gaussian bosonic channels are passive preserving (Theorem 36) and Fock-majorization preserving (Theorem 37), we conclude that they preserve regular majorization over the set of passive states. □

For completeness, let us mention that our results can also be extended to the set of phase-conjugate Gaussian bosonic channels, which can be expressed as a concatenation of a pure-loss channel $\mathcal{B}_\eta$ and a quantum-limited phase-conjugate channel $\tilde{\mathcal{A}}_G$. The interested reader is referred to Appendix F.1.

Corollary 6 nicely complements the property found in [116]. There, it was shown that among all isospectral states ρ at the input of a phase-insensitive Gaussian bosonic channel $\mathcal{G}$, the passive state, denoted as $\rho^\downarrow$, produces an output state that majorizes all other output states, namely $\mathcal{G}[\rho^\downarrow] \succ \mathcal{G}[\rho]$. Here, we consider instead two input states that have different spectra but are both passive, $\rho^\downarrow$ and $\sigma^\downarrow$, and have demonstrated that $\rho^\downarrow \succ \sigma^\downarrow$ implies $\mathcal{G}[\rho^\downarrow] \succ \mathcal{G}[\sigma^\downarrow]$. This reflects the fact that Gaussian bosonic channels exhibit quite a wide variety of majorization properties, going well beyond what was originally expected in Reference [7]. As a matter of fact, Corollary 6 may be combined together with the result of [116], giving what can be viewed as the *fundamental majorization-preservation property*

$$\rho^\downarrow \succ \sigma \quad \Rightarrow \quad \mathcal{G}[\rho^\downarrow] \succ \mathcal{G}[\sigma], \tag{9.39}$$

valid for any phase-insensitive Gaussian bosonic channel $\mathcal{G}$. Interestingly, this property (unlike the one of [116]) is transitive if we concatenate several passive-preserving channels. In particular, it means that proving it for an infinitesimal channel (*e.g.*, using the Lindbladian) suffices to prove it for any concatenated channel.

9.3 Preservation of Fock-majorization in passive-environment channels

We now focus on general passive-environment channels. One may wonder why we began by investigating Fock-majorization in Gaussian bosonic channels in the previous section, as they can be seen as a particular case of passive-environment channels. As it happens, we will make use of the results obtained in the Gaussian case in order to prove the results for more general passive-environment channels. We begin by focusing on beam-splitter passive-environment channels $\mathcal{B}_\eta^\downarrow$. In order to prove the preservation of Fock-majorization, we again recourse to Theorem 34. In order to do so, we prove in Appendix F.2 that any channel $\mathcal{B}_\eta^\downarrow$ satisfies

$$\langle n| \left(\mathcal{B}_\eta^\downarrow \left[|i\rangle \langle j|\right]\right) |n\rangle = 0, \quad \forall i \neq j. \tag{9.40}$$

Thus, we are left with having to prove the following lemma.

Lemma 15. *The passive-environment bosonic channel $\mathcal{B}_\eta^\downarrow$ exhibits a ladder of Fock-majorization relations*

$$\mathcal{B}_\eta^\downarrow\left[\,|i\rangle\langle i|\,\right] \succ_{\mathrm{F}} \mathcal{B}_\eta^\downarrow\left[\,|i+1\rangle\langle i+1|\,\right], \quad \forall i \geq 0. \tag{9.41}$$

Proof. We begin by proving the ladder of Fock-majorization relations for an extremal-passive channel $\mathcal{B}_\eta^{[K]}$ characterised by an environment which is a projector onto the space spanned by the $K+1$ first Fock states $|k\rangle$, i.e,

$$\mathcal{B}_\eta^{[K]}(\rho) = \mathrm{Tr}_2\left[U_\eta^{\mathrm{BS}}\left(\rho \otimes P_K^\downarrow\right) U_\eta^{\mathrm{BS}\dagger}\right], \tag{9.42}$$

where $P_K^\downarrow = \frac{1}{K+1}\sum_{k=0}^{K}|k\rangle\langle k|$. We need to show that

$$\mathcal{B}_\eta^{[K]}\left[\,|i\rangle\langle i|\,\right] \succ_{\mathrm{F}} \mathcal{B}_\eta^{[K]}\left[\,|i+1\rangle\langle i+1|\,\right], \forall i \geq 0, \tag{9.43}$$

or,

$$\mathrm{Tr}\left[\mathrm{Q}_{\mathrm{n}}^\downarrow\left(\mathcal{B}_\eta^{[\mathrm{K}]}\left[\,|\mathrm{i}\rangle\langle \mathrm{i}|\,\right] - \mathcal{B}_\eta^{[\mathrm{K}]}\left[\,|\mathrm{i}+1\rangle\langle \mathrm{i}+1|\,\right]\right)\right] \geq 0, \quad \forall \mathrm{i} \geq 0, \forall \mathrm{n} \geq 0. \tag{9.44}$$

where $Q_n^\downarrow = \sum_{m=0}^{n}|m\rangle\langle m|$. Using the definition of the transition probabilities in Equation (5.67), what we need to prove here is that

$$\Delta_n^{(i,K)} = \sum_{k=0}^{K}\sum_{m=0}^{n}\left[B_m^{(i,k)} - B_m^{(i+1,k)}\right] \geq 0, \quad \forall i \geq 0, \forall n \geq 0, n \leq i+k. \tag{9.45}$$

Using the recurrence relation of Equation (6.1) for $j=1$, we have that

$$\begin{aligned}\Delta_n^{(i,K)} = &\sum_{k=0}^{K}\sum_{m=0}^{n}\left[B_m^{(i,k)} - (1-\eta)B_m^{(i,k)}\right] \\ &- \sum_{k=0}^{K}\sum_{m=0}^{n}\left[\eta B_{m-1}^{(i,k)} + (1-\eta)B_{m-1}^{(i+1,k-1)} + \eta B_m^{(i+1,k-1)} - B_{m-1}^{(i,k-1)}\right],\end{aligned}$$

$$\begin{aligned}\Delta_n^{(i,K)} = &\eta\sum_{k=0}^{K}\sum_{m=0}^{n}\left(B_m^{(i,k)} - B_{m-1}^{(i,k)}\right) + \eta\sum_{k=0}^{K}\sum_{m=0}^{n}\left(B_{m-1}^{(i+1,k-1)} - B_m^{(i+1,k-1)}\right) \\ &+ \sum_{k=0}^{K}\sum_{m=0}^{n}\left(B_{m-1}^{(i,k-1)} - B_{m-1}^{(i+1,k-1)}\right),\end{aligned}$$

$$\Delta_n^{(i,K)} = \eta \sum_{k=0}^{K} B_n^{(i,k)} - \eta \sum_{k=0}^{K} B_n^{(i+1,k-1)} + \sum_{k=0}^{K-1} \sum_{m=0}^{n-1} \left(B_m^{(i,k)} - B_m^{(i+1,k)} \right)$$

$$\begin{aligned}\Delta_n^{(i,K)} &= \eta \sum_{k=0}^{K-1} B_n^{(i,k)} + \eta B_n^{(i,k)} - \eta \sum_{k=0}^{K-1} B_n^{(i+1,k)} \\ &\quad + \eta \sum_{k=0}^{K-1} \sum_{m=0}^{n-1} \left(B_m^{(i,k)} - B_m^{(i+1,k)} \right) + (1-\eta)\Delta_{n-1}^{(i,K-1)},\end{aligned}$$

so that

$$\Delta_n^{(i,K)} = \eta B_n^{(i,k)} + \eta \Delta_n^{(i,K-1)} + (1-\eta)\Delta_{n-1}^{(i,K-1)}. \tag{9.46}$$

We know that $\Delta_n^{(i,0)} \geq 0, \forall i \geq 0, \forall n \geq 0$, since it corresponds to a Gaussian pure-loss channel, and was proven in the previous section. We are then able to prove Equation (9.45) by using a recursive argument on K, since $B_n^{(i,k)} \geq 0, \forall i \geq 0, \forall n \geq 0, \forall K \geq 0$. This shows that

$$\mathcal{B}_\eta^{[K]}\left[|i\rangle\langle i| \right] \succ_F \mathcal{B}_\eta^{[K]}\left[|i+1\rangle\langle i+1| \right], \quad \forall i \geq 0. \tag{9.47}$$

Now, since any passive state can be written as a convex sum of projectors $P_K^{\downarrow}$, it also means that

$$\mathcal{B}_\eta^{\downarrow}\left[|i\rangle\langle i| \right] \succ_F \mathcal{B}_\eta^{\downarrow}\left[|i+1\rangle\langle i+1| \right], \quad \forall i \geq 0, \tag{9.48}$$

which concludes the proof of Lemma 15. □

Using Theorem 34 and Lemma 15, we are able to conclude with the following Theorem.

Theorem 38. *Passive-environment bosonic channels $\mathcal{B}_\eta^{\downarrow}$ are Fock-majorization preserving; that is, for all states ρ and σ,*

$$\rho \succ_F \sigma \quad \Rightarrow \quad \mathcal{B}_\eta^{\downarrow}[\rho] \succ_F \mathcal{B}_\eta^{\downarrow}[\sigma]. \tag{9.49}$$

In Section 9.1, we explained that passive-environment channels do not preserve regular majorization, namely that if $\rho \succ \sigma$, one cannot conclude that $\mathcal{B}_\eta^{\downarrow}[\rho] \succ \mathcal{B}_\eta^{\downarrow}[\sigma]$. Nevertheless, similarly to the Gaussian case, one can prove that passive-environment channels are majorization preserving when restricting to the set of passive states. As a result of the equivalence between majorization and Fock-majorization for this set, we simply need to verify that passive states remain passive after evolving through a passive-environment bosonic channel. This is the content of the following Theorem.

Theorem 39. *Passive-environment bosonic channel $\mathcal{B}_\eta^{\downarrow}$ are passive preserving; that is, if $\rho^{\downarrow}$ is passive then $\mathcal{B}_\eta^{\downarrow}[\rho^{\downarrow}]$ is passive as well.*

Proof. We begin by showing that this Theorem is true for any passive channel $\mathcal{B}_\eta^{[K]}$, but when the input is a projector $Q_I^{\downarrow}$. We need to prove that

$$\mathrm{Tr}\left[\left(|n\rangle\langle n| - |n+1\rangle\langle n+1|\right)\mathcal{B}_\eta^{[K]}\left[Q_I^{\downarrow}\right]\right] \geq 0, \quad \forall I \geq 0, \forall n \geq 0, \tag{9.50}$$

or,

$$\Gamma_n^{(I,K)} = \sum_{i=0}^{I}\sum_{k=0}^{K}\left(B_n^{(i,k)} - B_{n+1}^{(i,k)}\right) \geq 0, \quad \forall I \geq 0, \forall n \geq 0. \tag{9.51}$$

Using the recurrence relation of Equation (6.1) for $j = 1$, we have that

$$\begin{aligned}\Gamma_n^{(I,K)} &= \sum_{i=0}^{I}\sum_{k=0}^{K}\left(B_n^{(i,k)} - \eta B_n^{(i-1,k)}\right)\\ &\quad - \sum_{i=0}^{I}\sum_{k=0}^{K}\left((1-\eta)B_{n+1}^{(i-1,k)} + \eta B_{n+1}^{(i,k-1)} + (1-\eta)B_n^{(i,k-1)} - B_n^{(i-1,k-1)}\right),\end{aligned}$$

$$\begin{aligned}\Gamma_n^{(I,K)} &= \eta\sum_{i=0}^{I}\sum_{k=0}^{K}\left(B_n^{(i,k)} - B_n^{(i-1,k)}\right) + (1-\eta)\sum_{i=0}^{I}\sum_{k=0}^{K}\left(B_n^{(i,k)} - B_n^{(i,k-1)}\right)\\ &\quad - (1-\eta)\sum_{i=0}^{I}\sum_{k=0}^{K}B_{n+1}^{(i-1,k)} - \eta\sum_{i=0}^{I}\sum_{k=0}^{K}B_{n+1}^{(i,k-1)} + \sum_{i=0}^{I}\sum_{k=0}^{K}B_n^{(i-1,k-1)},\end{aligned}$$

$$\begin{aligned}\Gamma_n^{(I,K)} &= \eta\sum_{k=0}^{K}B_n^{(i,k)} + (1-\eta)\sum_{i=0}^{I}B_n^{(i,k)}\\ &\quad - (1-\eta)\sum_{i=0}^{I-1}\sum_{k=0}^{K}B_{n+1}^{(i,k)} - \eta\sum_{i=0}^{I}\sum_{k=0}^{K-1}B_{n+1}^{(i,k)} + \sum_{i=0}^{I-1}\sum_{k=0}^{K-1}B_n^{(i,k)},\end{aligned}$$

$$\begin{aligned}\Gamma_n^{(I,K)} &= B_n^{(i,k)} - (1-\eta)\sum_{i=0}^{I-1}\sum_{k=0}^{K}B_{n+1}^{(i,k)}\\ &\quad - \eta\sum_{i=0}^{I}\sum_{k=0}^{K-1}B_{n+1}^{(i,k)} + \eta\sum_{i=0}^{I}\sum_{k=0}^{K-1}B_n^{(i,k)} + (1-\eta)\sum_{i=0}^{I-1}\sum_{k=0}^{K}B_n^{(i,k)},\end{aligned}$$

$$\Gamma_n^{(I,K)} = B_n^{(i,k)} + \eta\sum_{i=0}^{I}\sum_{k=0}^{K-1}\left(B_n^{(i,k)} - B_{n+1}^{(i,k)}\right) + (1-\eta)\sum_{i=0}^{I-1}\sum_{k=0}^{K}\left(B_n^{(i,k)} - B_{n+1}^{(i,k)}\right),$$

so that

$$\Gamma_n^{(I,K)} = B_n^{(i,k)} + \eta\Gamma_n^{(I,K-1)} + (1-\eta)\Gamma_n^{(I-1,K)}. \tag{9.52}$$

We know that $\Gamma_n^{(I,0)} \geq 0, \forall I \geq 0, \forall n \geq 0$, since it corresponds to a Gaussian pure-loss channel,

and was proven in the previous section. We also know, because of the symmetry of the beam splitter, that $\Gamma_n^{(0,K)} \geq 0, \forall K \geq 0, \forall n \geq 0$. We are then able to prove Equation (9.51) by using a recursive argument on both I and K, since $B_n^{(i,k)} \geq 0, \forall I \geq 0, \forall K \geq 0, \forall n \geq 0$. This shows that $\mathcal{B}_\eta^{[K]}[P_I^\downarrow]$ is passive. Again, we conclude the proof of Theorem 39 by using the fact that any passive state can be written as a convex sum of extremal-passive states $P_l^\downarrow$. □

Before discussing the implications of Theorems 38 and 39 in the context of regular majorization, we take advantage of the property of duality of extremal-passive channels in order to express the same results for passive-environment channels $\mathcal{A}_G^\downarrow$ defined as

$$\mathcal{A}_G^\downarrow[\rho] = \mathrm{Tr}_2\left[U_\lambda^{\mathrm{TMS}}\left(\rho \otimes \sigma^\downarrow\right) U_\lambda^{\mathrm{TMS}\dagger}\right], \tag{9.53}$$

for any passive environment $\sigma^\downarrow$. This is encompassed in the two following theorems, which we prove together.

Theorem 40. *Passive-environment bosonic channels $\mathcal{A}_G^\downarrow$ are Fock-majorization preserving; that is, for all states ρ and σ,*

$$\rho \succ_{\mathrm{F}} \sigma \quad \Rightarrow \quad \mathcal{A}_G^\downarrow[\rho] \succ_{\mathrm{F}} \mathcal{A}_G^\downarrow[\sigma]. \tag{9.54}$$

Theorem 41. *Passive-environment bosonic channel $\mathcal{A}_G^\downarrow$ are passive preserving; that is, if $\rho^\downarrow$ is passive then $\mathcal{A}_\eta^\downarrow[\rho^\downarrow]$ is passive as well.*

Proof. According to Theorems 38 and 39, extremal-passive channels $\mathcal{B}_\eta^{(k)}$ are passive preserving and Fock-majorization preserving. If we combine the remarks of Appendix F.2 with the duality properties of Theorems 30 and 35, we have that extremal-passive channels $\mathcal{A}_G^{(k)}$ are passive preserving and Fock-majorization preserving as well. Using the fact that any passive state can be written as a convex sum of extremal-passive states, we end up with the two theorems. □

Using Theorems 38, 39, 40 and 41 we are now able to state the following for any passive-environment Gaussian-dilatable channel $\mathcal{C}^\downarrow$ defined in Equation (7.8).

Corollary 7. *Passive-environment bosonic channels $\mathcal{C}^\downarrow$ are majorization preserving over the set of passive states; that is, for any passive states $\rho^\downarrow$ and $\sigma^\downarrow$,*

$$\rho^\downarrow \succ \sigma^\downarrow \quad \Rightarrow \quad \mathcal{C}^\downarrow[\rho^\downarrow] \succ \mathcal{C}^\downarrow[\sigma^\downarrow]. \tag{9.55}$$

Since phase-insensitive Gaussian channels are part of the set of passive-environment channels, Corollary 7 generalises Corollary 6. Our main motivation in proving Corollary 7 was provided by Conjecture 4, which is itself a consequence of the precursor of the entropy photon-number inequality stated by means of the theory of majorization. Similarly to the Gaussian

case, Conjecture 4 (generalised to all passive-environment channels) and Corollary 7 may be combined in order to state a generalised conjecture of the following form.

Conjecture 5. *Consider a passive state $\rho^{\downarrow}$, and a state σ. In this case,*

$$\text{if} \quad \rho^{\downarrow} \succ \sigma, \quad \text{then} \quad \mathcal{C}^{\downarrow}[\rho^{\downarrow}] \succ \mathcal{C}^{\downarrow}[\sigma]. \tag{9.56}$$

One readily sees that the last property is transitive by concatenation of passive-environment channels. Indeed, we expect a concatenation of two of the latter to be such that it can be written in the form of Equation (7.8) as well. Nevertheless, unlike Gaussian channels, passive-environment channels do not possess a semi-group structure of the form of Equation (3.112), as shown for instance by the decomposition in Equation (7.39). As a result, one can unfortunately not use an infinitesimal argument in order to prove Conjecture 5.

Finally, we would like to stress that all proofs in this chapter are independent of the specific nature of the system (i.e., the harmonic oscillator Hamiltonian for a bosonic mode). Therefore, we believe that the application of Fock-majorization could be extended to other quantum systems and arbitrary Hamiltonians, yielding a general tool that could prove very useful in quantum information theory, more specifically in quantum thermodynamics.

IV

Perspectives

10

A resource theory of local activity for bosonic systems

Passive states are the least energetic states for a fixed spectrum. As such, they can be understood as those states from which no work can be extracted under unitary operations. If a general quantum state is constituted of several systems, one needs to apply global unitaries to extract work in general, as local unitaries might not be sufficient in order to do so. A paradigm of such a situation is found in a state $\left(\rho^{\downarrow}\right)^{\otimes n}$, whose subsystems are described by states $\rho^{\downarrow}$, which happen to be such that they are not completely passive. Since $\rho^{\downarrow}$ is passive, no work can be extracted locally using unitary transformations, while this can be circumvented by exploiting global unitaries.

It seems natural to try and quantify the amount of work which can be extracted under local unitaries only. In order to avoid a trivial situation, one may ask whether the work extraction can be enhanced by allowing energy-preserving global unitaries. In the framework of bosonic quantum systems, a reasonable choice of energy preserving unitaries is readily found in passive Gaussian transformations, or passive interferometers.

In this chapter, building on the knowledge of Gaussian unitary transformations gathered in Part II, and inspired by the study of passive states of the harmonic oscillator and passive-environment channels introduced in Part III, we lay out the basis for a resource theory centred on the notion of passivity of a quantum state. We introduce a concept of local-activity distance, which we compare with the notion of work that can be extracted using only local unitary transformations, but also assisted by passive global Gaussian unitaries. We begin by introducing the

concept of a general resource theory in Section 10.1, before presenting the basic framework of our resource theory in Section 10.2. We also attempt to compare it with other resource theories involving passives states. In Section 10.3, we provide our definition of the local-activity distance, and compute if for a state of one mode. Finally, we hint at future directions of research in Section 10.4.

10.1 Introduction to resource theories

We begin by giving a short introduction on resource theories. For a more detailed one, the interested reader is for instance referred to [118]. As suggested by its name, the goal of a resource theory is to build a framework in which a specific resource can be characterised completely. Classical and quantum information theories can be understood as examples of such resource theories allowing one to describe the interconversions among specific resources. The best-known paradigm of a resource theory studied in quantum information concerns the interconversion of entangled states [53]. Indeed, since entanglement can be exploited in order to perform several tasks which cannot be achieved with classical resources, such as quantum teleportation [60], it qualifies as resource.

In order to build a resource theory, one may begin by choosing a set of free states, in the sense that they do not contain any resource. In the context of a resource theory of entanglement, the free states are chosen to be all separable states as defined by Equation (2.79). The next step is to choose the set of allowed operations, which may not create any resource. If they did, the theory would not be consistent as the resource would no longer be considered as such. If we go back to the example of entanglement theory, there actually happens to be several ways to choose the allowed operations. For instance, local operations with classical communications (LOCCs) introduced in Section 2.3.3 will never increase entanglement. Similarly, separable operations, which are characterised by product Kraus operators, do not increase entanglement either. LOCCs operations are actually a subclass of separable operations. Here, a choice has to be performed. However, the choice should not be independent of the set of free states. Indeed, since the allowed operations should not create any resource, free states must remain free when transformed by such operations. This is the case for separable states, which remain so after begin transformed by LOCCs, and more generally separable operations.

Since the allowed operations are characterised by the fact that they cannot increase the resource, one should be able to quantify the latter. In order to do so, one defines a measure, or quantifier of the resource, which must be monotonically non-increasing under the action of allowed operations. In the context of entanglement theory, several quantifiers have been introduced in the past decade [53]. In the case of pure quantum states, such a quantifier can be found in the entropy of entanglement defined through Equation (2.80).

Finally, let us mention that there are other aspects of resource theories which can be con-

sidered, such as the study of the interconversion between different quantum states containing different amounts of resource. However, since the goal of the present Chapter is to lay out the basis for a resource theory for bosonic systems, we do not focus on such aspects for the moment.

10.2 Basic framework

10.2.1 Free states and operations

In the most common resource theory of quantum thermodynamics, thermal states are considered to be free, in the sense that they do not contain any resource. This seems like a natural choice, as thermal states characterise systems in thermal equilibrium. In this case, the resource can simply be viewed as athermality. One then describes the allowed operations, which do not create any resource, as follows: adding ancillae in free states (which are then treated as an environment), applying energy conserving unitaries, and tracing out the environment. In this case, the thermal operations, defined as

$$\mathcal{C}_{\mathrm{TO}}(\rho_{\mathrm{S}}) = \mathrm{Tr}_{\mathrm{E}}\left[U_{\mathrm{SE}}\left(\rho_{\mathrm{S}}\otimes\tau_{\mathrm{E}}\right)U_{\mathrm{SE}}^{\dagger}\right], \tag{10.1}$$

where τ_{E} is a thermal state of the environment and U_{SE} is an energy conserving unitary acting on both the system and the environment, are the most general free operations.

In the present work, we relax this resource theory of thermal operations by allowing passive states as free states. This choice is not arbitrary, as passive states have the least possible energy for a fixed spectrum. In that case, a tensor product of passives states is considered free as well. However, if one applies an energy conserving unitary to such a product state, the resulting state might be correlated, or even entangled in general. Since this state should nevertheless still be free, one then needs to extend the set of free states, in order for energy conserving unitaries to still be considered as free. In the context of bosonic quantum systems, an interesting choice of energy conserving unitaries can be found in passive Gaussian untiaries, knowing the huge role they play in continuous-variables quantum information. By doing so, we generalise passive-environment channels involving a beam splitter in their dilation, which can always be written as

$$\mathcal{B}_{\eta}^{\downarrow}(\bullet) = \mathrm{Tr}_{2}\left[U_{\eta}^{\mathrm{BS}}\left(\bullet\otimes\sigma^{\downarrow}\right)U_{\eta}^{\mathrm{BS}\dagger}\right], \tag{10.2}$$

where $\sigma^{\downarrow}$ is a passive state and U_{η}^{BS} is a beam-splitter unitary. Since passive interferometers constitute free operations, our set of free states, which we denote as $\mathcal{I}_{\mathrm{f}}$, is now composed of all possible products of Fock-passive states, transformed by any energy conserving Gaussian unitary U^{PG} (any passive interferometer), *i.e.*,

$$\mathcal{I}_{\mathrm{f}} = \bigcup_{N\in\mathbb{N}_{0}}\{U^{\mathrm{PG}}(\sigma_{1}^{\downarrow}\otimes\ldots\otimes\sigma_{N}^{\downarrow})U^{\mathrm{PG}\dagger}\}, \tag{10.3}$$

where $\mathbb{N}_0$ is the set of all natural number not including 0. Furthermore, we also allow post-selection on free states. In that case, we need to show that a state that belongs in $\mathcal{I}_f$ should remain so after a post selection on any other state belonging in $\mathcal{I}_f$. We can actually prove this for a state of two modes, *i.e.*, we show the following.

Lemma 16. *Let $\sigma_1^\downarrow$ and $\sigma_2^\downarrow$ be any two one-mode passive states, then the one-mode state*

$$\rho = \frac{\tilde{\rho}}{\mathrm{Tr}\left[\tilde{\rho}\right]}, \qquad \text{where } \tilde{\rho} = \mathrm{Tr}_2\left[\left(\mathbb{I}\otimes\sigma_3^\downarrow\right) U_\eta^{\mathrm{BS}}\left(\sigma_1^\downarrow\otimes\sigma_2^\downarrow\right) U_\eta^{\mathrm{BS}\dagger}\right], \tag{10.4}$$

is passive as well, for any transmittance $\eta \in [0,1]$ and any one-mode passive state $\sigma_3^\downarrow$, $\mathbb{I}$ being the identity operator.

For a proof of Lemma 16, we refer the interested reader to Appendix G. Under the hypothesis that it can be extended to an arbitrary number of modes, our free operations are then constituted of the following operations:

(a) Applying any energy preserving Gaussian unitary,

(b) Allowing ancillae in free states,

(c) Partial tracing,

(d) Post-selecting on free states.

We denote the set of free operations as Λ_f.

10.2.2 Properties of the set of free states

In the case of one-mode bosonic systems, the only possible free unitaries allowed in the context of this resource theory are in fact phase rotations $U_\theta^{\mathrm{R}} = \exp[-i\theta\hat{a}^\dagger\hat{a}]$, which do not affect passive states $\sigma^\downarrow$. As a consequence, a mixture of one-mode free states is still a free state. This is however no longer the case when one considers systems involving more than one mode. Take for instance a general two-mode free state σ_η^f, which can always be written as

$$\sigma_\eta^f = U_\eta^{\mathrm{BS}}\left(\sigma_1^\downarrow\otimes\sigma_2^\downarrow\right) U_\eta^{\mathrm{BS}\dagger}, \tag{10.5}$$

where U_η^{BS} is a beam splitter, and $\sigma_1^\downarrow$ and $\sigma_2^\downarrow$ are two one-mode passive states, and consider the mixture

$$\gamma = \frac{1}{2}\left[\sigma_\eta^f + \sigma_\nu^f\right] = \frac{1}{2}\left[U_\eta^{\mathrm{BS}}\left(\sigma_1^\downarrow\otimes\sigma_2^\downarrow\right) U_\eta^{\mathrm{BS}\dagger} + U_\nu^{\mathrm{BS}}\left(\sigma_1^\downarrow\otimes\sigma_2^\downarrow\right) U_\nu^{\mathrm{BS}\dagger}\right]. \tag{10.6}$$

The covariance matrix $\mathbf{V}_\eta$ of σ_η^{f} can be computed to be

$$\mathbf{V}_\eta = \begin{pmatrix} v_{11}^{(\eta)} & 0 & v_{13}^{(\eta)} & 0 \\ 0 & v_{11}^{(\eta)} & 0 & v_{13}^{(\eta)} \\ v_{13}^{(\eta)} & 0 & v_{33}^{(\eta)} & 0 \\ 0 & v_{13}^{(\eta)} & 0 & v_{33}^{(\eta)} \end{pmatrix}, \tag{10.7}$$

where $v_{11}^{(\eta)} = \eta(2n_1+1) + (1-\eta)(2n_2+1)$, $v_{33}^{(\eta)} = \eta(2n_2+1) + (1-\eta)(2n_1+1)$, and $v_{13}^{(\eta)} = 2\sqrt{\eta(1-\eta)}(n_2-n_1)$, n_1 and n_2 being the respecting mean numbers of photons of $\sigma_1^{\downarrow}$ and $\sigma_2^{\downarrow}$. As a result, the covariance matrix $\mathbf{V}$ of the mixture γ is such that its first diagonal element v_{11} verifies

$$v_{11} = \frac{\eta+\nu}{2}(2n_1+1) + \left(1 - \frac{\eta+\nu}{2}\right)(2n_2+1), \tag{10.8}$$

while

$$v_{13} = 2\frac{\sqrt{\eta(1-\eta)} + \sqrt{\nu(1-\nu)}}{2}(n_2-n_1). \tag{10.9}$$

One understands that it will never be possible to write γ as a free state, at the level of covariance matrices. This is sufficient to state that our set of free states is not convex in general, unfortunately.

10.2.3 Comparison with resource theories of passivity

It seems interesting to compare the set of free states $\mathcal{I}_{\mathrm{f}}$ with other types of passive states, involved in other resource theories. Lenard's definition of passive states relates them with the concept of extractable work (or ergotropy) $W_{\max}$ [107], which is defined as

$$W_{\max}(\rho) = \max_U \mathrm{Tr}\left[\hat{H}\left(\rho - U\rho U^\dagger\right)\right], \tag{10.10}$$

where the maximisation is taken over all possible unitaries U. Lenard showed that a state σ will be passive if and only if $W_{\max}(\sigma) = 0$. Since the definition can involve a state of any number of modes in the framework of bosonic systems, we will call such a state *σ globally passive*. As a consequence, we define the set of globally passive states as

$$\mathcal{I}_{\mathrm{gp}} = \left\{\sigma | W_{\max}(\sigma) = 0\right\}. \tag{10.11}$$

In the same line of thought, one can define the work which is extractable locally, as

$$W_{\max}^{\mathrm{l}}(\rho) = \max_{U_{\mathrm{l}}} \mathrm{Tr}\left[\hat{H}\left(\rho - U_{\mathrm{l}}\rho U_{\mathrm{l}}^\dagger\right)\right], \tag{10.12}$$

where the maximisation is taken this time over all possible local unitaries U_{l}. We will call a state σ of N modes *locally passive* if each of its modes is in a one-mode passive state, which will be true

if and only if $W_{\max}^{\mathrm{l}}(\sigma) = 0$. We then define the set of locally passive states as

$$\mathcal{I}_{\mathrm{lp}} = \left\{\sigma | W_{\max}^{\mathrm{l}}(\sigma) = 0\right\}. \tag{10.13}$$

Since the set of all unitaries U obviously includes the set of all local unitaries U_{l}, we have that

$$\mathcal{I}_{\mathrm{gp}} \subset \mathcal{I}_{\mathrm{lp}}. \tag{10.14}$$

Furthermore, a state which is locally passive does not have to be globally passive, so that the two sets $\mathcal{I}_{\mathrm{gp}}$ and $\mathcal{I}_{\mathrm{lp}}$ are distinct. Now we showed that in the case of two modes, a state which is free according to our resource theory is always locally passive. If we suppose this result can be extended to N-modes (which we do), then a free state will always be locally passive, so that

$$\mathcal{I}_{\mathrm{f}} \subset \mathcal{I}_{\mathrm{lp}}. \tag{10.15}$$

However, the two sets are again distinct, since one can always find a state which is locally passive, but not free. An example of such a state is given by $\gamma = U_{\lambda}^{\mathrm{TMS}}\left(\sigma_1^{\downarrow} \otimes \sigma_2^{\downarrow}\right) U_{\lambda}^{\mathrm{TMS}\dagger}$, where $\sigma_1^{\downarrow}$ and $\sigma_2^{\downarrow}$ are two one-mode passive states, and $U_{\lambda}^{\mathrm{TMS}}$ is a two-mode squeezer unitary. Indeed, γ can be shown to be locally passive, even though it is not free. Finally, the two sets $\mathcal{I}_{\mathrm{f}}$ and $\mathcal{I}_{\mathrm{gp}}$ are distinct, and neither of them is included in the other. One can always find a state which is in $\mathcal{I}_{\mathrm{f}}$ while not being in $\mathcal{I}_{\mathrm{gp}}$, and vice-versa. Note that a product of one-mode thermal Gaussian states is comprised in both $\mathcal{I}_{\mathrm{f}}$ and $\mathcal{I}_{\mathrm{gp}}$.

In the hope of introducing a definition of work which is consistent with the resource theory of free states contained in $\mathcal{I}_{\mathrm{f}}$, one may consider what could be understood as a halfway situation. Define the locally extractable work under passive Gaussian unitaries of a state ρ as

$$W_{\max}^{\mathrm{l,PG}}(\rho) = \max_{U_{\mathrm{l}}} \max_{U^{\mathrm{PG}}} \mathrm{Tr}\left[\hat{H}\left(\rho - U_{\mathrm{l}} U^{\mathrm{PG}} \rho U^{\mathrm{PG}\dagger} U_{\mathrm{l}}{}^{\dagger}\right)\right], \tag{10.16}$$

where the maximisation is taken this time over all possible passive Gaussian unitaries U^{PG}, as well as all local unitaries U_{l}. This can be viewed as the work that is extractable locally, *assisted* by passive Gaussian global unitaries. In other words, one may apply global Gaussien unitaries without modifying the energy, before actually extracting work locally. Now, define the set

$$\mathcal{I}_{\mathrm{lp}}^{\mathrm{GP}} = \left\{\sigma | W_{\max}^{\mathrm{l,PG}}(\sigma) = 0\right\}. \tag{10.17}$$

One readily understands that

$$\mathcal{I}_{\mathrm{gp}} \subset \mathcal{I}_{\mathrm{lp}}^{\mathrm{GP}} \subset \mathcal{I}_{\mathrm{lp}}. \tag{10.18}$$

Furthermore, the three sets $\mathcal{I}_{\mathrm{gp}}$, $\mathcal{I}_{\mathrm{lp}}^{\mathrm{GP}}$ and $\mathcal{I}_{\mathrm{lp}}$ are clearly distinct. Note that another way to write

$\mathcal{I}_{\text{lp}}^{\text{GP}}$ would be

$$\mathcal{I}_{\text{lp}}^{\text{GP}} = \left\{\sigma | W_{\text{max}}^{1}\left(U^{\text{PG}}\sigma U^{\text{PG}\dagger}\right) = 0, \quad \forall\, U^{\text{PG}}\right\}. \tag{10.19}$$

Proof. Trivially,

$$\begin{aligned}
& W_{\text{max}}^{1}\left(U^{\text{PG}}\sigma U^{\text{PG}\dagger}\right) = 0, \quad \forall\, U^{\text{PG}} \\
\Leftrightarrow \quad & \max_{U_1} \text{Tr}\left[\hat{H}\left(U^{\text{PG}}\sigma U^{\text{PG}\dagger} - U_1 U^{\text{PG}}\sigma U^{\text{PG}\dagger} U_1{}^{\dagger}\right)\right] = 0, \quad \forall\, U^{\text{PG}} \\
\Leftrightarrow \quad & \max_{U_1} \text{Tr}\left[\hat{H}\left(\sigma - U_1 U^{\text{PG}}\sigma U^{\text{PG}\dagger} U_1{}^{\dagger}\right)\right] = 0, \quad \forall\, U^{\text{PG}} \\
\Leftrightarrow \quad & \max_{U^{\text{PG}}} \max_{U_1} \text{Tr}\left[\hat{H}\left(\sigma - U_1 U^{\text{PG}}\sigma U^{\text{PG}\dagger} U_1{}^{\dagger}\right)\right] = 0 \\
\Leftrightarrow \quad & W_{\text{max}}^{1,\text{PG}}(\sigma) = 0,
\end{aligned}$$

where we used the fact that U^{PG} commutes with the Hamiltonian. The third line is equivalent to the second because of the fact that

$$\max_{U_1} \text{Tr}\left[\hat{H}\left(\sigma - U_1 U^{\text{PG}}\sigma U^{\text{PG}\dagger} U_1{}^{\dagger}\right)\right] = \max_{U_1} \text{Tr}\left[\hat{H}\left(U^{\text{PG}}\sigma U^{\text{PG}\dagger} - U_1 U^{\text{PG}}\sigma U^{\text{PG}\dagger} U_1{}^{\dagger}\right)\right] \tag{10.20}$$

is always non-negative. □

One could ask the question whether our set of free states $\mathcal{I}_{\text{f}}$ is the same as $\mathcal{I}_{\text{lp}}^{\text{GP}}$. A direct answer is found in the fact that, as already explained, $\mathcal{I}_{\text{f}}$ is not convex, while $\mathcal{I}_{\text{lp}}^{\text{GP}}$ can easily be proved to be convex, as a simple consequence of the fact the quantity $W_{\text{max}}^{1,\text{PG}}$ is itself convex. Unfortunately, it means that $\mathcal{I}_{\text{f}} \neq \mathcal{I}_{\text{lp}}^{\text{GP}}$. Furthermore, $\mathcal{I}_{\text{gp}} \subset \mathcal{I}_{\text{lp}}^{\text{GP}}$, while $\mathcal{I}_{\text{gp}} \not\subset \mathcal{I}_{\text{f}}$, so that $\mathcal{I}_{\text{lp}}^{\text{GP}} \not\subset \mathcal{I}_{\text{f}}$. Now, since the concatenation of two passive Gaussian unitaries gives another passive Gaussian unitary, and using Equation (10.15) (which is supposed to be true for N modes), we have that $\mathcal{I}_{\text{f}} \subset \mathcal{I}_{\text{lp}}^{\text{GP}}$. In order to circumvent the fact that $\mathcal{I}_{\text{f}}$ is not convex, one can take is convex hull $\text{conv}(\mathcal{I}_{\text{f}})$. Using similar arguments as before, $\text{conv}(\mathcal{I}_{\text{f}}) \subset \mathcal{I}_{\text{lp}}^{\text{GP}}$. What remains to be proven is whether $\text{conv}(\mathcal{I}_{\text{f}}) \stackrel{?}{=} \mathcal{I}_{\text{lp}}^{\text{GP}}$. All of this is summarised on Figure 10.2.1.

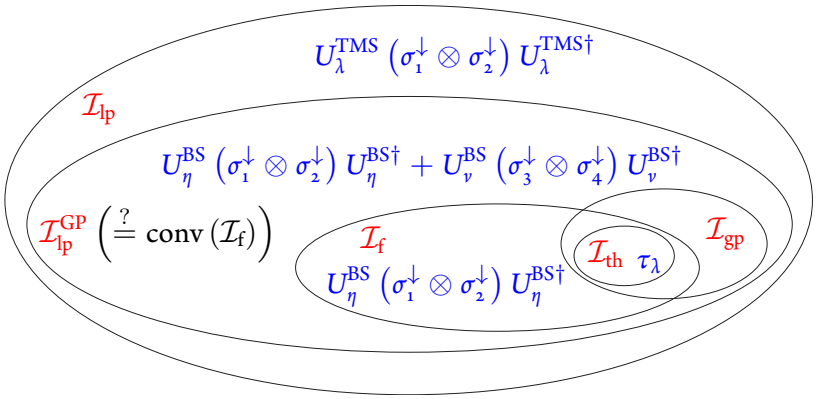

Figure 10.2.1: Relation between the different theories of "activity".

10.3 Local-activity distance

In this section, we introduce a definition of local-activity distance, which we will compare with the locally extractable work assisted by passive Gaussian unitaries in Section 10.4.

10.3.1 Definition of the local-activity distance

We call local-activity distance any "smallest distance" to a free state which never increases under free operations. In other words, the local-activity distance is a monotone of the free operations of our resource theory. It can be quantified using any contractive distance D as

$$A_l(\rho) = \min_{\sigma \in \mathcal{I}_f} D(\rho, \sigma). \tag{10.21}$$

By contractive distance, we mean that

$$D(A_l[\rho], A_l[\sigma]) \leq D(\rho, \sigma), \tag{10.22}$$

for any free operation A_l.

Proof. Suppose the minimum of the distance is achieved at some free state σ^*. Then,

$$\begin{aligned} A_l(\rho) &= S(\rho \parallel \sigma^*) \\ &\geq S(\Lambda_f[\rho] \parallel \Lambda_f[\sigma^*]) \\ &= S(\Lambda_f[\rho] \parallel \tilde{\sigma}]) \\ &\geq \min_{\sigma \in \mathcal{I}_f} S(\Lambda_f[\rho] \parallel \sigma]) \\ &= A_l(\Lambda_f[\rho]), \end{aligned}$$

where $\tilde{\sigma} = \Lambda_f[\sigma^*]$ is a free state. □

A natural choice of distance is the quantum relative entropy $S(\bullet \parallel \bullet)$, as it is contractive. From now on, we define the local-activity distance of any state ρ to be

$$A_l(\rho) = \min_{\sigma \in \mathcal{I}_f} S(\rho \parallel \sigma). \tag{10.23}$$

10.3.2 Properties of the local-activity distance

We prove the following properties of the local-activity distance.

Property 29 (Convexity of the local-activity distance for one mode)**.** *Consider a one-mode state $\rho = \sum_{i=1}^{d} p_i \rho_i$. We have*

$$A_l\left(\sum_{i=1}^{d} p_i \rho_i\right) \leq \sum_{i=1}^{d} p_i A_l(\rho_i). \tag{10.24}$$

Proof. Let us consider that

$$A_l\left(\rho_i\right) = \min_{\sigma\in\mathcal{I}_f} S\big(\rho_i \,\big\|\, \sigma\big) = S\big(\rho_i \,\big\|\, \sigma_i^*\big), \quad \forall i. \tag{10.25}$$

Then,

$$\begin{aligned}
\sum_{i=1}^{d} p_i A_l\left(\rho_i\right) &= \sum_{i=1}^{d} p_i S\big(\rho_i \,\big\|\, \sigma_i^*\big) \\
&\geq S\left(\sum_{i=1}^{d} p_i\rho_i \,\middle\|\, \sum_{i=1}^{d} p_i\sigma_i^*\right) \\
&= S\left(\sum_{i=1}^{d} p_i\rho_i \,\middle\|\, \sigma^*\right) \\
&\geq \min_{\sigma\in\mathcal{I}_f} S\left(\sum_{i=1}^{d} p_i\rho_i \,\middle\|\, \sigma\right) \\
&= A_l\left(\sum_{i=1}^{d} p_i\rho_i\right).
\end{aligned}$$

We used the fact that a mixture of passive state $\sum_{i=1}^{d} p_i\sigma_i^* = \sigma^*$ is still passive. □

Property 30 (Sub-additivity of the local-activity distance)**.** *Let $\rho^{(1)}$ and $\rho^{(2)}$ be two quantum states, then*

$$A_l\left(\rho^{(1)}\otimes\rho^{(2)}\right) \leq A_l\left(\rho^{(1)}\right) + A_l\left(\rho^{(2)}\right). \tag{10.26}$$

Proof. Let us first consider that

$$A_l\left(\rho^{(1)}\right) = \min_{\sigma\in\mathcal{I}_f} S\big(\rho^{(1)} \,\big\|\, \sigma\big) = S\big(\rho^{(1)} \,\big\|\, \tilde{\sigma}^{(1)}\big), \tag{10.27}$$

and

$$A_l\left(\rho^{(2)}\right) = \min_{\sigma\in\mathcal{I}_f} S\big(\rho^{(2)} \,\big\|\, \sigma\big) = S\big(\rho^{(2)} \,\big\|\, \tilde{\sigma}^{(2)}\big). \tag{10.28}$$

Now,

$$\begin{aligned}
A_l\left(\rho^{(1)}\right) + A_l\left(\rho^{(2)}\right) &= S\big(\rho^{(1)} \,\big\|\, \tilde{\sigma}^{(1)}\big) + S\big(\rho^{(2)} \,\big\|\, \tilde{\sigma}^{(2)}\big) \\
&= S\big(\rho^{(1)}\otimes\rho^{(2)} \,\big\|\, \tilde{\sigma}^{(1)}\otimes\tilde{\sigma}^{(2)}\big) \\
&= S\big(\rho^{(1)}\otimes\rho^{(2)} \,\big\|\, \sigma^*\big) \\
&\geq \min_{\sigma\in\mathcal{I}_f} S\big(\rho^{(1)}\otimes\rho^{(2)} \,\big\|\, \sigma\big) \\
&= A_l\left(\rho^{(1)}\otimes\rho^{(2)}\right),
\end{aligned}$$

where $\sigma^* = \tilde{\sigma}^{(1)} \otimes \tilde{\sigma}^{(2)}$ is some free state. Thus,

$$A_l\left(\rho^{(1)} \otimes \rho^{(2)}\right) \leq A_l\left(\rho^{(1)}\right) + A_l\left(\rho^{(2)}\right). \tag{10.29}$$

□

10.3.3 Calculation of the local-activity distance for a single mode

We are now going to compute the local-activity distance for a one-mode state, by finding an expression for its closest free state. As we already mentioned, in the case of one mode, the only energy preserving Gaussian unitary is the phase rotation, which does not modify a passive state. Thus, the local-activity distance of a one-mode state ρ is given by

$$A_l(\rho) = \min_{\sigma^{\downarrow}} S\big(\rho \,\big\|\, \sigma^{\downarrow}\big), \tag{10.30}$$

where the minimum is taken over all one-mode passive states. This minimisation problem can actually be solved analytically. Let $\sigma^{\downarrow} = \sum_{n=0}^{M-1} s_n |n\rangle\langle n|$, then

$$A_l(\rho) = \min_{\{s_n\}} \left[-S(\rho) - \sum_{n=0}^{M-1} r_n \ln s_n\right], \tag{10.31}$$

where $r_n = \langle n|\rho|n\rangle$. Since there are constraints on $\sigma^{\downarrow}$, namely, $(s_{n+1} - s_n) \leq 0$ (for $0 \leq n \leq M-2$) and $\sum_{n=0}^{M-1} s_n = 1$, we will minimise the quantity $\mathcal{L}^{\mathrm{KKT}}(\mathbf{s})$ defined as

$$\mathcal{L}^{\mathrm{KKT}}(\mathbf{s}) = \left[-S(\rho) - \sum_{n=0}^{M-1} r_n \ln s_n\right] + \sum_{n=0}^{M-2} \mu_n (s_{n+1} - s_n) + \bar{\mu}\left(\sum_{n=0}^{M-1} s_n - 1\right). \tag{10.32}$$

If $\mathbf{s}^*$ is a local minimum (and $(s_{n+1} - s_n)$ are continuously differentiable, which is the case here) then there exist Karush–Kuhn–Tucker (KKT) coefficients such that:

1. Stationarity: $\frac{\partial}{\partial s_n}\mathcal{L}^{\mathrm{KKT}}(\mathbf{s}) = 0$ for $n = 0, \ldots, M-1$,
2. Primal feasibility: $s_{n+1} - s_n \leq 0$ and $\sum_{n=0}^{M-1} s_n = 1$,
3. Dual feasibility: $\mu_n \geq 0$ for $n = 0, \ldots, M-1$,
4. Complementary slackness: $\mu_n (s_{n+1} - s_n) = 0$ for $n = 0, \ldots, M-2$.

Now, from stationarity, we have

$$0 = -\frac{r_n}{s_n} + \mu_{n-1} - \mu_n + \bar{\mu}, \quad \text{for} \quad n = 1, 2, \ldots, M-2, \tag{10.33}$$

or,

$$s_n = \frac{r_n}{\mu_{n-1} - \mu_n + \bar{\mu}}. \tag{10.34}$$

For $n = 0$ and $n = M - 1$, we have, respectively,

$$0 = -\frac{r_0}{s_0} - \mu_0 + \bar{\mu} \quad \Rightarrow \quad s_0 = \frac{r_0}{\bar{\mu} - \mu_0}, \tag{10.35}$$

and

$$0 = -\frac{r_{M-1}}{s_{M-1}} + \mu_{M-2} + \bar{\mu} \quad \Rightarrow \quad s_{M-1} = \frac{r_{M-1}}{\bar{\mu} + \mu_{M-2}}. \tag{10.36}$$

From complementary slackness, we have

$$\mu_n (s_{n+1} - s_n) = 0 \quad \text{for} \quad n = 0, \ldots, M-2. \tag{10.37}$$

The solution of above equations is nontrivial and can be discussed geometrically, more appropriately. Before we go into details, let us consider a simple example where $M = 3$. In this case, the above conditions look like.

$$s_0 = \frac{r_0}{\bar{\mu} - \mu_0}, \tag{10.38}$$

$$s_1 = \frac{r_1}{\mu_0 - \mu_1 + \bar{\mu}}, \tag{10.39}$$

$$s_2 = \frac{r_2}{\mu_1 + \bar{\mu}}. \tag{10.40}$$

From complementary slackness conditions $\mu_0 (\sigma_1 - s_0) = 0$ and $\mu_1 (\sigma_2 - s_1) = 0$, we end up with the following four cases:

Case 1: $\mu_0 = 0$ and $\mu_1 = 0$. In this case, we assume $s_1 - s_0 \leq 0$ and $s_2 - s_1 \leq 0$. Then,

$$s_0 = \frac{r_0}{\bar{\mu}}, \tag{10.41}$$

$$s_1 = \frac{r_1}{\bar{\mu}}, \tag{10.42}$$

$$s_2 = \frac{r_2}{\bar{\mu}}. \tag{10.43}$$

Therefore, $\bar{\mu} = 1$, $s_0 = r_0$, $s_1 = r_1$, $s_2 = r_2$ and $r_1 - r_0 \leq 0$ and $r_2 - r_1 \leq 0$.

Case 2: $\mu_0 = 0$ and $s_1 = s_2$. In this case, $s_1 - s_0 \leq 0$ is assumed. Then, we have

$$-\frac{r_0}{s_0} + \bar{\mu} = 0, \tag{10.44}$$

$$-\frac{r_1}{s_1} - \mu_1 + \frac{r_0}{s_0} = 0, \tag{10.45}$$

$$-\frac{r_2}{s_2} + \mu_1 + \frac{r_0}{s_0} = 0. \tag{10.46}$$

This leads to

$$s_1 = \frac{s_0(r_1 + r_2)}{2r_0}. \tag{10.47}$$

From normalisation $s_0 + 2s_1 = 1$, we have

$$s_0 = r_0. \tag{10.48}$$

This implies, $\bar{\mu} = 1$. The solutions are given by

$$s_0 = r_0, \quad s_1 = \frac{r_1 + r_2}{2} = s_2, \quad \bar{\mu} = 1, \quad \mu_0 = 0, \quad \mu_1 = \frac{r_2 - r_1}{r_2 + r_1}. \tag{10.49}$$

Case 3: $s_0 = s_1$ and $\mu_1 = 0$. In this case, $s_2 - s_1 \leq 0$ is assumed. Then, we have

$$-\frac{r_0}{s_0} - \mu_0 + \bar{\mu} = 0, \tag{10.50}$$

$$-\frac{r_1}{s_1} + \mu_0 + \bar{\mu} = 0, \tag{10.51}$$

$$-\frac{r_2}{s_2} + \bar{\mu} = 0. \tag{10.52}$$

The solution is given by

$$s_0 = \frac{r_0 + r_1}{2} = s_1, \quad s_2 = r_2, \quad \bar{\mu} = 1, \quad \mu_0 = \frac{r_1 - r_0}{r_1 + r_0}, \quad \mu_1 = 0. \tag{10.53}$$

Case 4: $s_0 = s_1$ and $s_1 = s_2$. Then, we have $s_0 = s_1 = s_2 = 1/3$ and

$$\bar{\mu} = 2(r_1 + r_2), \quad \mu_0 = -3r_0 + 2(r_1 - r_2), \quad \mu_1 = 2(r_2 - r_1). \tag{10.54}$$

In the general case, we can compute the closest passive state starting from ρ, iteratively as follows. Suppose the diagonal elements of ρ in the Fock basis are given by $\{r_0, \ldots, r_{M-1}\}$. We go through the diagonal starting from r_0. When we reach an element r_k which is smaller than the following one, *i.e.*, $r_k \leq r_{k+1}$, we compute the average $r_k^{(1)} = (r_k + r_{k+1})/2$ of the two, and set $r_k = r_{k+1} = r_k^{(1)}$. If we still have $r_{k-1} \leq r_k = r_k^{(1)}$, we take the average of $r_k = r_k^{(1)}$, $r_{k+1} = r_k^{(1)}$, r_{k-1}, *i.e.*, $r_k^{(2)} = (2r_k^{(1)} + r_{k-1})/3$, and set $r_k = r_{k+1} = r_{k-1} = r_k^{(2)}$. We continue until we reach a situation in which we have $r_l \geq r_{l+1}$ for $l = 0, \ldots, k$. When this is the case, we keep going through the new vector starting from $l = k + 1$, and do the same until we reach a passive state. The latter is the closest passive state to the initial state ρ in terms of the local-activity distance we defined.

Example 2. Consider the state $\rho = |M\rangle \langle M|$. The diagonal elements of ρ in the Fock basis

are given by $\{\overbrace{0,\ldots,0}^{M-times},1,0,\ldots,0\}$. The closest passive state is then given by

$$\sigma^{\downarrow}=\frac{1}{M+1}\sum_{i=0}^{M}|i\rangle\langle i|\,. \tag{10.55}$$

The local-activity distance is then computed to be $A_l\left(|M\rangle\langle M|\right)=\ln(M+1)$.

Obviously, the local-activity distance is different from the locally extractable work in the case of one mode. The two notions also behaves differently, as seen from their properties. We discuss these differences in Section 10.4.

10.4 Future directions of research

As already mentioned, the local-activity distance A_l is different from the work $W_{\max}^{l,\mathrm{PG}}$ which can be extracted from a quantum state under local unitaries, assisted by passive global unitaries (Equation (10.16)). In the case of product states, we showed that the distance A_l is sub-additive, while $W_{\max}^{l,\mathrm{PG}}$ is super-additive, since

$$\begin{aligned}W_{\max}^{l,\mathrm{PG}}(\rho_1\otimes\rho_2)&\geq\max_{U_l}\max_{U_1^{\mathrm{PG}},U_2^{\mathrm{PG}}}\mathrm{Tr}\left[\hat{H}\left(\rho_1\otimes\rho_2-U_lU_1^{\mathrm{PG}}\otimes U_2^{\mathrm{PG}}\rho_1\otimes\rho_2U_1^{\mathrm{PG}\dagger}\otimes U_2^{\mathrm{PG}\dagger}U_l{}^{\dagger}\right)\right]\\&=W_{\max}^{l,\mathrm{PG}}(\rho_1)+W_{\max}^{l,\mathrm{PG}}(\rho_2).\end{aligned}$$

This is not a surprise, since the local-activity distance is the relative-entropy distance to the closest free state. In the case of one mode, it is given by the distance with the closest passive state. The extractable work $W_{\max}^{l,\mathrm{PG}}(\rho)$ of a state ρ, however, can be interpreted as the energy "distance" with the corresponding passive state $\rho^{\downarrow}$. It may be interesting to find a connection between these two resource quantifiers.

Furthermore, as we already explained, the set of free states $\mathcal{I}_{\mathrm{f}}$ we introduced is not convex. Since our interpretation of the resource is related to the energy of a quantum state, one understands that mixing free states should not create any resource. Furthermore, such a mixture does not put different subsystems in contact in general, so that the role played by the passive global unitary is different from the mixing operation in the context of our theory. In other words, mixing free states should not activate the extraction of work. This can be witnessed when considering globally passive states, as a mixture of passive states stays passive, as exhibited by Property 29 (with our definition of the notion of local-activity distance) for one-mode free states, which are actually passive. In conclusion, a mixture of two free states should remain free. This can be achieved here by considering the convex closure of $\mathcal{I}_{\mathrm{f}}$, as already hinted at earlier. In that case, one possible direction of research is to prove whether $\mathrm{conv}(\mathcal{I}_{\mathrm{f}})\stackrel{?}{=}\mathcal{I}_{\mathrm{lp}}^{\mathrm{GP}}$.

11

Conclusions and future work

In quantum optics, the so-called phase-space representation provides us with a convenient way to describe states of the electromagnetic field as well as common optical components. In quantum information theory, one often needs to investigate information-theoretic characteristics of quantum states, especially the way they behave when processed through quantum channels. This is for instance the case for the von Neumann entropy of states, which is arguably the central notion of the field. One then turns to the state-space representation, based on the density operator of quantum systems, in order to achieve such a task. At the crossroad of quantum optics and quantum information theory lies continuous-variable quantum information theory, in which one undertakes the study of information-theoretic features of optical systems, and more generally bosonic systems. As long as Gaussian states and transformations are the sole objects involved, this analysis does not give rise to any specific difficulty. Indeed, the symplectic formalism of phase space allows one to fully characterise Gaussian states by means of their first two statistical moments only, while Gaussian transformations are described by simple affine mappings. The von Neumann entropy of Gaussian states, for instance, has a simple closed expression. This is however no longer the case whenever either of the state or the transformation in question is not Gaussian, in which case the symplectic formalism cannot be exploited any more.

The present thesis is dedicated to the development of new tools which allow us to leave the realm of Gaussian states and transformations. The prime motivation for such a purpose results from the fact that one often needs access to non-Gaussian resources in order to perform quan-

tum information processing tasks of importance. Indeed, universal quantum computing, quantum entanglement distillation and quantum error correction have been demonstrated to require non-Gaussian states or transformations in order to be achieved. As a first step towards our goal, we focus on objects which still involve Gaussian unitary operations. In doing so, it turns out we are still able to exploit the symplectic formalism in phase-space, while we end up accessing intrinsically non-Gaussian features of bosonic systems.

By bringing the generating function of non-Gaussian objects into the picture, we showed in the first part of this thesis how they could be characterised by means of the symplectic formalism specific to Gaussian systems. Specifically, we applied our method in order to describe the matrix elements of Gaussian unitaries in the Fock basis. In the case of two-mode bosonic systems, we were able to derive surprisingly simple recurrence relations for the transition probabilities in both a beam splitter and a two-mode squeezer, regardless of their complexity. Interestingly, our equations bring forth the effect of quantum interferences resulting from the indistinguishability of bosons in the form of suppression terms. In the case of a balanced beam splitter, our equations recover the so-called Hong-Ou-Mandel effect undergone by two indistinguishable photons entering the beam splitter. They also generalise the Hong-Ou-Mandel effect by describing the interference effects taking place in a situation involving multiple photons. Unexpectedly, our relations exhibit the existence of similar suppression effects in an active transformation in the form of a two-mode squeezer. The phenomenon can nicely be interpreted as the indistinguishability between an incoming photon pair that is absorbed in the two-mode squeezer and a photon pair which is created in the latter. Finally, we generalised our method to the description of multiple-mode passive Gaussian transformations. We demonstrated the existence of a relation that turns out to be quite elegant, even though it involves quantities as complex as multi-mode transition probabilities in Fock space. We anticipate that our multi-mode equation can be exploited in order to uncover situations in which the parameters of the passive transformation may give rise to total suppression events, generalising the Hong-Ou-Mandel effect to multiple modes.

The second part of our thesis was concerned with the description of Gaussian-dilatable channels involving a passive environment. Building on the generating function of extremal-passive channels, we derived a decomposition of these maps in terms of quantum-limited Gaussian channels and studied their duality properties. We then hinted at their importance in the context of the entropy photon-number inequality. This inspired us to investigate majorization relations in passive-environment channels. In recent years, the algebraic theory of majorization has been shown to naturally appear in the context of various topics related to quantum mechanics. The Shannon and von Neumann entropies being the central notions of information theory and its quantum counterpart, their connection to majorization suggests that the latter is ideal for

the study of information processing tasks. This is for instance illustrated in continous-variable quantum information by the proofs of the minimum output entropy conjectures for Gaussian channels [78, 81, 103, 116, 117].

Our aim was to study whether majorization relations are conserved in passive-environment channels $\mathcal{C}^{\downarrow}$, which led us to introduce the new pre-order relation of Fock-majorization. Unlike regular majorization, it turned out to be conserved in passive-environment channels $\mathcal{C}^{\downarrow}$, *i.e.*,

$$\rho \succ_{\mathrm{F}} \sigma \quad \Rightarrow \quad \mathcal{C}^{\downarrow}[\rho] \succ_{\mathrm{F}} \mathcal{C}^{\downarrow}[\sigma]. \tag{11.1}$$

As a side product, this also implies

$$\rho^{\downarrow} \succ \sigma^{\downarrow} \quad \Rightarrow \quad \mathcal{C}^{\downarrow}[\rho^{\downarrow}] \succ \mathcal{C}^{\downarrow}[\sigma^{\downarrow}], \tag{11.2}$$

for passive states. This line of research – and in particular the last equation – was inspired by our conjectured precursor for the entropy photon-number inequality, which boils down to

$$\Phi_{\eta}\left[\rho_a^{\downarrow}, \rho_b^{\downarrow}\right] \succ \Phi_{\eta}\left[\rho_a, \rho_b\right], \tag{11.3}$$

where $\Phi_{\eta}\left[\rho_a, \rho_b\right] = \mathrm{Tr}_2\left[U_{\eta}^{\mathrm{BS}}\left(\rho_a \otimes \rho_b\right) U_{\eta}^{\mathrm{BS}\dagger}\right]$. As a future direction of research, one of our main goals is to prove Equation (11.3). If shown, the latter would considerably simplify the proof of the original entropy photon-number inequality. Indeed, the precursor of the EPnI naturally implies that

$$S\Big(\Phi_{\eta}\left[\rho_a^{\downarrow}, \rho_b^{\downarrow}\right]\Big) \leq S\Big(\Phi_{\eta}\left[\rho_a, \rho_b\right]\Big). \tag{11.4}$$

This basically means that for any couple of states ρ_a and ρ_b at the input of the map Φ_{η}, the two corresponding passive states will always do better, in the sense that they will gives a smaller entropy at the output of the channel. One would then only need to investigate passive states in order to prove the entropy photon-number inequality. Indeed, if one demonstrates that among all passive state with fixed entropies at the two inputs of the channel Φ_{η}, thermal states give the minimum entropy at its output, the reasoning would be that

$$S\Big(\Phi_{\eta}\left[\tau_a, \tau_b\right]\Big) \leq S\Big(\Phi_{\eta}\left[\rho_a^{\downarrow}, \rho_b^{\downarrow}\right]\Big) \leq S\Big(\Phi_{\eta}\left[\rho_a, \rho_b\right]\Big), \tag{11.5}$$

where τ_a and τ_b are thermal states with the same entropies as ρ_a and ρ_b. This last step may be easier than proving the original EPnI since we deal with Fock-diagonal states only, hence it is an intrinsically classical problem.

Finally, the last part of our thesis was centred on laying the foundation of a resource theory related to the work that can be extracted using local unitaries assisted by global passive unitaries.

We introduced a notion of local-activity distance, and compared it with the latter. We also compared our set of free states with free states in other resource theories involving passive states. The current status of our work concerns the fact that our set of free states is not convex and should be extended using its convex closure. With this in mind, one of our future directions of research consists in further refining this resource theory of passivity.

Appendix

A Continuous majorization

A.1 Alternative definition of the rearrangement of function

In Section 2.2.3 of Chapter 2, we introduced the concept of spherically decreasing symmetric rearrangement of a non-negative function, which we chose to define in terms of the rearrangement of a Borel set. There actually exists a different way to define the rearrangement, which we present here for completeness. The definition we give in this section can also be used to compare two non-negative functions in terms of majorization. There are however some slight differences we discuss hereafter.

Let $(\mathcal{A}, \mathcal{F}, \nu)$ be a measure space. Consider a non-negative ν-integrable function f defined on $\mathcal{A}$, and write [5]

$$m_f(t) = \nu\left(\{x : f(x) > t\}\right), \quad t \geq 0. \tag{A.1}$$

This function generalises Equation (2.41) defined in Chapter 2, as it represents the Lebesgue measure of the set of elements x such that $f(x) > t$. Using this, we define the decreasing rearrangement as follows.

Definition 32 (Decreasing rearrangement of a non-negative function)**.** *For a non-negative integrable function f defined on $(\mathcal{A}, \mathcal{F}, \nu)$, one can define its decreasing rearrangement $f_\downarrow$ by*

$$f_\downarrow(u) = \sup\{t : m_t(t) > u\}, \quad 0 \leq u \leq \nu(\mathcal{A}). \tag{A.2}$$

Notice that we chose to denote by $f_\downarrow$ the rearrangement in this section, while it was denoted as $f^\downarrow$ in Section 2.2.3. The first difference between the two rearrangements resides in the fact that $f^\downarrow$ lives in the same space as f, while $f_\downarrow$ is defined in a new space and is always a function of an argument $u \in [0, \infty)$. Furthermore, $f^\downarrow$ and $f_\downarrow$ are both decreasing, but $f^\downarrow$ is symmetric and "spherical" (as its name suggests), while $f_\downarrow$ can never be as such. Even when f is defined on $\mathbb{R}$, $f^\downarrow$ will be defined on $\mathbb{R}$ while $f^\downarrow$ will be defined on $[0, \infty)$. A consequence of this is that f and $f_\downarrow$ can never be compared, while this can be done for f and $f^\downarrow$, like in Theorem 10 for instance.

In this context, majorization can be defined as follows [119].

Definition 33 (Continuous majorization)**.** *For two non-negative integrable function f and g defined on $(\mathcal{A}, \mathcal{F}, \nu)$ such that $\int f \, \mathrm{d}\nu = \int g \, \mathrm{d}\nu$, we say that f majorizes g, i.e. $f \succ g$ if*

$$\int_0^t f_\downarrow(u)\mathrm{d}\nu(u) \geq \int_0^t g_\downarrow(u)\mathrm{d}\nu(u), \quad \forall\, t \in [0, \nu(\mathcal{A})). \tag{A.3}$$

This definition is very similar to Definition 15. Interestingly, majorization can also be defined in this context using function $m_f(t)$ of Equation (A.1). This can be done as such [119].

Property 31. *For two non-negative integrable function f and g defined on $(\mathcal{A}, \mathcal{F}, \nu)$ such that $\int f \mathrm{d}\nu =$*

$\int g \, \mathrm{d}\nu$, *f majorizes g, i.e.* $f \succ g$ *if and only if*

$$\int_t^\infty m_f(s)\mathrm{d}s \geq \int_t^\infty m_g(s)\mathrm{d}s, \quad \forall\, t \geq 0. \tag{A.4}$$

A.2 Proof of Lemma 2

According to Lemma 2 of Section 2.2.3, If f and g are probability densities, and there exists a distribution $K : \mathbb{R} \to \mathbb{R}_{\geq 0}$ such that

$$\int_{-\infty}^{\infty} \mathrm{d}y\, K(y) = 1, \tag{A.5}$$

and

$$g(x) = \int_{-\infty}^{\infty} \mathrm{d}y\, K(y) f_y(x), \quad \forall x \in \mathbb{R}, \tag{A.6}$$

where $f_y \equiv f$, or $f_y^{\downarrow} = f^{\downarrow}$, for all $y \in \mathbb{R}$, then $f \succ g$.

Proof. Using Jensen's inequality, we have that for any convex function φ,

$$\begin{aligned}
\int_{-\infty}^{\infty} \mathrm{d}x\, \varphi\left[g(x)\right] &= \int_{-\infty}^{\infty} \mathrm{d}x\, \varphi\left[\int_{-\infty}^{\infty} \mathrm{d}y\, K(y) f_y(x)\right] \\
&\leq \int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}y\, K(y) \varphi\left[f_y(x)\right] \\
&= \int_{-\infty}^{\infty} \mathrm{d}y\, K(y) \int_{-\infty}^{\infty} \mathrm{d}x\, \varphi\left[f_y(x)\right].
\end{aligned}$$

Now, according to Theorem 10,

$$\int_{-\infty}^{\infty} \mathrm{d}x\, \varphi\left[f_y(x)\right] = \int_{-\infty}^{\infty} \mathrm{d}x\, \varphi\left[f(x)\right], \quad \forall y \in \mathbb{R}, \tag{A.7}$$

so that

$$\begin{aligned}
\int_{-\infty}^{\infty} \mathrm{d}x\, \varphi\left[g(x)\right] &\leq \int_{-\infty}^{\infty} \mathrm{d}y\, K(y) \int_{-\infty}^{\infty} \mathrm{d}x\, \varphi\left[f(x)\right] \\
&= \int_{-\infty}^{\infty} \mathrm{d}x\, \varphi\left[f(x)\right].
\end{aligned}$$

Using Theorem 11, we end up with $f \succ g$. □

B The Wigner function of a one-mode quantum state

In Section 3.2 of Chapter 3, we introduced the concept of Wigner function of a quantum state. In the present appendix, our goal is to go a bit more into details in the case of a one-mode quantum state. In particular, we introduce the notion of wave function of the state, which contains all the information encompassed in the Wigner function.

We begin by noting that the quadrature field operators $\hat{q}$ and $\hat{p}$ have continuous eigenspectra. Their respective eigenstates, which actually represent plane waves, are denoted by $|q\rangle$ and $|p\rangle$. As such, they have continuous eigenvalues, which verify

$$\hat{q}\,|q\rangle = q\,|q\rangle\,, \quad \hat{p}\,|p\rangle = p\,|p\rangle\,, \tag{B.1}$$

where $q \in \mathbb{R}$ and $p \in \mathbb{R}$. The two eigenbases $\{|q\rangle\}_{q\in\mathbb{R}}$ and $\{|p\rangle\}_{p\in\mathbb{R}}$ can be related via a Fourier transform, *i.e.*,

$$|q\rangle = \frac{1}{2\sqrt{\pi}} \int \mathrm{d}p\, \mathrm{e}^{-iqp/2}\,|p\rangle\,, \tag{B.2}$$

and

$$|p\rangle = \frac{1}{2\sqrt{\pi}} \int \mathrm{d}q\, \mathrm{e}^{iqp/2}\,|q\rangle\,. \tag{B.3}$$

For a pure quantum state $|\psi\rangle$, the wave function in the quadrature q is given by

$$\psi(q) = \langle q|\,\psi\rangle\,, \tag{B.4}$$

whereas the wave function in the quadrature p is written as

$$\tilde{\psi}(p) = \langle p|\,\psi\rangle\,. \tag{B.5}$$

They are related through a Fourier transform,

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2}} \int \mathrm{d}q\, \mathrm{e}^{-iqp/2}\psi(q). \tag{B.6}$$

The Wigner function of $|\psi\rangle$ can be related to its wave function $\tilde{\psi}$ through

$$W(q,p) = \frac{1}{2}\int \mathrm{d}u\, \mathrm{e}^{iqu/2}\tilde{\psi}\left(p+\frac{u}{2}\right)\tilde{\psi}^*\left(p-\frac{u}{2}\right), \tag{B.7}$$

and to its wavefunction ψ through

$$W(q,p) = \frac{1}{2}\int \mathrm{d}y\, \mathrm{e}^{ipy/2}\psi\left(q+\frac{y}{2}\right)\psi^*\left(q-\frac{y}{2}\right). \tag{B.8}$$

In order to recover either of the wave functions starting from the Wigner function, one can use

the following relation:

$$\begin{aligned}\int \mathrm{d}p W(q,p)\mathrm{e}^{ipq'/2} &= \psi\left(q+\frac{q'}{2}\right)\psi^*\left(q-\frac{q'}{2}\right)\\ &= \left\langle q+\frac{q'}{2}\middle|\,\psi\right\rangle\left\langle\psi\middle|\,q-\frac{q'}{2}\right\rangle,\end{aligned}$$

which leads to

$$\psi(q) = \frac{1}{\psi^*(0)}\int \mathrm{d}p\, W\left(\frac{q}{2},p\right)\mathrm{e}^{ipq/2}, \tag{B.9}$$

where the constant $\psi^*(0)$ can be obtained through normalisation of $\psi(q)$. In order to reobtain the pure state $|\psi\rangle$ itself, we do the following:

$$|\psi\rangle = \int \mathrm{d}q\,\langle q|\,\psi\rangle\,|q\rangle = \frac{1}{\psi^*(0)}\int \mathrm{d}q\mathrm{d}p\,\mathrm{e}^{ipq/2}W\left(\frac{q}{2},p\right)|q\rangle \tag{B.10}$$

For the density matrix $|\psi\rangle\langle\psi|$, it becomes

$$\begin{aligned}|\psi\rangle\langle\psi| &= \int \mathrm{d}q\mathrm{d}q'\,\langle q|\,\psi\rangle\langle\psi|\,q'\rangle\,|q\rangle\langle q'|\\ &= \int \mathrm{d}q\mathrm{d}q'\mathrm{d}p\,W\left(\frac{q+q'}{2},p\right)\mathrm{e}^{ip(q-q')/2}\,|q\rangle\langle q'|\,.\end{aligned}$$

Since the Wigner function of a mixture of pure states is given by the same mixture of the Wigner functions of the individual pure states, this relation stays true for any mixed state ρ whose Wigner function is W, i.e,

$$\rho = \int \mathrm{d}q\mathrm{d}q'\left(\int \mathrm{d}p\,W\left(\frac{q+q'}{2},p\right)\mathrm{e}^{ip(q-q')/2}\right)|q\rangle\langle q'|\,. \tag{B.11}$$

If one needs the representation of ρ in the Fock basis, one has to compute the matrix elements of ρ in the Fock basis given by

$$\langle i|\,\rho\,|j\rangle = \int \mathrm{d}q\mathrm{d}q'\int \mathrm{d}p\,W\left(\frac{q+q'}{2},p\right)\mathrm{e}^{ip(q-q')/2}\,\langle i|\,q\rangle\langle q'|\,j\rangle\,. \tag{B.12}$$

C Quantum-limited decomposition of one-mode Gaussian channels

In Section 3.4.4, we explained that any channel belonging to one of the canonical classes C(Loss), C(Amp) and D could be decomposed in terms of quantum-limited channels. In this section, we give some calculation details in phase space, for completeness.

C.1 Phase-insensitive channels

A phase-insensitive channel $\mathcal{G}_\kappa^{(\varepsilon)}$, which involves either a lossy channel $\mathcal{B}_\kappa^{(\varepsilon)}$ or an amplifier channel $\mathcal{A}_\kappa^{(\varepsilon)}$, can always be decomposed using

$$\mathcal{G}_\kappa^{(\varepsilon)} = \mathcal{A}_{G_0} \circ \mathcal{B}_{\eta_0}, \tag{C.1}$$

where ε is related to the thermal number $\bar{n}$ of $\mathcal{G}_\kappa^{(\varepsilon)}$ as

$$2\bar{n} + 1 = \frac{1+\varepsilon}{1-\varepsilon}, \tag{C.2}$$

with $\eta_0 \in (0,1)$ and $G_0 > 1$. Consider a state with covariance matrix $\mathbf{V}_0$. The pure-loss channel acts on it, outputting a state of covariance matrix $\mathbf{V}_1$ given by

$$\mathbf{V}_1 = \eta_0 \mathbf{V}_0 + (1-\eta_0)\mathbb{1}_2, \tag{C.3}$$

and is followed by the quantum-limited amplifier, which outputs

$$\mathbf{V}_2 = G_0 \mathbf{V}_1 + (G_0 - 1)\mathbb{1}_2. \tag{C.4}$$

This means that we have

$$\mathbf{V}_2 = G_0\eta_0 \mathbf{V}_0 + \left[G_0(1-\eta_0) + (G_0-1)\right]\mathbb{1}_2. \tag{C.5}$$

For the channel to be physical, one should have

$$G_0(1-\eta_0) + (G_0 - 1) \overset{?}{\geq} |G_0\eta_0 - 1|. \tag{C.6}$$

Suppose first that $G_0\eta_0 \geq 1$. In this case, one should check if

$$G_0(1-\eta_0) + (G_0 - 1) \overset{?}{\geq} G_0\eta_0 - 1. \tag{C.7}$$

This is true if $\eta_0 \leq 1$, which is satisfied by definition. In this case, the channel $\mathcal{G}_\kappa^{(\varepsilon)}$ is an amplifier channel, *i.e.*,

$$\mathcal{G}_\kappa^{(\varepsilon)} = \mathcal{A}_\kappa^{(\varepsilon)}, \tag{C.8}$$

which satisfies

$$\begin{cases} \kappa = G_0\eta_0 \geq 1, \\ \bar{n} = \dfrac{G_0(1-\eta_0)}{\eta_0 G_0 - 1} \geq 0. \end{cases} \tag{C.9}$$

Now, if $G_0\eta_0 \leq 1$, one should check if

$$G_0(1-\eta_0) + (G_0 - 1) \overset{?}{\geq} 1 - G_0\eta_0. \tag{C.10}$$

This is true if $G_0 \geq 1$, which is also verified by definition. In this case, the channel $\mathcal{G}_\kappa^{(\varepsilon)}$ is a lossy channel, *i.e.*,

$$\mathcal{G}_\kappa^{(\varepsilon)} = \mathcal{B}_\kappa^{(\varepsilon)}, \tag{C.11}$$

which satisfies

$$\begin{cases} \kappa = G_0\eta_0 \leq 1, \\ \bar{n} = \dfrac{G_0 - 1}{1 - \eta_0 G_0} \geq 0. \end{cases} \tag{C.12}$$

C.2 Phase-conjugate channels

A phase-conjugate channel $\tilde{\mathcal{G}}_\kappa^{(\varepsilon)}$ can always be decomposed as

$$\tilde{\mathcal{G}}_\kappa^{(\varepsilon)} = \tilde{\mathcal{A}}_{G_0} \circ \mathcal{B}_{\eta_0}, \tag{C.13}$$

where ε is related to the thermal number $\bar{n}$ of $\tilde{\mathcal{G}}_\kappa^{(\varepsilon)}$ as

$$2\bar{n} + 1 = \frac{1+\varepsilon}{1-\varepsilon}, \tag{C.14}$$

with $\eta_0 \in (0, 1)$ and $G_0 > 1$. Consider a state with covariance matrix $\mathbf{V}_0$. The pure-loss channel acts on it, outputting a state of covariance matrix $\mathbf{V}_1$ given by

$$\mathbf{V}_1 = \eta_0\mathbf{V}_0 + (1-\eta_0)\mathbb{1}_2, \tag{C.15}$$

and is followed by the quantum-limited phase-conjugate channel, which outputs

$$\mathbf{V}_2 = (G_0 - 1)\mathbf{V}_1 + G_0\mathbb{1}_2. \tag{C.16}$$

This means that we have

$$\mathbf{V}_2 = (G_0 - 1)\eta_0\mathbf{V}_0 + \left[(G_0 - 1)(1-\eta_0) + G_0\right]\mathbb{1}_2. \tag{C.17}$$

For the channel to be physical, one should have

$$(G_0 - 1)(1 - \eta_0) + G_0 \overset{?}{\geq} 1 + (G_0 - 1)\eta_0. \tag{C.18}$$

This is true if $G_0 \geq 1$, which is always satisfied. In this case, the channel $\mathcal{G}_\kappa^{(\varepsilon)}$ is a phase-conjugate channel, *i.e.*,

$$\tilde{\mathcal{G}}_\kappa^{(\varepsilon)} = \tilde{\mathcal{A}}_G^{(\varepsilon)}, \tag{C.19}$$

which satisfies

$$\begin{cases} G = \eta_0(G_0 - 1) + 1 \geq 1, \\ \bar{n} = \dfrac{(G_0 - 1)(1 - \eta_0)}{1 + (G_0 - 1)\eta_0} \geq 0. \end{cases} \tag{C.20}$$

D Description of a beam splitter in Fock space

D.1 Transition amplitudes

In Chapter 5, we computed the transition amplitudes in a beam splitter using the generating function. The purpose of this section is to calculate them without the help of the generating function. This will allow us to understand the usefulness of the latter by comparing the two methods. The action of the beam-splitter unitary on the product of two arbitrary Fock states is

$$|\psi^{\mathrm{BS}}_{i,k}\rangle = U^{\mathrm{BS}}_{\eta}\,|i,k\rangle = \frac{1}{\sqrt{k!}}U^{\mathrm{BS}}_{\eta}(\hat{b}^{\dagger})^{k}\,|i,0\rangle = \frac{1}{\sqrt{k!}}U^{\mathrm{BS}}_{\eta}(\hat{b}^{\dagger})^{k}U^{\mathrm{BS}\dagger}_{\eta}U^{\mathrm{BS}}_{\eta}\,|i,0\rangle\,, \tag{D.1}$$

$$\Rightarrow \quad |\psi^{\mathrm{BS}}_{i,k}\rangle = \frac{1}{\sqrt{k!}}\left(U^{\mathrm{BS}}_{\eta}\hat{b}^{\dagger}U^{\mathrm{BS}\dagger}_{\eta}\right)^{k}|\psi^{\mathrm{BS}}_{i,0}\rangle\,. \tag{D.2}$$

Exploiting the effect of the beam splitter in the Heisenberg picture, we get

$$\begin{aligned}|\psi^{\mathrm{BS}}_{i,k}\rangle &= \frac{1}{\sqrt{k!}}\left(\sqrt{1-\eta}\hat{a}^{\dagger}+\sqrt{\eta}\hat{b}^{\dagger}\right)^{k}|\psi^{\mathrm{BS}}_{i,0}\rangle \\ &= \frac{1}{\sqrt{k!}}\sum_{m=0}^{k}\binom{k}{m}\left(\sqrt{1-\eta}\hat{a}^{\dagger}\right)^{m}\left(\sqrt{\eta}\hat{b}^{\dagger}\right)^{k-m}|\psi^{\mathrm{BS}}_{i,0}\rangle\,.\end{aligned} \tag{D.3}$$

Now, in the simpler case in which one of the Fock state corresponds to the vacuum state, and using the same techniques as above, we end up with

$$\begin{aligned}|\psi^{\mathrm{BS}}_{i,0}\rangle &= \frac{1}{\sqrt{i!}}\left(\sqrt{\eta}\hat{a}^{\dagger}-\sqrt{1-\eta}\hat{b}^{\dagger}\right)^{i}|0,0\rangle \\ &= \frac{1}{\sqrt{i!}}\sum_{n=0}^{i}\binom{i}{n}\left(\sqrt{\eta}\hat{a}^{\dagger}\right)^{n}\left(-\sqrt{1-\eta}\hat{b}^{\dagger}\right)^{i-n}|0,0\rangle \\ &= \frac{1}{\sqrt{i!}}\sum_{n=0}^{i}\binom{i}{n}\sqrt{n!(i-n)!}\left(\sqrt{\eta}\right)^{n}\left(-\sqrt{1-\eta}\right)^{i-n}|n,i-n\rangle\,,\end{aligned}$$

$$\Rightarrow \quad |\psi^{\mathrm{BS}}_{i,0}\rangle = \sum_{n=0}^{i}(-1)^{i-n}\sqrt{\binom{i}{n}\eta^{n}(1-\eta)^{i-n}}\,|n,i-n\rangle\,. \tag{D.4}$$

The effect of creation operators on individual Fock states is such that

$$\begin{cases}(\hat{a}^{\dagger})^{m}\,|n\rangle = \sqrt{(n+1)\dots(n+m)}\,|n+m\rangle\,, \\ (\hat{b}^{\dagger})^{k-m}\,|i-n\rangle = \sqrt{(i-n+1)\dots(i-n+k-m)}\,|i-n+k-m\rangle\,.\end{cases} \tag{D.5}$$

Combining this with Equations (D.3) and (D.4), we obtain

$$|\psi_{i,k}^{\mathrm{BS}}\rangle = \sum_{n,m=0}^{i,k} (-1)^{i-n} \sqrt{\Gamma_{n,m}^{(i,k)} \eta^{n+k-m} (1-\eta)^{i-n+m}} \, |n+m, i-n+k-m\rangle \,, \tag{D.6}$$

where we defined

$$\Gamma_{n,m}^{(i,k)} = \binom{i}{n}\binom{k}{m}\binom{n+m}{n}\binom{i-n+k-m}{i-n}. \tag{D.7}$$

If one defines the amplitude $b_n^{(i,k)} = \langle n, m| \, U_\eta^{\mathrm{BS}} \, |i,k\rangle$ (noting that the index $m = i+k-n$ is redundant), Equation (D.6) can be rewritten as

$$|\psi_{i,k}^{\mathrm{BS}}\rangle = \sum_{n=0}^{i+k} b_n^{(i,k)} \, |n, i+k-n\rangle \,, \tag{D.8}$$

In this case, the amplitude is found to be equal to

$$b_n^{(i,k)} = \sum_{m=\max(0,n-k)}^{\min(i,n)} (-1)^{i-m} \sqrt{\Gamma_{m,n-m}^{(i,k)} \eta^{2m+k-n} (1-\eta)^{i-2m+n}}. \tag{D.9}$$

The method we used here to obtain an expression for $b_n^{(i,k)}$ is to be compared with the approach developed in Chapter 5 in terms of generating functions.

D.2 Transition probabilities

We now compute the probability

$$B_n^{(i,k)} = |\, \langle n, m| \, U_\eta^{\mathrm{BS}} \, |i,k\rangle \,|^2 = \left(b_n^{(i,k)}\right)^2, \tag{D.10}$$

since $b_n^{(i,k)}$ is real, meaning that

$$B_n^{(i,k)} = \sum_{m,j=\max(0,n-k)}^{\min(i,n)} (-1)^{m+j} \sqrt{\Gamma_{m,n-m}^{(i,k)} \Gamma_{j,n-j}^{(i,k)}} \eta^{k-n+m+j} (1-\eta)^{i+n-m-j}. \tag{D.11}$$

Now, notice that

$$\Gamma_{m,n-m}^{(i,k)} \Gamma_{j,n-j}^{(i,k)} = \binom{i}{m}\binom{k}{n-m}\binom{n}{m}\binom{i+k-n}{i-m} \times \binom{i}{j}\binom{k}{n-j}\binom{n}{j}\binom{i+k-n}{i-j},$$

$$\Gamma^{(i,k)}_{m,n-m}\Gamma^{(i,k)}_{j,n-j} == \underbrace{\binom{i}{m}\binom{k}{n-m}\binom{n}{j}\binom{i+k-n}{i-j}}_{\gamma^{(i,k)}_{n,m,j}} \times \underbrace{\binom{i}{j}\binom{k}{n-j}\binom{n}{m}\binom{i+k-n}{i-m}}_{\gamma^{(i,k)}_{n,j,m}}, \tag{D.12}$$

where

$$\begin{aligned}\gamma^{(i,k)}_{n,m,j} &= \binom{i}{m}\binom{k}{n-m}\binom{n}{j}\binom{i+k-n}{i-j} \\ &= \frac{i!}{m!(i-m)!}\frac{k!}{(n-m)!(k-n+m)!}\frac{n!}{j!(n-j)!}\frac{(i+k-n)!}{(i-j)!(k-n+j)!} \\ &= \frac{i!}{j!(i-j)!}\frac{k!}{(n-j)!(k-n+j)!}\frac{n!}{m!(n-m)!}\frac{(i+k-n)!}{(i-m)!(k-n+m)!} \\ &= \gamma^{(i,k)}_{n,j,m}.\end{aligned}$$

Using this, we get

$$\Gamma^{(i,k)}_{m,n-m}\Gamma^{(i,k)}_{j,n-j} = (\gamma^{(i,k)}_{n,m,j})^2, \tag{D.13}$$

allowing us to simplify Equation (D.11) into

$$B^{(i,k)}_n = \sum_{m,j=\max(0,n-k)}^{\min(i,n)} (-1)^{m+j}\gamma^{(i,k)}_{n,m,j}\eta^{k-n+m+j}(1-\eta)^{i+n-m-j}, \tag{D.14}$$

or,

$$B^{(i,k)}_n = \sum_{m,j=\max(0,n-k)}^{\min(i,n)} (-1)^{m+j}\binom{i}{m}\binom{k}{n-m}\binom{n}{j}\binom{i+k-n}{i-j}\eta^{k-n+m+j}(1-\eta)^{i+n-m-j}. \tag{D.15}$$

This last expression is quite cumbersome. Therefore, we are tempted to find a relation connecting these probabilities that would be easier to handle. The generating functions allow to prove such a relation.

E Transition probabilities of N-mode passive Gaussian unitaries

E.1 Derivation of the generating function of the transition probabilities

In this section, we prove Lemma 7, by computing the $2N$-variate generating function of the transition probabilities $B_{\mathbf{n}}^{(\mathbf{i})}$ defined in Section 5.3. We explained that the generating function could be written as

$$f^{\mathrm{PI}}(\mathbf{x},\mathbf{z}) = \left(\prod_{s=1}^{N}(1-x_s)^{-1}\right)\left(\prod_{r=1}^{N}(1-z_r)^{-1}\right)\mathrm{Tr}\left[U^{\mathrm{PI}}\Gamma_{\mathbf{x}}U^{\mathrm{PI}\dagger}\Gamma_{\mathbf{z}}\right], \tag{E.1}$$

The overlap between the two Gaussian states $U^{\mathrm{PI}}\Gamma_{\mathbf{x}}U^{\mathrm{PI}\dagger}$ and $\Gamma_{\mathbf{z}}$ of respective covariance matrices $\mathbf{V}'_{\mathbf{x}}$ and $\mathbf{V}_{\mathbf{z}}$ is given by

$$\mathrm{Tr}\left[U^{\mathrm{PI}}\Gamma_{\mathbf{x}}U^{\mathrm{PI}\dagger}\Gamma_{\mathbf{z}}\right] = \left(\det\left[\frac{\mathbf{V}'_{\mathbf{x}}+\mathbf{V}_{\mathbf{z}}}{2}\right]\right)^{-1/2} = 2^N\left(\det\left[\mathbf{V}'_{\mathbf{x}}+\mathbf{V}_{\mathbf{z}}\right]\right)^{-1/2}. \tag{E.2}$$

If we define

$$\tilde{\mathbf{X}} = \bigoplus_{s=1}^{N}\left(\frac{1+x_s}{1-x_s}\right), \quad \tilde{\mathbf{Z}} = \bigoplus_{r=1}^{N}\left(\frac{1+z_r}{1-z_r}\right), \tag{E.3}$$

$$\begin{aligned}
\det\left[\mathbf{V}'_{\mathbf{x}}+\mathbf{V}_{\mathbf{z}}\right] &= \det\left[\left(\mathbf{U}\tilde{\mathbf{X}}\mathbf{U}^{\mathrm{T}}\right)\oplus\left(\mathbf{U}\tilde{\mathbf{X}}\mathbf{U}^{\mathrm{T}}\right)+\tilde{\mathbf{Z}}\oplus\tilde{\mathbf{Z}}\right]\\
&= \det\left[\left(\mathbf{U}\tilde{\mathbf{X}}\mathbf{U}^{\mathrm{T}}+\tilde{\mathbf{Z}}\right)\oplus\left(\mathbf{U}\tilde{\mathbf{X}}\mathbf{U}^{\mathrm{T}}+\tilde{\mathbf{Z}}\right)\right]\\
&= \det\left[\left(\mathbf{U}\tilde{\mathbf{X}}\mathbf{U}^{\mathrm{T}}+\tilde{\mathbf{Z}}\right)\right]\det\left[\left(\mathbf{U}\tilde{\mathbf{X}}\mathbf{U}^{\mathrm{T}}+\tilde{\mathbf{Z}}\right)\right]\\
&= \left(\det\left[\left(\mathbf{U}\tilde{\mathbf{X}}\mathbf{U}^{\mathrm{T}}+\tilde{\mathbf{Z}}\right)\right]\right)^2.
\end{aligned}$$

Consequently,

$$\mathrm{Tr}\left[U^{\mathrm{PI}}\Gamma_{\mathbf{x}}U^{\mathrm{PI}\dagger}\Gamma_{\mathbf{z}}\right] = 2^N\left(\det\left[\left(\mathbf{U}\tilde{\mathbf{X}}\mathbf{U}^{\mathrm{T}}+\tilde{\mathbf{Z}}\right)\right]\right)^{-1}. \tag{E.4}$$

Using this, we end up with

$$f^{\mathrm{PI}}(\mathbf{x},\mathbf{z}) = 2^N\left(\det\left[\mathbb{1}_N-\mathbf{X}\right]\det\left[\mathbb{1}_N-\mathbf{Z}\right]\det\left[\mathbf{U}\tilde{\mathbf{X}}\mathbf{U}^{\mathrm{T}}+\tilde{\mathbf{Z}}\right]\right)^{-1}, \tag{E.5}$$

where $\mathbb{1}_N$ is the identity matrix of dimension N, $\mathbf{X} = \bigoplus_{s=1}^{N} x_s$ and $\mathbf{Z} = \bigoplus_{r=1}^{N} z_r$. The generating function can be further simplified as follows,

$$\begin{aligned}
f^{\mathrm{PI}}(\mathbf{x},\mathbf{z}) &= 2^N\left(\det\left[\mathbb{1}_N-\mathbf{X}\right]\det\left[\left(\mathbb{1}_N-\mathbf{Z}\right)\mathbf{U}\tilde{\mathbf{X}}\mathbf{U}^{\mathrm{T}}+\left(\mathbb{1}_N-\mathbf{Z}\right)\tilde{\mathbf{Z}}\right]\right)^{-1}\\
&= 2^N\left(\det\left[\mathbb{1}_N-\mathbf{X}\right]\det\left[\left(\mathbb{1}_N-\mathbf{Z}\right)\mathbf{U}\tilde{\mathbf{X}}\mathbf{U}^{\mathrm{T}}+\left(\mathbb{1}_N+\mathbf{Z}\right)\right]\right)^{-1}\\
&= 2^N\left(\det\left[\mathbb{1}_N-\mathbf{X}\right]\det\left[\mathbf{U}^{\mathrm{T}}\left(\mathbb{1}_N-\mathbf{Z}\right)\mathbf{U}\tilde{\mathbf{X}}+\mathbf{U}^{\mathrm{T}}\left(\mathbb{1}_N+\mathbf{Z}\right)\mathbf{U}\right]\right)^{-1}\\
&= 2^N\left(\det\left[\mathbf{U}^{\mathrm{T}}\left(\mathbb{1}_N-\mathbf{Z}\right)\mathbf{U}\tilde{\mathbf{X}}\left(\mathbb{1}_N-\mathbf{X}\right)+\mathbf{U}^{\mathrm{T}}\left(\mathbb{1}_N+\mathbf{Z}\right)\mathbf{U}\left(\mathbb{1}_N-\mathbf{X}\right)\right]\right)^{-1},
\end{aligned}$$

$$f^{\mathrm{PI}}(\mathbf{x},\mathbf{z}) = 2^N\left(\det\left[\mathbf{U}^{\mathrm{T}}\left(\mathbb{1}_N-\mathbf{Z}\right)\mathbf{U}\left(\mathbb{1}_N+\mathbf{X}\right)+\mathbf{U}^{\mathrm{T}}\left(\mathbb{1}_N+\mathbf{Z}\right)\mathbf{U}\left(\mathbb{1}_N-\mathbf{X}\right)\right]\right)^{-1}. \tag{E.6}$$

The determinant of a matrix $\mathbf{M}$ of dimension N can be defined using the Leibniz formula

$$\det[\mathbf{M}] = \sum_{\sigma \in S_N} \left(\operatorname{sgn}(\sigma) \prod_{i=1}^{N} M_{i,\sigma_i} \right), \tag{E.7}$$

where the sum is computed over all permutations σ of the set $\sigma^{(0)} = \{1, 2, \ldots, N\}$, S_N denoting the symmetric group on N elements. $\operatorname{sgn}(\sigma)$ denotes the signature of σ, which is equal to 1 whenever the reordering given by σ can be achieved by successively interchanging two entries an even number of times, and -1 otherwise.

Example 3. For instance, if $N = 2$, $\sigma^{(0)} = \{1, 2\}$, and

$$S_2 = \{\{1, 2\}, \{2, 1\}\}, \tag{E.8}$$

with

$$sgn(\{1, 2\}) = 1, \quad sgn(\{2, 1\}) = -1, \tag{E.9}$$

so that

$$\begin{aligned} \det[\mathbf{M}] &= \left(\prod_{i=1}^{2} M_{i,\sigma_i} \right)_{\sigma=\{1,2\}} - \left(\prod_{i=1}^{2} M_{i,\sigma_i} \right)_{\sigma=\{2,1\}} \\ &= M_{11}M_{22} - M_{12}M_{21}. \end{aligned}$$

We begin by calculating

$$\begin{aligned} \left(\mathbf{U}^{\mathrm{T}} (\mathbb{1}_N - \mathbf{Z}) \mathbf{U} (\mathbb{1}_N + \mathbf{X})\right)_{ij} &= \sum_k \left(\mathbf{U}^{\mathrm{T}} (\mathbb{1}_N - \mathbf{Z})\right)_{ik} \left(\mathbf{U} (\mathbb{1}_N + \mathbf{X})\right)_{kj} \\ &= \sum_{krs} u_{ri} (\mathbb{1}_N - \mathbf{Z})_{rk} u_{ks} (\mathbb{1}_N + \mathbf{X})_{sj} \\ &= \sum_k u_{ki} (\mathbb{1}_N - \mathbf{Z})_{kk} u_{kj} (\mathbb{1}_N + \mathbf{X})_{jj} \\ &= \sum_k u_{ki} u_{kj} (1 - z_k)(1 + x_j), \end{aligned}$$

where we chose to define the elements of the matrix $\mathbf{U}$ as u_{ij}. If

$$\mathbf{M} = \mathbf{U}^{\mathrm{T}} (\mathbb{1}_N - \mathbf{Z}) \mathbf{U} (\mathbb{1}_N + \mathbf{X}) + \mathbf{U}^{\mathrm{T}} (\mathbb{1}_N + \mathbf{Z}) \mathbf{U} (\mathbb{1}_N - \mathbf{X}), \tag{E.10}$$

then

$$M_{ij} = \sum_k u_{ki} u_{kj} \left[(1 - z_k)(1 + x_j) + (1 + z_k)(1 - x_j) \right] = 2 \sum_k u_{ki} u_{kj} (1 - x_j z_k). \tag{E.11}$$

We now have

$$f^{\mathrm{PI}}(\mathbf{x},\mathbf{z}) = 2^N \left(\sum_{\sigma\in S_N} \left(\operatorname{sgn}(\sigma) \prod_{i=1}^{N} 2 \sum_k u_{ki} u_{k\sigma_i} (1 - x_{\sigma_i} z_k) \right) \right)^{-1} \tag{E.12}$$

so that, if we define $\bar{f}(\mathbf{x},\mathbf{z}) = \left(f^{\mathrm{PI}}(\mathbf{x},\mathbf{z})\right)^{-1}$,

$$\bar{f}(\mathbf{x},\mathbf{z}) = \sum_{\sigma\in S_N} \left(\operatorname{sgn}(\sigma) \prod_{i=1}^{N} \sum_k u_{ki} u_{k\sigma_i} (1 - x_{\sigma_i} z_k) \right). \tag{E.13}$$

Since $\mathbf{U}$ is orthogonal,

$$\begin{aligned}
\bar{f}(\mathbf{x},\mathbf{z}) &= \sum_{\sigma\in S_N} \left(\operatorname{sgn}(\sigma) \prod_{i=1}^{N} \left[\sum_k u_{ki} u_{k\sigma_i} - \sum_k u_{ki} u_{k\sigma_i} x_{\sigma_i} z_k \right] \right) \\
&= \sum_{\sigma\in S_N} \left(\operatorname{sgn}(\sigma) \prod_{i=1}^{N} \left[\delta_{i,\sigma_i} - \sum_k u_{ki} u_{k\sigma_i} x_{\sigma_i} z_k \right] \right) \\
&= \sum_{\sigma\in S_N} \left(\operatorname{sgn}(\sigma) \prod_{i=1}^{N} \left[\delta_{i,\sigma_i} - c_{i,\sigma_i} \right] \right),
\end{aligned}$$

where we defined

$$c_{i,\sigma_i} = \sum_{k=1}^{N} u_{ki} u_{k\sigma_i} x_{\sigma_i} z_k. \tag{E.14}$$

Notice that this means

$$\bar{f}(\mathbf{x},\mathbf{z}) = \det\left[\mathbb{1}_N - C\right] = \det\left[\mathbb{1}_N - \mathbf{U}^{\mathrm{T}}\mathbf{Z}\mathbf{U}\mathbf{X}\right]. \tag{E.15}$$

We need to show that

$$\bar{f}(\mathbf{x},\mathbf{z}) = \sum_{m=0}^{N} (-1)^m \sum_{\alpha\in\mathcal{R}_m^{(N)}} \sum_{\beta\in\mathcal{R}_m^{(N)}} \left(\det\left[\mathbf{U}(\beta,\alpha)\right]\right)^2 \det\left[\mathbf{X}(\alpha)\right] \det\left[\mathbf{Z}(\beta)\right]. \tag{E.16}$$

We begin by expanding

$$\prod_{i=1}^{N} \left[\delta_{i,\sigma_i} - c_{i,\sigma_i}\right] = \sum_{\nu\in\mathcal{P}_n} \left(\prod_{i\in\nu} \delta_{i,\sigma_i} \right) \left(\prod_{j\in\bar{\nu}} (-c_{j,\sigma_j}) \right), \tag{E.17}$$

where $\sigma^{(0)} = \{1, 2, \ldots, N\}$ was defined earlier, $\mathcal{P}_N$ is the power set of $\sigma^{(0)}$ (the set of all subsets of $\sigma^{(0)}$, including the empty set and $\sigma^{(0)}$ itself), and $\bar{\nu}$ is the relative complement of ν in $\sigma^{(0)}$, i.e,

$\bar{\nu} = \sigma^{(o)} \setminus \nu = \left\{ i \in \sigma^{(o)} | i \notin \nu \right\}$. We also choose the convention

$$\prod_{i \in \emptyset} t_i = 1, \tag{E.18}$$

for any sequence t_i. Using the definition of the cardinality $|\nu|$ of a set ν, we get

$$\prod_{i=1}^{N} [\delta_{i,\sigma_i} - c_{i,\sigma_i}] = \sum_{\nu \in \mathcal{P}_N} (-1)^{|\bar{\nu}|} \left(\prod_{i \in \nu} \delta_{i,\sigma_i} \right) \left(\prod_{j \in \bar{\nu}} c_{j,\sigma_j} \right). \tag{E.19}$$

Example 4. For instance, if $N = 2$, $\sigma^{(o)} = \{1, 2\}$, and

$$\mathcal{P}_2 = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}, \tag{E.20}$$

so that

$$\begin{aligned}
\prod_{i=1}^{2} [\delta_{i,\sigma_i} - c_{i,\sigma_i}] = & \left(\prod_{i \in \emptyset} \delta_{i,\sigma_i} \right) \left(\prod_{j \in \{1,2\}} (-c_{j,\sigma_j}) \right) + \left(\prod_{i \in \{1\}} \delta_{i,\sigma_i} \right) \left(\prod_{j \in \{2\}} (-c_{j,\sigma_j}) \right) \\
& + \left(\prod_{i \in \{2\}} \delta_{i,\sigma_i} \right) \left(\prod_{j \in \{1\}} (-c_{j,\sigma_j}) \right) + \left(\prod_{i \in \{1,2\}} \delta_{i,\sigma_i} \right) \left(\prod_{j \in \emptyset} (-c_{j,\sigma_j}) \right) \\
= & \, c_{1,\sigma_1} c_{2,\sigma_2} - \delta_{1,\sigma_1} c_{2,\sigma_2} - \delta_{2,\sigma_2} c_{1,\sigma_1} + \delta_{1,\sigma_1} \delta_{2,\sigma_2},
\end{aligned}$$

and $\sigma \in S_2 = \{\{1, 2\}, \{2, 1\}\}$, meaning that

$$\left(\prod_{i=1}^{2} [\delta_{i,\sigma_i} - c_{i,\sigma_i}] \right) \Bigg|_{\sigma = \{1,2\}} = c_{11} c_{22} - c_{22} - c_{11} + 1, \tag{E.21}$$

and

$$\left(\prod_{i=1}^{2} [\delta_{i,\sigma_i} - c_{i,\sigma_i}] \right) \Bigg|_{\sigma = \{2,1\}} = c_{12} c_{21}. \tag{E.22}$$

In the end, since

$$c_{ij} = x_j (u_{1i} u_{1j} z_1 + u_{2i} u_{2j} z_2), \tag{E.23}$$

we end up with

$$\begin{aligned}\bar{f}(\mathbf{x},\mathbf{z}) &= c_{11}c_{22} - c_{22} - c_{11} + 1 - c_{12}c_{21}\\ &= x_1(u_{11}u_{11}z_1 + u_{21}u_{21}z_2)x_2(u_{12}u_{12}z_1 + u_{22}u_{22}z_2) - x_2(u_{12}u_{12}z_1 + u_{22}u_{22}z_2)\\ &\quad - x_1(u_{11}u_{11}z_1 + u_{21}u_{21}z_2) + 1 - x_2(u_{11}u_{12}z_1 + u_{21}u_{22}z_2)x_1(u_{12}u_{11}z_1 + u_{22}u_{21}z_2),\end{aligned}$$

$$\begin{aligned}\bar{f}(\mathbf{x},\mathbf{z}) &= 1 - u_{11}^2x_1z_1 - u_{22}^2x_2z_2 - u_{21}^2x_1z_2 - u_{12}^2x_2z_1\\ &\quad + (u_{11}^2u_{22}^2 + u_{21}^2u_{12}^2 - 2u_{11}u_{12}u_{22}u_{21})x_1x_2z_1z_2\\ &\quad + (u_{11}^2u_{12}^2 - u_{11}^2u_{12}^2)x_1x_2z_1^2 + (u_{21}^2u_{22}^2 - u_{21}^2u_{22}^2)x_1x_2z_2^2\\ &= 1 - u_{11}^2x_1z_1 - u_{22}^2x_2z_2 - u_{21}^2x_1z_2 - u_{12}^2x_2z_1 + x_1x_2z_1z_2.\end{aligned}$$

One can understand that

$$\prod_{i\in\nu}\sum_{k=1}^{N} b_{ik} = \sum_{a\in(\sigma^{(0)})^{|\nu|}}\prod_{k=1}^{|\nu|} b_{\nu_k a_k}, \tag{E.24}$$

where, in the notation we used, the set $(\sigma^{(0)})^{|\nu|}$ represents the $|\nu|$-fold Cartesian product of the set $\sigma^{(0)}$ with itself. In this case,

$$\prod_{j\in\bar{\nu}} c_{j,\sigma_j} = \prod_{j\in\bar{\nu}}\sum_{k=1}^{N} u_{kj}u_{k\sigma_j}x_{\sigma_j}z_k = \sum_{a\in(\sigma^{(0)})^{|\bar{\nu}|}}\prod_{k=1}^{|\bar{\nu}|} u_{a_k\bar{\nu}_k}u_{a_k\sigma_{\bar{\nu}_k}}x_{\sigma_{\bar{\nu}_k}}z_{a_k}. \tag{E.25}$$

This leads to

$$\prod_{i=1}^{N}[\delta_{i,\sigma_i} - c_{i,\sigma_i}] = \sum_{\nu\in\mathcal{P}_N}(-1)^{|\bar{\nu}|}\left(\prod_{i\in\nu}\delta_{i,\sigma_i}\right)\left(\sum_{a\in(\sigma^{(0)})^{|\bar{\nu}|}}\prod_{k=1}^{|\bar{\nu}|} u_{a_k\bar{\nu}_k}u_{a_k\sigma_{\bar{\nu}_k}}x_{\sigma_{\bar{\nu}_k}}z_{a_k}\right), \tag{E.26}$$

and

$$\begin{aligned}\bar{f}(\mathbf{x},\mathbf{z}) = \sum_{\sigma\in S_N}\Bigg[&\operatorname{sgn}(\sigma)\sum_{\nu\in\mathcal{P}_N}(-1)^{|\bar{\nu}|}\left(\prod_{i\in\nu}\delta_{i,\sigma_i}\right)\\ &\times\left(\sum_{a\in(\sigma^{(0)})^{|\bar{\nu}|}}\prod_{k=1}^{|\bar{\nu}|} u_{a_k\bar{\nu}_k}u_{a_k\sigma_{\bar{\nu}_k}}x_{\sigma_{\bar{\nu}_k}}z_{a_k}\right)\Bigg].\end{aligned}$$

Since the power set $\mathcal{P}_N$ does not depend on $\sigma\in S_N$, we can exchange the summations on ν and σ, so that

$$\bar{f}(\mathbf{x},\mathbf{z}) = \sum_{\nu\in\mathcal{P}_N}(-1)^{|\bar{\nu}|}\sum_{\sigma\in S_N}\operatorname{sgn}(\sigma)\left(\prod_{i\in\nu}\delta_{i,\sigma_i}\right)\sum_{a\in(\sigma^{(0)})^{|\bar{\nu}|}}\prod_{k=1}^{|\bar{\nu}|} u_{a_k\bar{\nu}_k}u_{a_k\sigma_{\bar{\nu}_k}}x_{\sigma_{\bar{\nu}_k}}z_{a_k}. \tag{E.27}$$

We divide the summation over $\nu\in\mathcal{P}_N$ in two summations, one over the cardinality $m = |\nu|$

going from 0 to N, and one over the sets ν of cardinality m, getting

$$\bar{f}(\mathbf{x},\mathbf{z}) = \sum_{m=0}^{N}(-1)^{N-m}\sum_{\substack{\nu\in\mathcal{P}_N\\|\nu|=m}}\sum_{\sigma\in S_N}\operatorname{sgn}(\sigma)\left(\prod_{i\in\nu}\delta_{i,\sigma_i}\right)\sum_{a\in(\sigma^{(0)})^{|\bar{\nu}|}}\prod_{k=1}^{|\bar{\nu}|}u_{a_k\bar{\nu}_k}u_{a_k\sigma_{\bar{\nu}_k}}x_{\sigma_{\bar{\nu}_k}}z_{a_k}. \tag{E.28}$$

Define

$$\bar{f}_m(\mathbf{x},\mathbf{z}) = \sum_{\substack{\nu\in\mathcal{P}_N\\|\nu|=m}}\sum_{\sigma\in S_N}\operatorname{sgn}(\sigma)\left(\prod_{i\in\nu}\delta_{i,\sigma_i}\right)\sum_{a\in(\sigma^{(0)})^{|\bar{\nu}|}}\prod_{k=1}^{|\bar{\nu}|}u_{a_k\bar{\nu}_k}u_{a_k\sigma_{\bar{\nu}_k}}x_{\sigma_{\bar{\nu}_k}}z_{a_k}. \tag{E.29}$$

Since the set $\sigma^{(0)}$ is unique, we can exchange the summation on σ with the one on a, so that

$$\bar{f}_m(\mathbf{x},\mathbf{z}) = \sum_{\substack{\nu\in\mathcal{P}_N\\|\nu|=m}}\sum_{a\in(\sigma^{(0)})^{N-m}}\sum_{\sigma\in S_N}\operatorname{sgn}(\sigma)\left(\prod_{i\in\nu}\delta_{i,\sigma_i}\right)\prod_{k=1}^{N-m}u_{a_k\bar{\nu}_k}u_{a_k\sigma_{\bar{\nu}_k}}x_{\sigma_{\bar{\nu}_k}}z_{a_k}. \tag{E.30}$$

The set $(\sigma^{(0)})^{|\nu|}$ represents the $|\nu|$-fold Cartesian product of the set $\sigma^{(0)}$ with itself. It can be seen as the set of all possible $|\nu|$-element sets , whose elements are between 1 and N, as well as all their possible permutations. If we denote by $Q_{|\nu|}^{(N)}$ the set of all possible $|\nu|$-element sets whose elements are between 1 and N, we have

$$\sum_{a\in(\sigma^{(0)})^{N-m}} = \sum_{\beta\in Q_{N-m}^{(N)}}\sum_{a\in\pi(\beta)}, \tag{E.31}$$

where $\pi(\beta)$ represents the set of all possible permutations of β. With this, we have

$$\bar{f}_m(\mathbf{x},\mathbf{z}) = \sum_{\substack{\nu\in\mathcal{P}_N\\|\nu|=m}}\sum_{\beta\in Q_{N-m}^{(N)}}\sum_{a\in\pi(\beta)}\sum_{\sigma\in S_N}\operatorname{sgn}(\sigma)\left(\prod_{i\in\nu}\delta_{i,\sigma_i}\right)\prod_{k=1}^{N-m}u_{a_k\bar{\nu}_k}u_{a_k\sigma_{\bar{\nu}_k}}x_{\sigma_{\bar{\nu}_k}}z_{a_k}, \tag{E.32}$$

and define

$$\bar{f}_m^{(\bar{\nu},\beta)}(\mathbf{x},\mathbf{z}) = \sum_{a\in\pi(\beta)}\sum_{\sigma\in S_N}\operatorname{sgn}(\sigma)\left(\prod_{i\in\nu}\delta_{i,\sigma_i}\right)\prod_{k=1}^{N-m}u_{a_k\bar{\nu}_k}u_{a_k\sigma_{\bar{\nu}_k}}x_{\sigma_{\bar{\nu}_k}}z_{a_k}, \tag{E.33}$$

where β is a $(N-m)$-element set whose elements have values between 1 and N and $\bar{\nu}\in\mathcal{P}_N$ such that $|\bar{\nu}| = N-m$. Remember that this means

$$\bar{f}(\mathbf{x},\mathbf{z}) = \sum_{m=0}^{N}(-1)^{N-m}\bar{f}_m(\mathbf{x},\mathbf{z}) = \sum_{m=0}^{N}(-1)^{N-m}\sum_{\substack{\nu\in\mathcal{P}_N\\|\nu|=m}}\sum_{\beta\in Q_{N-m}^{(N)}}\bar{f}_m^{(\bar{\nu},\beta)}(\mathbf{x},\mathbf{z}). \tag{E.34}$$

Rearranging the summations,

$$\bar{f}_m^{(\bar{\nu},\beta)}(\mathbf{x},\mathbf{z}) = \sum_{\sigma\in S_N} \operatorname{sgn}(\sigma) \left(\prod_{i\in\nu} \delta_{i,\sigma_i}\right) \sum_{a\in\pi(\beta)} \prod_{k=1}^{N-m} u_{a_k\bar{\nu}_k} u_{a_k\sigma_{\bar{\nu}_k}} x_{\sigma_{\bar{\nu}_k}} z_{a_k}, \tag{E.35}$$

and, if $\tilde{N} = N - m$,

$$\begin{aligned}
\sum_{a\in\pi(\beta)} \prod_{k=1}^{\tilde{N}} u_{a_k\bar{\nu}_k} u_{a_k\sigma_{\bar{\nu}_k}} x_{\sigma_{\bar{\nu}_k}} z_{a_k} &= \sum_{a\in\pi(\beta)} \left(\prod_{k=1}^{\tilde{N}} u_{a_k\bar{\nu}_k} u_{a_k\sigma_{\bar{\nu}_k}}\right) \left(\prod_{j=1}^{\tilde{N}} x_{\sigma_{\bar{\nu}_j}}\right) \left(\prod_{l=1}^{\tilde{N}} z_{a_l}\right) \\
&= \sum_{a\in\pi(\beta)} \left(\prod_{k=1}^{\tilde{N}} u_{a_k\bar{\nu}_k} u_{a_k\sigma_{\bar{\nu}_k}}\right) \left(\prod_{j=1}^{\tilde{N}} x_{\sigma_{\bar{\nu}_j}}\right) \left(\prod_{l=1}^{\tilde{N}} z_{\beta_l}\right) \\
&= \sum_{a\in\pi(\beta)} \left(\prod_{k=1}^{\tilde{N}} u_{a_k\bar{\nu}_k} u_{a_k\sigma_{\bar{\nu}_k}}\right) \left(\prod_{j=1}^{\tilde{N}} x_{\sigma_{\bar{\nu}_j}}\right) \left(\prod_{l\in\beta} z_l\right) \\
&= \left(\sum_{a\in\pi(\beta)} \prod_{k=1}^{\tilde{N}} u_{a_k\bar{\nu}_k} u_{a_k\sigma_{\bar{\nu}_k}}\right) \left(\prod_{j=1}^{\tilde{N}} x_{\sigma_{\bar{\nu}_j}}\right) \left(\prod_{l\in\beta} z_l\right),
\end{aligned}$$

since β is chosen so that $|\beta| = \tilde{N} = N - m$, so that

$$\bar{f}_m^{(\bar{\nu},\beta)}(\mathbf{x},\mathbf{z}) = \left[\sum_{\sigma\in S_N} \operatorname{sgn}(\sigma) \left(\prod_{i\in\nu} \delta_{i,\sigma_i}\right) \left(\sum_{a\in\pi(\beta)} \prod_{k=1}^{\tilde{N}} u_{a_k\bar{\nu}_k} u_{a_k\sigma_{\bar{\nu}_k}}\right) \left(\prod_{j=1}^{\tilde{N}} x_{\sigma_{\bar{\nu}_j}}\right)\right] \left(\prod_{l\in\beta} z_l\right). \tag{E.36}$$

If we define $\mathbf{Z}(\beta)$ as the sub-matrix of Z corresponding to the rows and columns whose indices belong in β, we can rewrite $f_m^{(\bar{\nu},\beta)}$ as

$$\bar{f}_m^{(\bar{\nu},\beta)}(\mathbf{x},\mathbf{z}) = \left[\sum_{\sigma\in S_N} \operatorname{sgn}(\sigma) \left(\prod_{i\in\nu} \delta_{i,\sigma_i}\right) \left(\sum_{a\in\pi(\beta)} \prod_{k=1}^{\tilde{N}} u_{a_k\bar{\nu}_k} u_{a_k\sigma_{\bar{\nu}_k}}\right) \left(\prod_{j=1}^{\tilde{N}} x_{\sigma_{\bar{\nu}_j}}\right)\right] \det\left[\mathbf{Z}(\beta)\right]. \tag{E.37}$$

Remark 7. Since β is a $(N-m)$-element set whose elements have values between 1 and N in general, β can have elements with the same values (meaning that, if $N - m = 2$ for instance, we can have products like z_1^2). According to what we conjectured, these instances of β ($\beta = \{1,1\}$ in our example) have to simplify and disappear in the end of our calculation.

The product over the x_s can be rewritten as

$$\bar{f}_m^{(\bar{\nu},\beta)}(\mathbf{x},\mathbf{z}) = \det\left[\mathbf{Z}(\beta)\right] \sum_{\sigma\in S_N} \operatorname{sgn}(\sigma) \left(\prod_{i\in\nu} \delta_{i,\sigma_i}\right) \left(\sum_{a\in\pi(\beta)} \prod_{k=1}^{\tilde{N}} u_{a_k\bar{\nu}_k} u_{a_k\sigma_{\bar{\nu}_k}}\right) \left(\prod_{j\in\bar{\nu}} x_{\sigma_j}\right), \tag{E.38}$$

and, rearranging the summations again,

$$\bar{f}_m^{(\bar{\nu},\beta)}(\mathbf{x},\mathbf{z}) = \det\left[\mathbf{Z}\left(\beta\right)\right] \sum_{a\in\pi(\beta)} \sum_{\sigma\in S_N} \operatorname{sgn}(\sigma) \left(\prod_{i\in\nu} \delta_{i,\sigma_i}\right) \left(\prod_{k=1}^{\tilde{N}} u_{a_k\bar{\nu}_k} u_{a_k\sigma_{\bar{\nu}_k}}\right) \left(\prod_{j\in\bar{\nu}} x_{\sigma_j}\right). \tag{E.39}$$

The factor $\prod_{i\in\nu}\delta_{i,\sigma_i}$ eliminates some of the σ in the summation over all the elements of S_N, depending on the ν (or $\bar{\nu}$). The goal here is to clarify the index σ_j of $u_{l\sigma_j}$ and x_{σ_j}, for $j\in\bar{\nu}$. Taking into account the factor $\prod_{i\in\nu}\delta_{i,\sigma_i}$, one can understand that the index will actually span all the permutations θ of $\bar{\nu}$. It also happens that $sgn(\sigma)$ which appears in the summation over σ will be equal to $sgn(\theta)$, taking the convention that $\bar{\nu}$ is always ordered. Having this in mind, we get

$$\bar{f}_m^{(\bar{\nu},\beta)}(\mathbf{x},\mathbf{z}) = \det\left[\mathbf{Z}\left(\beta\right)\right] \sum_{a\in\pi(\beta)} \sum_{\theta\in\pi(\bar{\nu})} \operatorname{sgn}(\theta) \left(\prod_{k=1}^{\tilde{N}} u_{a_k\bar{\nu}_k} u_{a_k\theta_k}\right) \left(\prod_{j=1}^{|\bar{\nu}|} x_{\theta_j}\right), \tag{E.40}$$

where $\pi\left(\bar{\nu}\right)$ represents the set of all possible permutations of $\bar{\nu}$.

Remark 8. In Remark (7) we explained that we could have products z_iz_j with $i=j$, but that they would eventually disappear. Notice here that the same cannot be said about products of the form x_ix_j, for which we always have $i\neq j$, which is in accordance with what we conjectured.

The product $\left(\prod_{j=1}^{|\bar{\nu}|} x_{\theta_j}\right)$ is the same for all permutations θ of the same $\bar{\nu}$, so that

$$\bar{f}_m^{(\bar{\nu},\beta)}(\mathbf{x},\mathbf{z}) = \det\left[\mathbf{Z}\left(\beta\right)\right] \sum_{a\in\pi(\beta)} \sum_{\theta\in\pi(\bar{\nu})} \operatorname{sgn}(\theta) \left(\prod_{k=1}^{\tilde{N}} u_{a_k\bar{\nu}_k} u_{a_k\theta_k}\right) \left(\prod_{j=1}^{|\bar{\nu}|} x_{\bar{\nu}_j}\right), \tag{E.41}$$

$$\bar{f}_m^{(\bar{\nu},\beta)}(\mathbf{x},\mathbf{z}) = \det\left[\mathbf{Z}\left(\beta\right)\right] \left(\prod_{j\in\bar{\nu}} x_j\right) \sum_{a\in\pi(\beta)} \sum_{\theta\in\pi(\bar{\nu})} \operatorname{sgn}(\theta) \left(\prod_{k=1}^{\tilde{N}} u_{a_k\bar{\nu}_k} u_{a_k\theta_k}\right), \tag{E.42}$$

$$\bar{f}_m^{(\bar{\nu},\beta)}(\mathbf{x},\mathbf{z}) = \det\left[\mathbf{Z}\left(\beta\right)\right] \det\left[\mathbf{X}\left(\bar{\nu}\right)\right] \sum_{a\in\pi(\beta)} \sum_{\theta\in\pi(\bar{\nu})} \operatorname{sgn}(\theta) \left(\prod_{k=1}^{\tilde{N}} u_{a_k\bar{\nu}_k} u_{a_k\theta_k}\right). \tag{E.43}$$

This means that

$$\bar{f}_m(\mathbf{x},\mathbf{z}) = \sum_{\substack{\nu\in\mathcal{P}_N\\|\nu|=m}} \sum_{\beta\in Q_{N-m}^{(N)}} \det\left[\mathbf{Z}\left(\beta\right)\right] \det\left[\mathbf{X}\left(\bar{\nu}\right)\right] \sum_{a\in\pi(\beta)} \sum_{\theta\in\pi(\bar{\nu})} \operatorname{sgn}(\theta) \left(\prod_{k=1}^{\tilde{N}} u_{a_k\bar{\nu}_k} u_{a_k\theta_k}\right), \tag{E.44}$$

$$\bar{f}(\mathbf{x},\mathbf{z}) = \sum_{m=0}^{N} (-1)^{N-m} \sum_{\substack{\nu\in\mathcal{P}_N\\|\nu|=m}} \sum_{\beta\in Q_{N-m}^{(N)}} \det\left[\mathbf{Z}\left(\beta\right)\right] \det\left[\mathbf{X}\left(\bar{\nu}\right)\right] \sum_{a\in\pi(\beta)} \sum_{\theta\in\pi(\bar{\nu})} \operatorname{sgn}(\theta) \left(\prod_{k=1}^{\tilde{N}} u_{a_k\bar{\nu}_k} u_{a_k\theta_k}\right), \tag{E.45}$$

and, changing the summation over the $\nu \in \mathcal{P}_N$ into an equivalent summation over the $\bar{\nu} \in \mathcal{P}_N$,

$$\bar{f}(\mathbf{x}, \mathbf{z}) = \sum_{m=0}^{N} (-1)^{N-m} \sum_{\beta \in Q_{N-m}^{(N)}} \sum_{\substack{\bar{\nu} \in \mathcal{P}_N \\ |\bar{\nu}|=N-m}} \det\left[\mathbf{Z}\left(\beta\right)\right] \det\left[\mathbf{X}\left(\bar{\nu}\right)\right] \sum_{\theta \in \pi(\bar{\nu})} \operatorname{sgn}(\theta) \sum_{\alpha \in \pi(\beta)} \left(\prod_{k=1}^{\tilde{N}} u_{\alpha_k \bar{\nu}_k} u_{\alpha_k \theta_k} \right) . \tag{E.46}$$

The last product in $\bar{f}(\mathbf{x}, \mathbf{z})$ can be written as

$$\begin{aligned} \prod_{k=1}^{\tilde{N}} u_{\alpha_k \bar{\nu}_k} u_{\alpha_k \theta_k} &= \prod_{k=1}^{\tilde{N}} \mathbf{U}\left(\alpha, \bar{\nu}\right)_{kk} \mathbf{U}\left(\alpha, \theta\right)_{kk} \\ &= \left(\prod_{k=1}^{\tilde{N}} \mathbf{U}\left(\alpha, \bar{\nu}\right)_{kk} \right) \left(\prod_{j=1}^{\tilde{N}} \mathbf{U}\left(\alpha, \theta\right)_{jj} \right) , \end{aligned}$$

so that

$$\begin{aligned} \sum_{\theta \in \pi(\bar{\nu})} \operatorname{sgn}(\theta) \prod_{k=1}^{\tilde{N}} u_{\alpha_k \bar{\nu}_k} u_{\alpha_k \theta_k} &= \left(\prod_{k=1}^{\tilde{N}} \mathbf{U}\left(\alpha, \bar{\nu}\right)_{kk} \right) \sum_{\theta \in \pi(\bar{\nu})} \operatorname{sgn}(\theta) \left(\prod_{j=1}^{\tilde{N}} \mathbf{U}\left(\alpha, \theta\right)_{jj} \right) \\ &= \left(\prod_{k=1}^{\tilde{N}} \mathbf{U}\left(\alpha, \bar{\nu}\right)_{kk} \right) \det\left[\mathbf{U}(\alpha, \bar{\nu})\right] . \end{aligned}$$

Since exchanging two rows of a matrix changes the sign of its determinant, we have

$$\det\left[\mathbf{U}(\beta, \bar{\nu})\right] = \operatorname{sgn}(\alpha) \det\left[\mathbf{U}(\alpha, \bar{\nu})\right] , \quad \alpha \in \pi\left(\beta\right) , \tag{E.47}$$

so that

$$\sum_{\theta \in \pi(\bar{\nu})} \operatorname{sgn}(\theta) \prod_{k=1}^{\tilde{N}} u_{\alpha_k \bar{\nu}_k} u_{\alpha_k \theta_k} = \left(\prod_{k=1}^{\tilde{N}} \mathbf{U}\left(\alpha, \bar{\nu}\right)_{kk} \right) \operatorname{sgn}(\alpha) \det\left[\mathbf{U}(\beta, \bar{\nu})\right] , \quad \alpha \in \pi\left(\beta\right) , \tag{E.48}$$

$$\begin{aligned} \Rightarrow \quad \sum_{\alpha \in \pi(\beta)} \sum_{\theta \in \pi(\bar{\nu})} \operatorname{sgn}(\theta) \prod_{k=1}^{\tilde{N}} u_{\alpha_k \bar{\nu}_k} u_{\alpha_k \theta_k} &= \sum_{\alpha \in \pi(\beta)} \left(\prod_{k=1}^{\tilde{N}} \mathbf{U}\left(\alpha, \bar{\nu}\right)_{kk} \right) \operatorname{sgn}(\alpha) \det\left[\mathbf{U}(\beta, \bar{\nu})\right] \\ &= \det\left[\mathbf{U}(\beta, \bar{\nu})\right] \sum_{\alpha \in \pi(\beta)} \operatorname{sgn}(\alpha) \left(\prod_{k=1}^{\tilde{N}} \mathbf{U}\left(\alpha, \bar{\nu}\right)_{kk} \right) \\ &= \det\left[\mathbf{U}(\beta, \bar{\nu})\right] \det\left[\mathbf{U}(\beta, \bar{\nu})\right] \\ &= \left(\det\left[\mathbf{U}(\beta, \bar{\nu})\right]\right)^2 . \end{aligned}$$

Using this, we end up with

$$\bar{f}(\mathbf{x},\mathbf{z}) = \sum_{m=0}^{N}(-1)^{N-m}\sum_{\beta\in Q_{N-m}^{(N)}}\sum_{\substack{\bar{\nu}\in\mathcal{P}_N\\|\bar{\nu}|=N-m}}\det\left[\mathbf{Z}\left(\beta\right)\right]\det\left[\mathbf{X}\left(\bar{\nu}\right)\right]\left(\det\left[\mathbf{U}(\beta,\bar{\nu})\right]\right)^2, \tag{E.49}$$

$$\Leftrightarrow\quad \bar{f}(\mathbf{x},\mathbf{z}) = \sum_{m=0}^{N}(-1)^{m}\sum_{\alpha\in\mathcal{R}_m^{(N)}}\sum_{\beta\in\mathcal{R}_m^{(N)}}\left(\det\left[\mathbf{U}\left(\beta,\alpha\right)\right]\right)^2\det\left[\mathbf{Z}\left(\beta\right)\right]\det\left[\mathbf{X}\left(\alpha\right)\right], \tag{E.50}$$

where $\mathcal{R}_m^{(N)}$ is the set of all subsets of $\{1,2,\ldots,N\}$ of cardinality m. Finally, the generating function of the probabilities $B_{\mathbf{n}}^{(\mathbf{i})}, \mathbf{i}\in\mathbb{N}_+^N, \mathbf{n}\in\mathbb{N}_+^N$ is of the form

$$f^{\mathrm{PI}}(\mathbf{x},\mathbf{z}) = \left(\sum_{m=0}^{N}(-1)^{m}\sum_{\alpha\in\mathcal{R}_m^{(N)}}\sum_{\beta\in\mathcal{R}_m^{(N)}}\left(\det\left[\mathbf{U}\left(\beta,\alpha\right)\right]\right)^2\det\left[\mathbf{Z}\left(\beta\right)\right]\det\left[\mathbf{X}\left(\alpha\right)\right]\right)^{-1}. \tag{E.51}$$

This concludes the proof.

E.2 Proof of the recurrence relation for the transition probabilities

In this section, we prove Theorem 29, by showing that the probabilities $B_{\mathbf{n}}^{(\mathbf{i})}$ verify the recurrence relation

$$B_{\mathbf{n}}^{(\mathbf{i})} = \sum_{m=1}^{N}(-1)^{m-1}\sum_{\substack{\alpha\in\mathcal{R}_m^{(N)}\\ \mathbf{i}_s\neq 0, s\in\alpha}}\sum_{\substack{\beta\in\mathcal{R}_m^{(N)}\\ \mathbf{n}_r\neq 0, r\in\beta}}\left(\det\left[\mathbf{U}\left(\beta,\alpha\right)\right]\right)^2 B_{\mathbf{n}-1_N^{(\beta)}}^{\left(\mathbf{i}-1_N^{(\alpha)}\right)}. \tag{E.52}$$

Equation (E.51) implies that

$$f^{\mathrm{PI}}(\mathbf{x},\mathbf{z})\left(\sum_{m=0}^{N}(-1)^{m}\sum_{\alpha\in\mathcal{R}_m^{(N)}}\sum_{\beta\in\mathcal{R}_m^{(N)}}\left(\det\left[\mathbf{U}\left(\beta,\alpha\right)\right]\right)^2\det\left[\mathbf{Z}\left(\beta\right)\right]\det\left[\mathbf{X}\left(\alpha\right)\right]\right) = 1, \tag{E.53}$$

we explicitly write the determinants of $\mathbf{Z}\left(\beta\right)$ and $\mathbf{X}\left(\alpha\right)$,

$$\sum_{m=0}^{N}(-1)^{m}\sum_{\alpha\in\mathcal{R}_m^{(N)}}\sum_{\beta\in\mathcal{R}_m^{(N)}}\left(\det\left[\mathbf{U}\left(\beta,\alpha\right)\right]\right)^2\left(\prod_{j\in\alpha}x_j\right)\left(\prod_{m\in\beta}z_m\right)f^{\mathrm{PI}}(\mathbf{x},\mathbf{z}) = 1, \tag{E.54}$$

and divide both sides of the equation by $\prod_{l=1}^{N}x_l z_l$,

$$\Rightarrow\sum_{m=0}^{N}(-1)^{m}\sum_{\alpha\in\mathcal{R}_m^{(N)}}\sum_{\beta\in\mathcal{R}_m^{(N)}}\left(\det\left[\mathbf{U}\left(\beta,\alpha\right)\right]\right)^2\left(\prod_{j\in\bar{\alpha}}x_j^{-1}\right)\left(\prod_{m\in\bar{\beta}}z_m^{-1}\right)f^{\mathrm{PI}}(\mathbf{x},\mathbf{z}) = \prod_{l=1}^{N}x_l^{-1}z_l^{-1}, \tag{E.55}$$

where $\bar{a}(\bar{\beta})$ is the relative complement of $a(\beta)$ in $\sigma^{(o)} = \{1, 2, \ldots, N\}$. Now,

$$\begin{aligned} f^{\mathrm{PI}}(\mathbf{x}, \mathbf{z}) &= \mathcal{T}_{\mathbf{i},\mathbf{n}} \left[B^{(\mathbf{i})}_{\mathbf{n}}\right](\mathbf{x}, \mathbf{z}) \\ &= \mathcal{T}_{\mathbf{i}_{\bar{a}},\mathbf{n}_{\bar{\beta}}} \left[\mathcal{T}_{\mathbf{i}_a,\mathbf{n}_\beta} \left[B^{(\mathbf{i})}_{\mathbf{n}}\right](\mathbf{x}_a, \mathbf{z}_\beta)\right](\mathbf{x}_{\bar{a}}, \mathbf{z}_{\bar{\beta}}) \\ &= \sum_{\mathbf{i}_{\bar{a}} \in \mathbb{N}_o^{|\bar{a}|}} \sum_{\mathbf{n}_{\bar{\beta}} \in \mathbb{N}_o^{|\bar{\beta}|}} \mathcal{T}_{\mathbf{i}_a,\mathbf{n}_\beta} \left[B^{(\mathbf{i})}_{\mathbf{n}}\right](\mathbf{x}_a, \mathbf{z}_\beta) \left(\prod_{s \in \bar{a}} x_s^{i_s}\right) \left(\prod_{r \in \bar{\beta}} z_r^{n_r}\right), \end{aligned}$$

where $\mathbf{i}_a$ is the subset of $\mathbf{i}$ containing the elements whose index is in a. For any sequence $\left\{c_{\mathbf{i}_{\bar{a}},\mathbf{n}_{\bar{\beta}}}\right\}$, one can show that,

$$\sum_{\mathbf{i}_{\bar{a}} \in \mathbb{N}_o^{|\bar{a}|}} \sum_{\mathbf{n}_{\bar{\beta}} \in \mathbb{N}_o^{|\bar{\beta}|}} c_{\mathbf{i}_{\bar{a}},\mathbf{n}_{\bar{\beta}}} = \sum_{\gamma \in \mathcal{P}(\bar{a})} \sum_{\omega \in \mathcal{P}(\bar{\beta})} \sum_{\mathbf{i}_\gamma \in \mathbb{N}_1^{|\gamma|}} \sum_{\mathbf{n}_\omega \in \mathbb{N}_1^{|\omega|}} c_{\mathbf{i}_{\bar{a}},\mathbf{n}_{\bar{\beta}}}\Big|_{\mathbf{i}_{\bar{a}\setminus\gamma}=o, \mathbf{n}_{\bar{\beta}\setminus\omega}=o}, \tag{E.56}$$

where $\mathcal{P}(a)$ denotes the power set of a and $\mathbb{N}_1$ is the set of all natural numbers not including zero, so that

$$\begin{aligned} f^{\mathrm{PI}}(\mathbf{x}, \mathbf{z}) = \sum_{\gamma \in \mathcal{P}(\bar{a})} \sum_{\omega \in \mathcal{P}(\bar{\beta})} \sum_{\mathbf{i}_\gamma \in \mathbb{N}_1^{|\gamma|}} \sum_{\mathbf{n}_\omega \in \mathbb{N}_1^{|\omega|}} \Bigg(& \mathcal{T}_{\mathbf{i}_a,\mathbf{n}_\beta} \left[B^{(\mathbf{i})}_{\mathbf{n}}\right](\mathbf{x}_a, \mathbf{z}_\beta) \\ & \times \left(\prod_{s \in \bar{a}} x_s^{i_s}\right) \left(\prod_{r \in \bar{\beta}} z_r^{n_r}\right)\Bigg)\Bigg|_{\mathbf{i}_{\bar{a}\setminus\gamma}=o, \mathbf{n}_{\bar{\beta}\setminus\omega}=o}, \end{aligned}$$

$$\begin{aligned} f^{\mathrm{PI}}(\mathbf{x}, \mathbf{z}) = \sum_{\gamma \in \mathcal{P}(\bar{a})} \sum_{\omega \in \mathcal{P}(\bar{\beta})} \sum_{\mathbf{i}_\gamma \in \mathbb{N}_1^{|\gamma|}} \sum_{\mathbf{n}_\omega \in \mathbb{N}_1^{|\omega|}} & \left(\prod_{s \in \gamma} x_s^{i_s}\right) \left(\prod_{r \in \omega} z_r^{n_r}\right) \\ & \times \mathcal{T}_{\mathbf{i}_a,\mathbf{n}_\beta} \left[B^{\left(\mathbf{i},\mathbf{i}_{\bar{a}\setminus\gamma}=o\right)}_{\mathbf{n},\mathbf{n}_{\bar{\beta}\setminus\omega}=o}\right](\mathbf{x}_a, \mathbf{z}_\beta), \end{aligned}$$

$$\begin{aligned} f^{\mathrm{PI}}(\mathbf{x}, \mathbf{z}) = \sum_{\gamma \in \mathcal{P}(\bar{a})} \sum_{\omega \in \mathcal{P}(\bar{\beta})} \sum_{\mathbf{i}_\gamma \in \mathbb{N}_o^{|\gamma|}} \sum_{\mathbf{n}_\omega \in \mathbb{N}_o^{|\omega|}} & \left(\prod_{s \in \gamma} x_s^{i_s+1}\right) \left(\prod_{r \in \omega} z_r^{n_r+1}\right) \\ & \times \mathcal{T}_{\mathbf{i}_a,\mathbf{n}_\beta} \left[B^{\left(\mathbf{i}+1_N^{(\gamma)},\mathbf{i}_{\bar{a}\setminus\gamma}=o\right)}_{\mathbf{n}+1_N^{(\omega)},\mathbf{n}_{\bar{\beta}\setminus\omega}=o}\right](\mathbf{x}_a, \mathbf{z}_\beta). \end{aligned}$$

Since

$$\left(\prod_{j \in \bar{a}} x_j^{-1}\right) \left(\prod_{m \in \bar{\beta}} z_m^{-1}\right) = \left(\prod_{s \in \gamma} x_s^{-1}\right) \left(\prod_{r \in \gamma} z_r^{-1}\right) \left(\prod_{j \in \bar{a}\setminus\gamma} x_j^{-1}\right) \left(\prod_{m \in \bar{\beta}\setminus\omega} z_m^{-1}\right), \tag{E.57}$$

we get

$$\left(\prod_{j\in\bar{a}} x_j^{-1}\right)\left(\prod_{m\in\bar{\beta}} z_m^{-1}\right) f^{\mathrm{PI}}(\mathbf{x},\mathbf{z}) = \sum_{\gamma\in\mathcal{P}(\bar{a})}\sum_{\omega\in\mathcal{P}(\bar{\beta})}\left(\prod_{j\in\bar{a}\setminus\gamma} x_j^{-1}\right)\left(\prod_{m\in\bar{\beta}\setminus\omega} z_m^{-1}\right) \times \sum_{\mathbf{i}_\gamma\in\mathbb{N}_0^{|\gamma|}}\sum_{\mathbf{n}_\omega\in\mathbb{N}_0^{|\omega|}}\left(\prod_{s\in\gamma} x_s^{i_s}\right)\left(\prod_{r\in\omega} z_r^{n_r}\right) \times \mathcal{T}_{\mathbf{i}_a,\mathbf{n}_\beta}\left[B^{\left(\mathbf{i}+1_N^{(\gamma)},\mathbf{i}_{\bar{a}\setminus\gamma}=0\right)}_{\mathbf{n}+1_N^{(\omega)},\mathbf{n}_{\bar{\beta}\setminus\omega}=0}\right]\left(\mathbf{x}_a,\mathbf{z}_\beta\right).$$

From Equation (E.55),

$$\prod_{l=1}^{N}(x_l z_l)^{-1} = \sum_{m=0}^{N}(-1)^m \sum_{a\in\mathcal{R}_m^{(N)}}\sum_{\beta\in\mathcal{R}_m^{(N)}} |\mathbf{U}_{\beta,a}|^2 \sum_{\gamma\in\mathcal{P}(\bar{a})}\sum_{\omega\in\mathcal{P}(\bar{\beta})}\left(\prod_{j\in\bar{a}\setminus\gamma} x_j^{-1}\right)\left(\prod_{m\in\bar{\beta}\setminus\omega} z_m^{-1}\right) \times \sum_{\mathbf{i}_\gamma\in\mathbb{N}_0^{|\gamma|}}\sum_{\mathbf{n}_\omega\in\mathbb{N}_0^{|\omega|}}\left(\prod_{s\in\gamma} x_s^{i_s}\right)\left(\prod_{r\in\omega} z_r^{n_r}\right) \times \mathcal{T}_{\mathbf{i}_a,\mathbf{n}_\beta}\left[B^{\left(\mathbf{i}+1_N^{(\gamma)},\mathbf{i}_{\bar{a}\setminus\gamma}=0\right)}_{\mathbf{n}+1_N^{(\omega)},\mathbf{n}_{\bar{\beta}\setminus\omega}=0}\right]\left(\mathbf{x}_a,\mathbf{z}_\beta\right),$$

with $\mathbf{U}(\beta,a) = \mathbf{U}_{\beta,a}$ and $|\mathbf{U}_{\beta,a}| = \det[\mathbf{U}(\beta,a)]$. We define $\tilde{\gamma} = \bar{a}\setminus\gamma$, and change the summation over the γ into a summation over the $\tilde{\gamma}$, and do the same for the ω, before doing the change of variables $\tilde{\gamma}$ to γ and $\tilde{\omega}$ to ω, getting

$$\prod_{l=1}^{N}(x_l z_l)^{-1} = \sum_{m=0}^{N}(-1)^m \sum_{a\in\mathcal{R}_m^{(N)}}\sum_{\beta\in\mathcal{R}_m^{(N)}} |\mathbf{U}_{\beta,a}|^2 \sum_{\gamma\in\mathcal{P}(\bar{a})}\sum_{\omega\in\mathcal{P}(\bar{\beta})}\left(\prod_{j\in\gamma} x_j^{-1}\right)\left(\prod_{m\in\omega} z_m^{-1}\right) \times \sum_{\mathbf{i}_{\bar{a}\setminus\gamma}\in\mathbb{N}_0^{|\bar{a}|-|\gamma|}}\sum_{\mathbf{n}_{\bar{\beta}\setminus\omega}\in\mathbb{N}_0^{|\bar{\beta}|-|\omega|}}\left(\prod_{s\in\bar{a}\setminus\gamma} x_s^{i_s}\right)\left(\prod_{r\in\bar{\beta}\setminus\omega} z_r^{n_r}\right) \times \mathcal{T}_{\mathbf{i}_a,\mathbf{n}_\beta}\left[B^{\left(\mathbf{i}+1_N^{(\bar{a}\setminus\gamma)},\mathbf{i}_\gamma=0\right)}_{\mathbf{n}+1_N^{(\bar{\beta}\setminus\omega)},\mathbf{n}_\omega=0}\right]\left(\mathbf{x}_a,\mathbf{z}_\beta\right).$$

We add factors such as $\Delta_{\gamma\in\mathcal{P}(\bar{a})}$, which is equal to one if $\gamma \in \mathcal{P}(\bar{a})$ and zero else, in order to remove the dependence on $\bar{a}$ in the sum over the γ (same for the sum over the ω),

$$\begin{aligned}\prod_{l=1}^{N}(x_l z_l)^{-1} = \sum_{m=0}^{N}(-1)^m \sum_{a\in\mathcal{R}_m^{(N)}} \sum_{\beta\in\mathcal{R}_m^{(N)}} |\mathbf{U}_{\beta,a}|^2 \sum_{\gamma\in\mathcal{P}_N}\sum_{\omega\in\mathcal{P}_N} \Delta_{\gamma\in\mathcal{P}(\bar{a})}\Delta_{\omega\in\mathcal{P}(\bar{\beta})} \left(\prod_{j\in\gamma} x_j^{-1}\right)\left(\prod_{m\in\omega} z_m^{-1}\right) \\ \times \sum_{\mathbf{i}_{\bar{a}\setminus\gamma}\in\mathbb{N}_0^{|\bar{a}|-|\gamma|}} \sum_{\mathbf{n}_{\bar{\beta}\setminus\omega}\in\mathbb{N}_0^{|\bar{\beta}|-|\omega|}} \left(\prod_{s\in\bar{a}\setminus\gamma} x_s^{i_s}\right)\left(\prod_{r\in\bar{\beta}\setminus\omega} z_r^{n_r}\right) \\ \times \mathcal{T}_{\mathbf{i}_a,\mathbf{n}_\beta}\left[B^{\left(\mathbf{i}+1_N^{(\bar{a}\setminus\gamma)},\mathbf{i}_\gamma=0\right)}_{\mathbf{n}+1_N^{(\bar{\beta}\setminus\omega)},\mathbf{n}_\omega=0}\right]\left(\mathbf{x}_a,\mathbf{z}_\beta\right).\end{aligned}$$

Rearranging some terms, we get

$$\begin{aligned}\prod_{l=1}^{N}(x_l z_l)^{-1} = \sum_{\gamma\in\mathcal{P}_N}\sum_{\omega\in\mathcal{P}_N}\left(\prod_{j\in\gamma} x_j^{-1}\right)\left(\prod_{m\in\omega} z_m^{-1}\right)\sum_{m=0}^{N}(-1)^m \sum_{a\in\mathcal{R}_m^{(N)}}\sum_{\beta\in\mathcal{R}_m^{(N)}} |\mathbf{U}_{\beta,a}|^2\Delta_{\gamma\in\mathcal{P}(\bar{a})}\Delta_{\omega\in\mathcal{P}(\bar{\beta})} \\ \times \sum_{\mathbf{i}_{\bar{a}\setminus\gamma}\in\mathbb{N}_0^{|\bar{a}|-|\gamma|}} \sum_{\mathbf{n}_{\bar{\beta}\setminus\omega}\in\mathbb{N}_0^{|\bar{\beta}|-|\omega|}} \left(\prod_{s\in\bar{a}\setminus\gamma} x_s^{i_s}\right)\left(\prod_{r\in\bar{\beta}\setminus\omega} z_r^{n_r}\right) \\ \times \mathcal{T}_{\mathbf{i}_a,\mathbf{n}_\beta}\left[B^{\left(\mathbf{i}+1_N^{(\bar{a}\setminus\gamma)},\mathbf{i}_\gamma=0\right)}_{\mathbf{n}+1_N^{(\bar{\beta}\setminus\omega)},\mathbf{n}_\omega=0}\right]\left(\mathbf{x}_a,\mathbf{z}_\beta\right).\end{aligned}$$

Since $\bar{a}\cup a = \bar{\beta}\cup\beta = \sigma^{(0)}$, we have $(\bar{a}\setminus\gamma)\cup a = \sigma^{(0)}\setminus\gamma$, and similarly, $(\bar{\beta}\setminus\omega)\cup\beta = \sigma^{(0)}\setminus\omega$,

$$\begin{aligned}\prod_{l=1}^{N}(x_l z_l)^{-1} = \sum_{\gamma\in\mathcal{P}_N}\sum_{\omega\in\mathcal{P}_N}\left(\prod_{j\in\gamma} x_j^{-1}\right)\left(\prod_{m\in\omega} z_m^{-1}\right)\sum_{m=0}^{N}(-1)^m \sum_{a\in\mathcal{R}_m^{(N)}}\sum_{\beta\in\mathcal{R}_m^{(N)}} |\mathbf{U}_{\beta,a}|^2\Delta_{\gamma\in\mathcal{P}(\bar{a})}\Delta_{\omega\in\mathcal{P}(\bar{\beta})} \\ \times \mathcal{T}_{\mathbf{i}_{\sigma^{(0)}\setminus\gamma},\mathbf{n}_{\sigma^{(0)}\setminus\omega}}\left[B^{\left(\mathbf{i}+1_N^{(\bar{a}\setminus\gamma)},\mathbf{i}_\gamma=0\right)}_{\mathbf{n}+1_N^{(\bar{\beta}\setminus\omega)},\mathbf{n}_\omega=0}\right]\left(\mathbf{x}_{\sigma^{(0)}\setminus\gamma},\mathbf{z}_{\sigma^{(0)}\setminus\omega}\right).\end{aligned}$$

Since in the generating function, only the probability depends on either a, β or m, we can incorporate the summations over these three sets into the generating function, getting

$$\begin{aligned}\prod_{l=1}^{N}(x_l z_l)^{-1} = \sum_{\gamma\in\mathcal{P}_N}\sum_{\omega\in\mathcal{P}_N}\left(\prod_{j\in\gamma} x_j^{-1}\right)\left(\prod_{m\in\omega} z_m^{-1}\right) \\ \times \mathcal{T}_{\left(\mathbf{i}_{\sigma^{(0)}\setminus\gamma},\mathbf{n}_{\sigma^{(0)}\setminus\omega}\right)}\left[\sum_{m=0}^{N}(-1)^m \sum_{a\in\mathcal{R}_m^{(N)}}\sum_{\beta\in\mathcal{R}_m^{(N)}} |\mathbf{U}_{\beta,a}|^2\Delta_{\gamma\in\mathcal{P}(\bar{a})}\Delta_{\omega\in\mathcal{P}(\bar{\beta})} \right. \\ \left. \times B^{\left(\mathbf{i}+1_N^{(\bar{a}\setminus\gamma)},\mathbf{i}_\gamma=0\right)}_{\mathbf{n}+1_N^{(\bar{\beta}\setminus\omega)},\mathbf{n}_\omega=0}\right]\left(\mathbf{x}_{\sigma^{(0)}\setminus\gamma},\mathbf{z}_{\sigma^{(0)}\setminus\omega}\right).\end{aligned} \tag{E.58}$$

If $\gamma = \omega = \sigma^{(0)}$, the corresponding term in the right-hand side of Equation (E.58) is equal to the left-hand side of the latter. If $\gamma = \omega = \emptyset$, the corresponding term in the right-hand side of Equation (E.58) is actually the generating function of

$$\sum_{m=0}^{N} (-1)^m \sum_{a \in \mathcal{R}_m^{(N)}} \sum_{\beta \in \mathcal{R}_m^{(N)}} \left(\det\left[\mathbf{U}\left(\beta, a\right)\right]\right)^2 B_{\mathbf{n}+1_N^{(\bar{\beta})}}^{\left(\mathbf{i}+1_N^{(\bar{a})}\right)}, \tag{E.59}$$

which, according to what we conjectured, should be equal to zero. We can write Equation (E.58) as

$$\begin{aligned}\prod_{l=1}^{N} (x_l z_l)^{-1} = \sum_{\gamma \in \mathcal{P}_N} \sum_{\omega \in \mathcal{P}_N} \left(\prod_{j \in \gamma} x_j^{-1}\right) \left(\prod_{m \in \omega} z_m^{-1}\right) \\ \times \mathcal{T}_{\left(\mathbf{i}_{\sigma^{(0)} \setminus \gamma}, \mathbf{n}_{\sigma^{(0)} \setminus \omega}\right)} \left[\sum_{m=0}^{N} (-1)^m \sum_{a \in \mathcal{R}_m^{(N)}} \sum_{\beta \in \mathcal{R}_m^{(N)}} \left|\mathbf{U}_{\beta,a}\right|^2 \Delta_{\gamma \in \mathcal{P}(\bar{a})} \Delta_{\omega \in \mathcal{P}(\bar{\beta})} \right. \\ \left. \times B_{\mathbf{n}+1_N^{(\bar{\beta})} - 1_N^{(\omega)}, \mathbf{n}_\omega = 0}^{\left(\mathbf{i}+1_N^{(\bar{a})} - 1_N^{(\gamma)}, \mathbf{i}_\gamma = 0\right)} \right] \left(\mathbf{x}_{\sigma^{(0)} \setminus \gamma}, \mathbf{z}_{\sigma^{(0)} \setminus \omega}\right).\end{aligned} \tag{E.60}$$

Now, if when $\mathbf{i}_\gamma = 0$ (as it is the case in the probabilities), we only consider the a such that $\mathbf{i}_s \neq 0, \forall s \in a$, it will mean that $\gamma \subset \bar{a}$, which is true if and only if $\gamma \in \mathcal{P}(\bar{a})$ (same reasoning for the ω). We can therefore remove the Δ factors, and add a condition in the summations over the a and β, ending up with

$$\begin{aligned}\prod_{l=1}^{N} (x_l z_l)^{-1} = \sum_{\gamma \in \mathcal{P}_N} \sum_{\omega \in \mathcal{P}_N} \left(\prod_{j \in \gamma} x_j^{-1}\right) \left(\prod_{m \in \omega} z_m^{-1}\right) \\ \times \mathcal{T}_{\left(\mathbf{i}_{\sigma^{(0)} \setminus \gamma}, \mathbf{n}_{\sigma^{(0)} \setminus \omega}\right)} \left[\sum_{m=0}^{N} (-1)^m \sum_{\substack{a \in \mathcal{R}_m^{(N)} \\ \mathbf{i}_s \neq 0, s \in a}} \sum_{\substack{\beta \in \mathcal{R}_m^{(N)} \\ \mathbf{n}_r \neq 0, r \in \beta}} \left|\mathbf{U}_{\beta,a}\right|^2 \right. \\ \left. \times B_{\mathbf{n}+1_N^{(\bar{\beta})} - 1_N^{(\omega)}, \mathbf{n}_\omega = 0}^{\left(\mathbf{i}+1_N^{(\bar{a})} - 1_N^{(\gamma)}, \mathbf{i}_\gamma = 0\right)} \right] \left(\mathbf{x}_{\sigma^{(0)} \setminus \gamma}, \mathbf{z}_{\sigma^{(0)} \setminus \omega}\right).\end{aligned} \tag{E.61}$$

The relation we conjectured is

$$B_{\mathbf{n}}^{(\mathbf{i})} = \sum_{m=1}^{N} (-1)^{m-1} \sum_{\substack{a \in \mathcal{R}_m^{(N)} \\ \mathbf{i}_s \neq 0, s \in a}} \sum_{\substack{\beta \in \mathcal{R}_m^{(N)} \\ \mathbf{n}_r \neq 0, r \in \beta}} \left(\det\left[\mathbf{U}\left(\beta, a\right)\right]\right)^2 B_{\mathbf{n}-1_N^{(\beta)}}^{\left(\mathbf{i}-1_N^{(a)}\right)}. \tag{E.62}$$

It can be rewritten

$$\sum_{m=0}^{N}(-1)^m \sum_{\substack{a\in\mathcal{R}_m^{(N)}\\ \mathbf{i}_s\neq 0, s\in a}} \sum_{\substack{\beta\in\mathcal{R}_m^{(N)}\\ \mathbf{n}_r\neq 0, r\in\beta}} \left(\det\left[\mathbf{U}\left(\beta,a\right)\right]\right)^2 B_{\mathbf{n}-1_N^{(\beta)}}^{\left(\mathbf{i}-1_N^{(a)}\right)} = 0. \tag{E.63}$$

Choose $\mathbf{i}_\gamma = 0$ and $\mathbf{n}_\omega = 0$ in the last relation, *i.e.*,

$$\sum_{m=0}^{N}(-1)^m \sum_{\substack{a\in\mathcal{R}_m^{(N)}\\ \mathbf{i}_s\neq 0, s\in a}} \sum_{\substack{\beta\in\mathcal{R}_m^{(N)}\\ \mathbf{n}_r\neq 0, r\in\beta}} \left(\det\left[\mathbf{U}\left(\beta,a\right)\right]\right)^2 B_{\mathbf{n}-1_N^{(\beta)}}^{\left(\mathbf{i}-1_N^{(a)}\right)}\Bigg|_{\mathbf{n}_\omega=0,\mathbf{i}_\gamma=0}. \tag{E.64}$$

We can always choose to add 1 to all the elements of $\mathbf{i}_{\sigma^{(0)}\setminus\gamma}$ and $\mathbf{n}_{\sigma^{(0)}\setminus\omega}$, without changing the relation,

$$\sum_{m=0}^{N}(-1)^m \sum_{\substack{a\in\mathcal{R}_m^{(N)}\\ \mathbf{i}_s\neq 0, s\in a}} \sum_{\substack{\beta\in\mathcal{R}_m^{(N)}\\ \mathbf{n}_r\neq 0, r\in\beta}} \left(\det\left[\mathbf{U}\left(\beta,a\right)\right]\right)^2 B_{\mathbf{n}+1_N^{(\bar{\beta})}-1_N^{(\omega)}}^{\left(\mathbf{i}+1_N^{(\bar{a})}-1_N^{(\gamma)}\right)}\Bigg|_{\mathbf{n}_\omega=0,\mathbf{i}_\gamma=0}, \tag{E.65}$$

which is exactly the term appearing in the generating function in Equation (E.61). This actually means that, if the recurrence relation is true for $\mathbf{i}$ and $\mathbf{n}$ such that $\mathbf{i}_\gamma = 0$ and $\mathbf{n}_\omega = 0$, for all $(\gamma+\omega) \in ((\mathcal{P}_N + \mathcal{P}_N)\setminus\emptyset)$, then it is true for $\gamma = \emptyset$ and $\omega = \emptyset$. Since the relation is obviously true for $\gamma = \sigma^{(0)}$ and $\omega = \sigma^{(0)}$, we can proceed by recurrence to show that it is true for any γ and ω.

F Fock-majorization in passive-environment channels

F.1 Fock-majorization preservation in phase-conjugate channels

Lemma 17. *The quantum-limited phase-conjugate channel $\tilde{\mathcal{A}}_G$ of arbitrary gain G exhibits a ladder of Fock-majorization relations*

$$\tilde{\mathcal{A}}_G\big[\,|i\rangle\langle i|\,\big] \succ_{\mathrm{F}} \tilde{\mathcal{A}}_G\big[\,|i+1\rangle\langle i+1|\,\big], \quad \forall i \geq 0. \tag{F.1}$$

Proof. We have

$$\omega^{(i)} := \tilde{\mathcal{A}}_G\big[\,|i\rangle\langle i|\,\big] = \sum_{n=0}^{\infty} A_{i+n}^{(i,0)}\,|n\rangle\langle n|\,, \tag{F.2}$$

where the $A_{i+n}^{(i,0)}$ are defined in Equation (9.33). Note that the diagonal of $\mathcal{A}_G\big[\,|i\rangle\langle i|\,\big]$ in fact corresponds to the diagonal of $\tilde{\mathcal{A}}_G\big[\,|i\rangle\langle i|\,\big]$, shifted by an index i. The differences between the cumulated sums of eigenvalues are given by

$$\sum_{n=0}^{j} A_{i+n}^{(i,0)} - \sum_{n=0}^{j} A_{i+1+n}^{(i+1,0)} = (G-1)\,A_{i+1+j}^{(i+1,0)} \geq 0, \forall j \geq 0, \tag{F.3}$$

where we used Equation (9.35) again. This gives the Fock-majorization relation $\omega^{(i)} \succ_{\mathrm{F}} \omega^{(i+1)}$. □

Corollary 8. *Phase-conjugate Gaussian bosonic channels are majorization-preserving over the set of passive states.*

Proof. Using Lemma 17 together with Theorem 33, we obtain that the quantum-limited phase-conjugate channel $\tilde{\mathcal{A}}_G$, whose adjoint is $1/(G-1)$ times the quantum-limited phase-conjugate channel $\tilde{\mathcal{A}}_{G/(G-1)}$, is passive preserving. Since any phase-conjugate Gaussian bosonic channel $\mathcal{G}$ can be expressed as the concatenation of a pure loss channel $\mathcal{B}$ and a quantum-limited phase-conjugate channel $\tilde{\mathcal{A}}$, i.e., $\mathcal{G} = \tilde{\mathcal{A}} \circ \mathcal{B}$, and since passive preservation is transitive over channel composition, we deduce (following the reasoning of Theorems 36 and 37, and Corollary 6) that phase-conjugate Gaussian bosonic channels are passive preserving, Fock-majorization preserving, and majorization- reserving over the set of passive states. □

F.2 Theorem 34 for passive-environment channels

Here, we simply explain that any passive-environment channel $\mathcal{B}_\eta^{\downarrow}$ satisfies both

$$\langle n|\left(\mathcal{B}_\eta^{\downarrow}\left[|i\rangle\langle j|\right]\right)|n\rangle = 0, \quad \forall i \neq j, \tag{F.4}$$

and

$$\langle n| \left(\mathcal{B}_\eta^\downarrow \left[|i\rangle \langle i| \right] \right) |m\rangle = 0, \quad \forall n \neq m. \tag{F.5}$$

Proof. We need to show that if ρ is diagonal in the Fock basis, $\mathcal{B}_\eta^\downarrow[\rho]$ is also diagonal in the Fock basis, while if ρ is non-diagonal in the Fock basis, its non-diagonal elements do not contribute to the diagonal elements of $\mathcal{B}_\eta^\downarrow[\rho]$. We have that

$$U_\eta^{\mathrm{BS}} |i,k\rangle = \sum_{n=0}^{i+k} b_n^{(i,k)} |n, i+k-n\rangle , \tag{F.6}$$

where the $b_n^{(i,k)} \in \mathbb{C}$ are defined in Equation (5.29). If we define our passive channel as in Equation (7.7), we have

$$\begin{aligned} \mathcal{B}_\eta^\downarrow[|i\rangle \langle j|] &= \sum_k \lambda_k^\downarrow \mathrm{Tr}_E \left[U_\eta^{\mathrm{BS}} \left(|i\rangle \langle j| \otimes |k\rangle \langle k| \right) U_\eta^{\mathrm{BS}\dagger} \right] \\ &= \sum_k \lambda_k^\downarrow \sum_l \langle l|_E \left(\sum_{n=0}^{i+k} b_n^{(i,k)} |n, i+k-n\rangle \right) \\ &\qquad \times \left(\sum_{m=0}^{j+k} b_m^{(j,k)*} \langle m, j+k-m| \right) |l\rangle_E \\ &= \sum_k \lambda_k^\downarrow \sum_n b_n^{(i,k)} b_{n+j-i}^{(j,k)*} |n\rangle \langle n+j-i| . \end{aligned} \tag{F.7}$$

We end up with

$$\mathcal{B}_\eta^\downarrow[|i\rangle \langle i|] = \sum_k \lambda_k^\downarrow \sum_n |b_n^{(i,k)}|^2 |n\rangle \langle n| , \tag{F.8}$$

which means that if ρ is diagonal in the Fock basis, $\mathcal{B}_\eta^\downarrow[\rho]$ is also diagonal in the Fock basis. Furthermore, Equation (F.7) tells us that if ρ is non-diagonal in the Fock basis, its non-diagonal elements do not contribute to the diagonal elements of $\mathcal{B}_\eta^\downarrow[\rho]$. □

Now, the exact same reasoning can be applied to passive-environment channels $\mathcal{A}_G^\downarrow$ defined as

$$\mathcal{A}_G^\downarrow[\rho] = \mathrm{Tr}_2 \left[U_\lambda^{\mathrm{TMS}} \left(\rho \otimes \sigma^\downarrow \right) U_\lambda^{\mathrm{TMS}\dagger} \right] , \tag{F.9}$$

for any passive environment $\sigma^\downarrow$, by obviously replacing the beam splitter U_η^{BS} by the two-mode squeezer U_λ^{TMS}, and the amplitudes $b_n^{(i,k)}$ by $a_n^{(i,k)}$. This means that the channels $\mathcal{A}_G^\downarrow$ verify Equations (F.4) and (F.5) as well.

G Conservation of passivity after passive post-selection

In this section, we prove Lemma 16, which states that for any two passive states $\sigma_1^{\downarrow}$ and $\sigma_2^{\downarrow}$, the one-mode state

$$\rho = \frac{\tilde{\rho}}{\mathrm{Tr}\left[\tilde{\rho}\right]}, \qquad \text{where } \tilde{\rho} = \mathrm{Tr}_2\left[\left(\mathbb{I}\otimes\sigma_3^{\downarrow}\right) U_\eta^{\mathrm{BS}}\left(\sigma_1^{\downarrow}\otimes\sigma_2^{\downarrow}\right) U_\eta^{\mathrm{BS}\dagger}\right], \tag{G.1}$$

is passive as well, for any one-mode passive state $\sigma_3^{\downarrow}$. In order to show this, we simply generalise the proof of Theorem 39, by including a post-selection on a passive state. For the purpose of symmetry, we define $d_{n,m}^{(i,k)} = B_n^{(i,k)}$, where $m = i + k - n$. Since, $d_{n,m}^{(i,k)} = |\langle n,m| U_\eta^{\mathrm{BS}} |i,k\rangle|^2$, we need to show that

$$\Gamma_{n,M}^{(I,K)} = \sum_{i=0}^{I}\sum_{k=0}^{K}\sum_{m=0}^{M}\left(d_{n,m}^{(i,k)} - d_{n+1,m}^{(i,k)}\right) \geq 0. \quad \forall I,K,M,n \geq 0. \tag{G.2}$$

Using the recurrence relation of Equation (6.1) for $j = 1$ again, we have that

$$\begin{aligned}\Gamma_{n,M}^{(I,K)} = &\sum_{i=0}^{I}\sum_{k=0}^{K}\sum_{m=0}^{M}\left(d_{n,m}^{(i,k)} - \eta d_{n,m}^{(i-1,k)}\right)\\ &- \sum_{i=0}^{I}\sum_{k=0}^{K}\sum_{m=0}^{M}\left((1-\eta)d_{n+1,m-1}^{(i-1,k)} + \eta d_{n+1,m-1}^{(i,k-1)} + (1-\eta)d_{n,m}^{(i,k-1)} - d_{n,m-1}^{(i-1,k-1)}\right).\end{aligned}$$

After some calculations, we get

$$\begin{aligned}\Gamma_{n,M}^{(I,K)} = &\eta\sum_{i=0}^{I}\sum_{k=0}^{K}\sum_{m=0}^{M}\left(d_{n,m}^{(i,k)} - d_{n,m}^{(i-1,k)}\right) + (1-\eta)\sum_{i=0}^{I}\sum_{k=0}^{K}\sum_{m=0}^{M}\left(d_{n,m}^{(i,k)} - d_{n,m}^{(i,k-1)}\right)\\ &- (1-\eta)\sum_{i=0}^{I}\sum_{k=0}^{K}\sum_{m=0}^{M} d_{n+1,m-1}^{(i-1,k)} - \eta\sum_{i=0}^{I}\sum_{k=0}^{K}\sum_{m=0}^{M} d_{n+1,m-1}^{(i,k-1)} + \sum_{i=0}^{I}\sum_{k=0}^{K}\sum_{m=0}^{M} d_{n,m-1}^{(i-1,k-1)},\\ \Gamma_{n,M}^{(I,K)} = &\eta\sum_{k=0}^{K}\sum_{m=0}^{M} d_{n,m}^{(I,k)} + (1-\eta)\sum_{i=0}^{I}\sum_{m=0}^{M} d_{n,m}^{(i,K)}\\ &- (1-\eta)\sum_{i=0}^{I-1}\sum_{k=0}^{K}\sum_{m=0}^{M-1} d_{n+1,m}^{(i,k)} - \eta\sum_{i=0}^{I}\sum_{k=0}^{K-1}\sum_{m=0}^{M-1} d_{n+1,m}^{(i,k)} + \sum_{i=0}^{I-1}\sum_{k=0}^{K-1}\sum_{m=0}^{M-1} d_{n,m}^{(i,k)},\\ \Gamma_{n,M}^{(I,K)} = &\eta\sum_{k=0}^{K}\sum_{m=0}^{M} d_{n,m}^{(I,k)} + (1-\eta)\sum_{i=0}^{I}\sum_{m=0}^{M} d_{n,m}^{(i,K)}\\ &- (1-\eta)\sum_{i=0}^{I-1}\sum_{k=0}^{K}\sum_{m=0}^{M-1} d_{n+1,m}^{(i,k)} - \eta\sum_{i=0}^{I}\sum_{k=0}^{K-1}\sum_{m=0}^{M-1} d_{n+1,m}^{(i,k)}\\ &+ \eta\sum_{i=0}^{I-1}\sum_{k=0}^{K-1}\sum_{m=0}^{M-1} d_{n,m}^{(i,k)} + (1-\eta)\sum_{i=0}^{I-1}\sum_{k=0}^{K-1}\sum_{m=0}^{M-1} d_{n,m}^{(i,k)},\end{aligned}$$

$$\begin{aligned}
\Gamma^{(I,K)}_{n,M} &= \eta \sum_{k=0}^{K-1}\sum_{m=0}^{M} d^{(I,k)}_{n,m} + \eta \sum_{m=0}^{M} d^{(I,K)}_{n,m} + (1-\eta)\sum_{i=0}^{I-1}\sum_{m=0}^{M} d^{(i,K)}_{n,m} + (1-\eta)\sum_{m=0}^{M} d^{(I,K)}_{n,m} \\
&\quad - (1-\eta)\sum_{i=0}^{I-1}\sum_{k=0}^{K}\sum_{m=0}^{M-1} d^{(i,k)}_{n+1,m} - \eta \sum_{i=0}^{I}\sum_{k=0}^{K-1}\sum_{m=0}^{M-1} d^{(i,k)}_{n+1,m} \\
&\quad + \eta \sum_{i=0}^{I-1}\sum_{k=0}^{K-1}\sum_{m=0}^{M-1} d^{(i,k)}_{n,m} + (1-\eta)\sum_{i=0}^{I-1}\sum_{k=0}^{K-1}\sum_{m=0}^{M-1} d^{(i,k)}_{n,m}, \\
\Gamma^{(I,K)}_{n,M} &= \eta \sum_{k=0}^{K-1}\sum_{m=0}^{M-1} d^{(I,k)}_{n,m} + \eta \sum_{k=0}^{K-1} d^{(I,k)}_{n,M} + (1-\eta)\sum_{i=0}^{I-1}\sum_{m=0}^{M-1} d^{(i,K)}_{n,m} + (1-\eta)\sum_{i=0}^{I-1} d^{(i,K)}_{n,M} + \sum_{m=0}^{M} d^{(I,K)}_{n,m} \\
&\quad - (1-\eta)\sum_{i=0}^{I-1}\sum_{k=0}^{K}\sum_{m=0}^{M-1} d^{(i,k)}_{n+1,m} - \eta \sum_{i=0}^{I}\sum_{k=0}^{K-1}\sum_{m=0}^{M-1} d^{(i,k)}_{n+1,m} \\
&\quad + \eta \sum_{i=0}^{I-1}\sum_{k=0}^{K-1}\sum_{m=0}^{M-1} d^{(i,k)}_{n,m} + (1-\eta)\sum_{i=0}^{I-1}\sum_{k=0}^{K-1}\sum_{m=0}^{M-1} d^{(i,k)}_{n,m}, \\
\Gamma^{(I,K)}_{n,M} &= \eta \sum_{k=0}^{K-1} d^{(I,k)}_{n,M} + (1-\eta)\sum_{i=0}^{I-1} d^{(i,K)}_{n,M} + \sum_{m=0}^{M} d^{(I,K)}_{n,m} \\
&\quad - (1-\eta)\sum_{i=0}^{I-1}\sum_{k=0}^{K}\sum_{m=0}^{M-1} d^{(i,k)}_{n+1,m} - \eta \sum_{i=0}^{I}\sum_{k=0}^{K-1}\sum_{m=0}^{M-1} d^{(i,k)}_{n+1,m} \\
&\quad + \eta \sum_{i=0}^{I}\sum_{k=0}^{K-1}\sum_{m=0}^{M-1} d^{(i,k)}_{n,m} + (1-\eta)\sum_{i=0}^{I-1}\sum_{k=0}^{K}\sum_{m=0}^{M-1} d^{(i,k)}_{n,m}, \\
\Gamma^{(I,K)}_{n,M} &= \eta \sum_{k=0}^{K-1} d^{(I,k)}_{n,M} + (1-\eta)\sum_{i=0}^{I-1} d^{(i,K)}_{n,M} + \sum_{m=0}^{M} d^{(I,K)}_{n,m} \\
&\quad + \eta \sum_{i=0}^{I}\sum_{k=0}^{K-1}\sum_{m=0}^{M-1} \left(d^{(i,k)}_{n,m} - d^{(i,k)}_{n+1,m} \right) + (1-\eta)\sum_{i=0}^{I-1}\sum_{k=0}^{K}\sum_{m=0}^{M-1} \left(d^{(i,k)}_{n,m} - d^{(i,k)}_{n+1,m} \right),
\end{aligned}$$

so that,

$$\begin{aligned}
\Gamma^{(I,K)}_{n,M} &= \eta \sum_{k=0}^{K-1} d^{(I,k)}_{n,M} + (1-\eta)\sum_{i=0}^{I-1} d^{(i,K)}_{n,M} + \sum_{m=0}^{M} d^{(I,K)}_{n,m} \\
&\quad + \eta\Gamma^{(I,K-1)}_{n,M-1} + (1-\eta)\Gamma^{(I-1,K)}_{n,M-1}.
\end{aligned} \tag{G.3}$$

From Lemma 16, we know that $\Gamma^{(I,K)}_{n,I+K-n} \geq 0$, for all $I, K, n \geq 0$. Now, we know that the first three terms on the left-hand side of Equation (G.3) are always positive. Using a recursive argument for increasing values of I, K starting from 0 and decreasing values of M starting from $I + K - n$, one can prove that $\Gamma^{(I,K)}_{n,M} \geq 0$, for all $I, K, n, M \geq 0$. This, along with the fact that any passive state can be written as a convex sum of projectors $P^{\downarrow}_{l}$, ends the proof.

References

[1] C. E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27:379–423 and 623–656, 1948.

[2] J. von Neumann. *Mathematische Grundlagen der Quantenmechanik*. Springer, 1932.

[3] A. S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Transactions on Information Theory*, 44(1):269–273, 1998.

[4] B. Schumacher and M. D. Westmoreland. Strong no-go theorem for gaussian quantum bit commitment. *Phys. Rev. A*, 56(1):131–138, 1997.

[5] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and its Applications*. Academic Press, 1979.

[6] M. A. Nielsen. Conditions for a Class of Entanglement Transformations. *Phys. Rev. Lett.*, 83(2):436–439, 1999.

[7] S. Guha. *Multiple-User Quantum Information Theory for Optical Communication Channels*. PhD thesis, Massachusetts Institute of Technology, 2008.

[8] J. Eisert, S. Scheel, and M. B. Plenio. Distilling Gaussian States with Gaussian Operations is Impossible. *Phys. Rev. Lett.*, 89(13):137903, 4 pp, 2002.

[9] G. Giedke and J. I. Cirac. Characterization of Gaussian operations and distillation of Gaussian states. *Phys. Rev. A*, 66(3):032316, 7 pp, 2002.

[10] J. Fiurášek. Gaussian Transformations and Distillation of Entangled Gaussian States. *Phys. Rev. Lett.*, 89(13):137904, 4 pp, 2002.

[11] A. Peruzzo, A. Laing, A. Politi, T. Rudolph, and J. L. O'Brien. Multimode quantum interference of photons in multiport integrated devices. *Nature Comm.*, 2(224):6 pp, 2011.

[12] S. Aaronson and A. Arkhipov. The Computational Complexity of Linear Optics. In *Proceedings of STOC '11 43rd Annual ACM Symposium on Theory of Computing*, pages 333–342, 2011.

[13] M. Tillmann, B. Dakić, R. Heilmann, S. Nolte, A. Szameit, and P. Walther. Experimental boson sampling. *Nature Photonics*, 7:540–544, 2013.

[14] A. Crespi, R. 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. *Nature Photonics*, 7:545–549, 2013.

[15] 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(6121):794–798, 2013.

[16] J. Carolan, J. D. A. Meinecke, P. J. Shadbolt, N. J. Russell, N. Ismail, K. Wörhoff, T. Rudolph, M. G. Thompson, J. L. O'Brien, J. C. F. Matthews, and A. Laing. On the experimental verification of quantum complexity in linear optics. *Nature Photonics*, 8:621–626, 2014.

[17] C. K. Hong, Z. Y. Ou, and L. Mandel. Measurement of subpicosecond time intervals between two photons by interference. *Phys. Rev. Lett.*, 59(18):2044–2046, 1987.

[18] V. S. Shchesnovich. Universality of Generalized Bunching and Efficient Assessment of Boson Sampling. *Phys. Rev. Lett.*, 116(12):123601, 5 pp, 2016.

[19] L. Rigovacca, C. Di Franco, B. J. Metcalf, I. A. Walmsley, and M. S. Kim. Nonclassicality Criteria in Multiport Interferometry. *Phys. Rev. Lett.*, 117(21):213602, 6 pp, 2016.

[20] S. Agne, T. Kauten, J. Jin, E. Meyer-Scott, J. Z. Salvail, D. R. Hamel, K. J. Resch, G. Weihs, and T. Jennewein. Observation of Genuine Three-Photon Interference. *Phys. Rev. Lett.*, 118(15):153602, 4 pp, 2017.

[21] 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(15):153603, 6 pp, 2017.

[22] 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. *Nature Comm.*, 7(10469):8 pp, 2016.

[23] M. Tribus and E. C. McIrvine. Energy and Information. *Scientific American*, 224:179–188, 1971.

[24] T. M. Cover and J. A. Thomas. *Elements of information theory*. Wiley, 2 edition, 2006.

[25] M. M. Wilde. *Quantum information theory*. Cambridge University Press, 2013.

[26] D. Petz. Entropy, von Neumann and the von Neumann entropy. In M. Redei and M. Stoeltzner, editors, *John von Neumann and the Foundations of Quantum Physics*, pages 83–96. Springer Netherlands, 2013.

[27] R. Clausius. *Mechanical Theory of Heat*. London: Macmillan & Co., 2 edition, 1879.

[28] A. N. Kolmogorov. On the Shannon theory of information transmission in the case of continuous signals. *IRE Trans. Inf. Theory*, 2(4):102–108, 1956.

[29] A. Rényi. On measures of entropy and information. In *Fourth Berkeley Symp. on Math. Statist. and Prob.*, volume 1, pages 547–561, 1961.

[30] Y. He, A. B. Hamza, and H. Krim. A generalized divergence measure for robust image registration. *IEEE Transactions on Signal Processing*, 51(5):1211–1220, 2003.

[31] T. van Erven and P. Harremos. Rényi Divergence and Kullback-Leibler Divergence. *IEEE Transactions on Information Theory*, 60(7):3797–3820, 2014.

[32] R. F. Muirhead. Some methods applicable to identities and inequalities of symmetric algebraic functions of n letters. *Proc. Edinburgh Math. Soc.*, 21:144–162, 1902.

[33] G. H. Hardy, J. E. Littlewood, and G. Pólya. Some simple inequalities satisfied by convex functions. *Messenger Math.*, 58:145–152, 1929.

[34] I. Schur. Über eine Klasse von Mittelbildungen mit Anwendungen die Determinanten. *Theorie Sitzungsber. Berlin. Math. Gesellschaft*, 22:9–20, 1923.

[35] S. Ho and S. Verdú. On the Interplay Between Conditional Entropy and Error Probability. *IEEE Transactions on Information Theory*, 56(12):5930–5942, 2010.

[36] S. Ho and S. Verdú. Convexity/Concavity of Renyi Entropy and α-Mutual Information. In *IEEE International Symposium on Information Theory*, pages 745–749, 2015.

[37] G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge University Press, 2 edition, 1952.

[38] D. Jonathan and M. B. Plenio. Entanglement-Assisted Local Manipulation of Pure Quantum States. *Phys. Rev. Lett.*, 83(17):3566–3569, 1999.

[39] G. Aubrun and I. Nechita. Catalytic Majorization and l_p Norms. *Commun. Math. Phys.*, 278(1):133–144, 2008.

[40] A. S. Markus. The eigen- and singular values of the sum and product of linear operators. *Uspekhi Mat. Nauk*, 19(4):93–123, 1964.

[41] I. C. Gohberg and A. S. Markus. Some relations between eigenvalues and matrix elements of linear operators. *Mat. Sb.*, 64(4):481–496, 1964.

[42] V. Kaftal and G. Weiss. An infinite dimensional Schur–Horn Theorem and majorization theory. *J. Funct. Anal.*, 259(12):3115–3162, 2010.

[43] L. Wang and M. Madiman. Beyond the Entropy Power Inequality, via Rearrangements. *IEEE Transactions on Information Theory*, 60(9):5116–5137, 2014.

[44] R. Gardner. The Brunn–Minkowski inequality. *Bull. Amer. Math. Soc. (N.S.)*, 39(3):355–405, 2002.

[45] J. von Neumann. Thermodynamic quantummechanischer Gesamheiten. *Gött. Nach.*, 1:273–291, 1927.

[46] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge univeristy press, 2000.

[47] N.J. Cerf and C. Adami. Negative Entropy and Information in Quantum Mechanics. *Phys. Rev. Lett.*, 79(26):5194–5197, 1997.

[48] A. Uhlmann. Sätze über dichtematrizen. *Wiss. Z. Karl-Marx-Univ. Leipzig*, 20:633–637, 1971.

[49] A. Uhlmann. Endlich-dimensionale dichtematrizen I. *Wiss. Z. Karl-Marx-Univ. Leipzig*, 21:421–452, 1972.

[50] A. Uhlmann. Endlich-dimensionale dichtematrizen II. *Wiss. Z. Karl-Marx-Univ. Leipzig*, 22:139–177, 1973.

[51] C. B. Mendl and M. M. Wolf. Unital Quantum Channels – Convex Structure and Revivals of Birkhoff's Theorem. *Commun. Math. Phys.*, 289(3):1057–1086, 2009.

[52] B. Schumacher. Quantum coding. *Phys. Rev. A*, 51(4):2738–2747, 1995.

[53] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81(2):865–942, 2009.

[54] A. Einstein, B. Podolsky, and N. Rosen. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Phys. Rev.*, 47(10):777–780, 1935.

[55] E. Schrödinger. Die gegenwärtige Situation in der Quantenmechanik. *Naturwiss.*, 23:823–828, 1935.

[56] J. S. Bell. On The Einstein Podolsky Rosen Paradox. *Physics*, 1(3):195–200, 1964.

[57] R. F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40(8):4277–4281, 1989.

[58] A. K. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67(6):661–663, 1991.

[59] C. H. Bennett and S. J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.*, 69(20):2881–2884, 1992.

[60] 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(13):1895–1899, 1993.

[61] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher. Concentrating partial entanglement by local operations. *Phys. Rev. A*, 53(4):2046–2052, 1996.

[62] N. J. Cerf and C. Adami. Quantum extension of conditional probability. *Phys. Rev. A*, 60(2):893–897, 1999.

[63] M. A. Nielsen and J. Kempe. Separable States Are More Disordered Globally than Locally. *Phys. Rev. Lett.*, 86(22):5184–5187, 2001.

[64] T. Hiroshima. Majorization Criterion for Distillability of a Bipartite Quantum State. *Phys. Rev. Lett.*, 91(5):057902, 4 pp, 2003.

[65] C. Weedbrook, S. Pirandola, R. García-Patrón, T. Ralph, N. J. Cerf, J. H. Shapiro, and S. Lloyd. Gaussian quantum information. *Rev. Mod. Phys.*, 84(2):62–669, 2012.

[66] G. Adesso, S. Ragy, and A. R. Lee. Continuous Variable Quantum Information: Gaussian States and Beyond. *Open Syst. Inf. Dyn.*, 21:1440001, 47 pp, 2014.

[67] R. Simon, N. Mukunda, and B. Dutta. Quantum-noise matrix for multimode systems: U(n) invariance, squeezing, and normal forms. *Phys. Rev. A*, 49(3):1567–1583, 1994.

[68] R. L. Hudson. When is the Wigner quasi-probability density non-negative? *Rep. Math. Phys.*, 6(2):249–252, 1974.

[69] A. Hertz, M. G. Jabbour, and N. J. Cerf. Entropy-power uncertainty relations: towards a tight inequality for all Gaussian pure states. *J. Phys. A*, 50(38):385301, 20 pp, 2017.

[70] R. Haag, N. M. Hugenholtz, and M. Winnink. On the equilibrium states in quantum statistical mechanics. *Commun. Math. Phys.*, 5(3):215–236, 1967.

[71] H. P. Yuen. Two-photon coherent states of the radiation field. *Phys. Rev. A*, 13(6):2226–2243, 1976.

[72] Arvind, B. Dutta, N. Mukunda, and R. Simon. The real symplectic groups in quantum mechanics and optics. *Pramana J. Phys.*, 45(6):471–497, 1995.

[73] S. L. Braunstein. Squeezing as an irreducible resource. *Phys. Rev. A*, 71(5):055801, 4 pp, 2005.

[74] J. Williamson. On the Algebraic Problem Concerning the Normal Forms of Linear Dynamical Systems. *Am. J. Math.*, 58(1):141–163, 1936.

[75] G. Giedke, J. Eisert, J. I. Cirac, and M. B. Plenio. Entanglement transformations of pure Gaussian states. *Quant. Inf. Comp.*, 3(3):211–223, 2003.

[76] A. S. Holevo. *Quantum Systems, Channels, Information*. De Gruyter, 2012.

[77] A. S. Holevo. One-mode quantum Gaussian channels: Structure and quantum capacity. *Probl. Inf. Transm.*, 43(1):1–11, 2007.

[78] V. Giovannetti, R. García-Patrón, N. J. Cerf, and A. S. Holevo. Ultimate classical communication rates of quantum optical channels. *Nature Photonics*, 8:796–800, 2014.

[79] A. S. Holevo. Entanglement-breaking channels in infinite dimensions. *Probl. Inf. Transm.*, 44(3):171–184, 2008.

[80] F. Caruso, V. Giovannetti, and A. S. Holevo. One-mode bosonic Gaussian channels: a full weak-degradability classification. *New J. Phys.*, 8:310, 18 pp, 2006.

[81] R. García-Patrón, C. Navarrete-Benlloch, S. Lloyd, J. H. Shapiro, and N. J. Cerf. Majorization theory approach to the gaussian channel minimum entropy conjecture. *Phys. Rev. Lett.*, 108(11):110505, 5 pp, 2012.

[82] H.-P. Breuer and F. Petruccione. *The theory of open quantum systems*. Oxford university press, 2002.

[83] V. Giovannetti, A. S. Holevo, S. Lloyd, and L. Maccone. Generalized minimal output entropy conjecture for one-mode Gaussian channels: definitions and some exact results. *J. Phys. A: Math. Theor.*, 43(41):415305, 10 pp, 2010.

[84] O. Rioul. Information Theoretic Proofs of Entropy Power Inequalities. *IEEE Transactions on Information Theory*, 57(1):33–55, 2011.

[85] A. Stam. Some Inequalities Satisfied by the Quantities of Information of Fisher and Shannon. *Inform. and Control*, 2(2):101–112, 1959.

[86] N. Blachman. The Convolution Inequality for Entropy Powers. *IEEE Transactions on Information Theory*, 11(2):267–271, 1965.

[87] A. R. Barron. Entropy and the central limit theorem. *Ann. Probab.*, 14(1):336–342, 1986.

[88] A. Dembo, T. Cover, and J. Thomas. Information Theoretic Inequalities. *IEEE Transactions on Information Theory*, 37(6):1501–1518, 1991.

[89] O. Rioul. A Simple Proof of the Entropy-Power Inequality via Properties of Mutual Information. In *IEEE International Symposium on Information Theory*, pages 46–50, 2007.

[90] S. Guha, B. I. Erkmen, and J. H. Shapiro. The Entropy Photon-Number Inequality and its consequences. In *Information Theory and Applications Workshop*, 2008.

[91] R. König and G. Smith. The Entropy Power Inequality for Quantum Systems. *IEEE Transactions on Information Theory*, 60(3):1536–1548, 2014.

[92] G. De Palma, D. Trevisan, and V. Giovannetti. Gaussian States Minimize the Output Entropy of the One-Mode Quantum Attenuator. *IEEE Transactions on Information Theory*, 63(1):728–737, 2017.

[93] G. Pólya. *Mathematics and plausible reasoning*. Princeton University Press, 1954.

[94] J. Fiurášek J. Niset and N. J. Cerf. No-Go Theorem for Gaussian Quantum Error Correction. *Phys. Rev. Lett.*, 102(12):120501, 4 pp, 2009.

[95] L. Magnin, F. Magniez, A. Leverrier, and N. J. Cerf. Strong no-go theorem for gaussian quantum bit commitment. *Phys. Rev. A*, 81(1):010302, 4 pp, 2010.

[96] H. S. Wilf. *generatingfunctionology*. CRC Press, 2005.

[97] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[98] G. H. Hardy and J. E. Littlewood. New proofs of the prime-number theorem and similar theorems. *Quart. J. Math.*, 46:215–219, 1915.

[99] U. Leonhardt. *Essential Quantum Optics: From Quantum Measurements to Black Holes*. Cambridge University Press, 2010.

[100] N. J. Cerf. *The optical beam splitter under partial time reversal*. 9th Central European Workshop on Quantum Optics (CEWQO 2012), Sinaia, Romania, 2012.

[101] H. Scutaru. Suppression law of quantum states in a 3D photonic fast Fourier transform chip. *J. Phys. A: Math. Gen.*, 31:3659–3663, 1998.

[102] H. Nakazato, S. Pascazio, M. Stobińska, and K. Yuasa. Photon distribution at the output of a beam splitter for imbalanced input states. *Phys. Rev. A*, 93(2):023845, 6 pp, 2016.

[103] C. N. Gagatsos, O. Oreshkov, and N. J. Cerf. Majorization relations and entanglement generation in a beam splitter. *Phys. Rev. A*, 87(4):042307, 9 pp, 2013.

[104] W. Pusz and S. L. Woronowicz. Passive states and kms states for general quantum systems. *Commun. Math. Phys.*, 58(3):273–290, 1978.

[105] A. E. Allahverdyan, R. Balian, and Th. M. Nieuwenhuizen. Maximal work extraction form finite quantum systems. *Europhys. Lett.*, 67:565–571, 2004.

[106] M. Perarnau-Llobet, K. V. Hovhannisyan, M. Huber, P. Skrzypczyk, J. Tura, and A. Acín. Most energetic passive states. *Phys. Rev. E*, 92(4):042147, 5 pp, 2015.

[107] A. Lenard. Thermodynamical proof of the gibbs formula for elementary quantum systems. *J. Stat. Phys.*, 19(6):575–586, 1978.

[108] K. K. Sabapathy and A. Winter. Non-Gaussian operations on bosonic modes of light: Photon-added Gaussian channels. *Phys. Rev. A*, 95(6):062309, 17 pp, 2017.

[109] J. Goold, M. Huber, A. Riera, L. del Rio, and P. Skrzypczyk. The role of quantum information in thermodynamics—a topical review. *J. Phys. A: Math. Theor.*, 49(14):143001, 50 pp, 2016.

[110] V. Narasimhachar and G. Gour. Low-temperature thermodynamics with quantum coherence. *Nature Comm.*, 6(7689):6 pp, 2015.

[111] J. Solomon Ivan, K. K. Sabapathy, and R. Simon. Operator-sum representation for bosonic Gaussian channels. *Phys. Rev. A*, 84(4):042311, 26 pp, 2011.

[112] M. Horodecki and J. Oppenheim. Fundamental limitations for quantum and nanoscale thermodynamics. *Nature Comm.*, 4(2059):6 pp, 2013.

[113] A. S. Holevo and R. F. Werner. Evaluating capacities of bosonic Gaussian channels. *Phys. Rev. A*, 63(3):032312, 14 pp, 2001.

[114] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, and J. H. Shapiro. Minimum output entropy of bosonic channels: A conjecture. *Phys. Rev. A*, 70(3):032315, 14 pp, 2004.

[115] M. G. Jabbour, R. Garcia-Patron, and N. J. Cerf. Interconversion of pure Gaussian states requiring non-Gaussian operations. *Phys. Rev. A*, 91(1):012316, 8 pp, 2015.

[116] G. De Palma, D. Trevisan, and V. Giovannetti. Passive States Optimize the Output of Bosonic Gaussian Quantum Channels. *IEEE Transactions on Information Theory*, 62(5):2895–2906, 2016.

[117] A. Mari, V. Giovannetti, and A.S. Holevo. Quantum state majorization at the output of bosonic Gaussian channels. *Nature Comm.*, 5(3826):5 pp, 2014.

[118] F. G. S. L. Brandão and G. Gour. Reversible Framework for Quantum Resource Theories. *Phys. Rev. Lett.*, 115(7):070503, 5 pp, 2015.

[119] H. Joe. Majorization, Randomness and Dependence for Multivariate Distributions. *Ann. Probab.*, 15(3):1217–1225, 1987.

Index