Technische Universität München
Max-Planck-Institut für Quantenoptik

# Quantum Entanglement: Theory and Applications

Norbert Schuch

Vollständiger Abdruck der von der Fakultät für Physik
der Technischen Universität München
zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

| | |
|---|---|
| Vorsitzender: | Univ.-Prof. Dr. G. Abstreiter |
| Prüfer der Dissertation: | 1. Hon.-Prof. I. Cirac, Ph.D. |
| | 2. Univ.-Prof. Dr. M. Kleber |

Die Dissertation wurde am 30.5.2007
bei der Technischen Universität München eingereicht
und durch die Fakultät für Physik am 10.10.2007 angenommen.

# Zusammenfassung

Die vorliegende Arbeit behandelt Fragestellungen im Zusammenhang mit der Quantifizierung, der Erzeugung sowie der Anwendung von Verschränktheit.

Verschränktheit hat ihre Ursache in der Beschränkung auf lokale Operationen und klassische Kommunikation. Wir untersuchen, wie sich das Konzept von Verschränktheit unter der zusätzlichen Einschränkung durch Superauswahlregeln ändert und zeigen, dass diese zu einer neuen Ressource führen. Wir charakterisieren diese Ressource und demonstrieren, wie sie verwendet werden kann, um die Einschränkung zu überwinden, ebenso wie Verschränktheit verwendet werden kann, um die Einschränkung auf lokale Operationen mittels Teleportation zu überwinden.

Anschließend betrachten wir die optimale Erzeugung von Ressourcen. Wir zeigen, wie aus verrauschten Operationen unter Zuhilfenahme perfekter passiver Operationen Squeezing bestmöglich erzeugt werden kann und diskutieren die Implikationen dieses Ergebnisses für die optimale Erzeugung von Verschränktheit.

Die Komplexität korrelierter Vielteilchensysteme rührt letztlich von der komplizierten Verschränktheitsstruktur des zugrundeliegenden multipartiten Zustands her. Wir untersuchen die Grundzustandseigenschaften von Gittern harmonischer Oszillatoren unter Verwendung von Methoden der Quanteninformation. Wir zeigen, dass für Systeme mit einer Energielücke die Korrelationen exponentiell abfallen, leiten die Beziehung zwischen Lücke und Korrelationslänge her und untersuchen das Konzept von Kritikalität, indem wir die Verbindung zwischen verschwindender Energielücke und algebraisch abfallenden Korrelationen herstellen.

In letzter Zeit sind Konzepte aus der Verschränktheitstheorie verstärkt zur Beschreibung von Vielteilchensystemen verwendet worden. Matrixproduktzustände (MPS), die eine äußerst einfache quanteninformationstheoretische Interpretation haben, können Grundzustände lokaler Hamiltonians mit hervorragender Genauigkeit approximieren. Dies wird allgemein dem Umstand zugeschrieben, dass sowohl diese Grundzustände als auch MPS eine beschränkte Blockentropie haben. Wir untersuchen den Zusammenhang zwischen der Skalierung von Blockentropien und der Approximierbarkeit durch MPS und finden insbesondere, dass auch Zustände mit beschränkter Entropie nicht stets durch MPS approximiert werden können.

Ausgehend von der quanteninformationstheoretischen Beschreibung von MPS kann eine zweidimensionale Verallgemeinerung konstruiert werden, sog. *projected entangled pair states* (PEPS). Während MPS effizient präpariert und simuliert werden können, scheint dies für PEPS nicht mehr zu gelten. Wir gehen dieser Frage nach und bestimmen sowohl für die Präparation als auch für die Simulation von PEPS deren komplexitätstheoretischen Schwierigkeitsgrad.

Schließlich führen wir Gaußsche MPS ein, eine Verallgemeinerung von MPS und PEPS auf bosonische Vielteilchensysteme, und leiten ihre Eigenschaften in Analogie zum endlichdimensionalen Fall her.

# Summary

This thesis deals with various questions concerning the quantification, the creation, and the application of quantum entanglement.

Entanglement arises due to the restriction to local operations and classical communication. We investigate how the notion of entanglement changes if additional restrictions in form of a superselection rule are imposed and show that they give rise to a new resource. We characterize this resource and demonstrate that it can be used to overcome the restrictions, very much as entanglement can overcome the restriction to local operations by teleportation.

We next turn towards the optimal generation of resources. We show how squeezing can be generated as efficiently as possible from noisy squeezing operations supplemented by noiseless passive operations, and discuss the implications of this result to the optimal generation of entanglement.

The difficulty in describing the behaviour of correlated quantum many-body systems is ultimately due to the complicated entanglement structure of multipartite states. Using quantum information techniques, we investigate the ground state properties of lattices of harmonic oscillators. We derive an exponential decay of correlations for gapped systems, compute the dependence of correlation length and gap, and investigate the notion of criticality by relating a vanishing energy gap to an algebraic decay of correlations.

Recently, ideas from entanglement theory have been applied to the description of many-body systems. Matrix Product States (MPS), which have a particularly simple interpretation from the point of quantum information, perform extremely well in approximating the ground states of local Hamiltonians. It is generally believed that this is due to the fact that both ground states and MPS obey an entropic area law. We clarify the relation between entropy scaling laws and approximability by MPS, and in particular find that an area law does not necessarily imply approximability.

Using the quantum information perspective on MPS, a natural extension to two dimensions, so-called projected entangled pair states (PEPS), can be found. While MPS can be both created and simulated efficiently, this does not seem to hold for PEPS any more. We make this rigorous by deriving the exact computational complexity of both the creation and the simulation of PEPS.

Finally, motivated by the success of MPS and PEPS in describing lattices of finite-dimensional systems, we introduce Gaussian MPS, i.e. MPS for states with a Gaussian Wigner function, and derive their properties in analogy to the finite dimensional case.

# Contents

**1 Introduction** **3**

**2 Entanglement in the presence of superselection rules** **9**
2.1 Introduction . . . . . 9
2.2 Particle number conservation as a superselection rule . . . . . 11
2.3 Characterization of pure states . . . . . 12
2.4 Mixed states in the presence of superselection rules . . . . . 21
2.5 SiV as a resource . . . . . 32
2.6 Conclusions . . . . . 38

**3 Optimal generation of squeezing and entanglement** **39**
3.1 Introduction . . . . . 39
3.2 Gaussian states and operations . . . . . 40
3.3 Single iteration case . . . . . 42
3.4 Multiple iteration case . . . . . 44
3.5 General Gaussian maps . . . . . 45
3.6 An example . . . . . 46
3.7 Optimal entanglement generation . . . . . 47

**4 Gaussian states on harmonic lattices** **51**
4.1 Introduction . . . . . 51
4.2 Quadratic Hamiltonians and their ground states . . . . . 53
4.3 Translationally invariant systems . . . . . 57
4.4 Non-critical systems . . . . . 60
4.5 Correlation length and gap . . . . . 65
4.6 Critical systems . . . . . 70

**5 Entropy and approximability by Matrix Product States** **83**
5.1 Introduction . . . . . 83
5.2 Definitions . . . . . 84
5.3 Approximability and truncation error . . . . . 86
5.4 Conclusive cases . . . . . 87
5.5 Inconclusive cases . . . . . 88

5.6 Hardness of simulating time evolution . . . . . 91
5.7 Smooth Rényi entopies . . . . . 91
5.8 Conclusions . . . . . 92

**6 The computational complexity of PEPS** **93**
6.1 Introduction . . . . . 93
6.2 PEPS and postselection . . . . . 94
6.3 The power of creating PEPS . . . . . 95
6.4 The classical complexity of PEPS . . . . . 96
6.5 PEPS and ground states . . . . . 98
6.6 The power of creating ground states . . . . . 99

**7 Gaussian Matrix Product States** **101**
7.1 Introduction . . . . . 101
7.2 Definition of Gaussian MPS . . . . . 102
7.3 Completeness of Gaussian MPS . . . . . 104
7.4 GMPS with finitely entangled bonds . . . . . 105
7.5 Correlation functions of Gaussian MPS . . . . . 106
7.6 States with rational trigonometric functions as Fourier transforms 108
7.7 Correlation length . . . . . 109
7.8 GMPS as ground states of local Hamiltonians . . . . . 110

**8 Conclusions and outlook** **111**

**Acknowledgements** **115**

**Bibliography** **117**

# Chapter 1

# Introduction

Entanglement is one of the most intriguing features of quantum physics, and it is at the heart of applications such as quantum cryptography [1, 2], teleportation [3], dense coding [4], and quantum computation [5]. On the other hand, it is also a key ingredient for a variety of phenomena observed in many-particle physics, as e.g. superconductivity [6], quantum phase transitions [7, 8], or the fractional quantum Hall effect [9]. Quantum information, starting from the observation that information is inherently physical, aims to describe quantum mechanics as a theory of information processing, and, vice versa, investigates the possibilities and limitations of information processing in systems governed by the laws of quantum mechanics. Entanglement theory is one of the principal topics in quantum information, where research is carried out in several different directions.

A central aim is to get a better understanding of entanglement, both qualitatively and quantitatively. What different types of entanglement exist, and how can they be classified? Is there a well defined way to measure how much entanglement is contained in a given state, and what is the meaning of such a number? Another direction is to explore in which way entanglement can be utilized. In which situations does it help to have entanglement, and how can it be obtained in a given scenario under certain restrictions? Finally, since entanglement underlies the complexity of correlated quantum many-body systems, can entanglement theory help to get a better understanding of the behaviour of those systems, and to develop techniques for their description and simulation?

From the perspective of information theory, entanglement is viewed as a *resource*. Generally, a resource allows to do things which are otherwise impossible. Entanglement, for instance, can be used to teleport quantum states between two separated parties which are restricted to local operations and classical communication (LOCC), and thus to overcome the restriction. On the other hand, resources only exist due to the very restriction: entanglement, for instance, arises only due to the restriction to LOCC.

A central aim of information theory is to quantify resources. Given a noisy communication line between two parties, can it be used to simulate a differ-

ent communication line or to establish perfect communication, and how much bandwidth has to be used? Correspondingly, in entanglement theory one asks whether a non-perfectly entangled state can be used to substitute a perfect one, and at which cost. It has been a most important discovery that the amount of entanglement contained in any pure bipartite state can be quantified by a single number, the entropy of entanglement [10]. Asymptotically, any two states can be converted into each other and thus into a maximally entangled state as long as the total amount of entanglement is conserved. The entanglement of mixed states however cannot be quantified in a unique way: On the one hand, one can ask how much entanglement is needed to create a certain state [11], and on the other hand, how much entanglement can be extracted from it [12]. Different from the case of pure states, these measures do not coincide, which makes the entanglement of mixed states a more involved but also more rich subject.

Entanglement is used in many different tasks such as teleportation, dense coding, or quantum cryptography. In order to generate entanglement in a given scenario, one wants to use the available resources as efficiently as possible. Various ideas have been investigated: to establish entanglement between two distant parties, relay stations—so-called quantum repeaters—might prove useful [13], and percolation ideas can be applied to create entanglement using a network of intermediate parties [14]. On the other hand, if one considers two particles which can interact directly via some—possibly noisy—evolution, protocols to create entanglement as efficiently as possible from the given interaction are being sought [15, 16].

The complexity of correlated many-particle states, as e.g. ground or thermal states of local Hamiltonians in solid state systems, has its source in the exponential dimension of the underlying Hilbert space, and thus ultimately in the complicated entanglement structure of the quantum state under investigation. In the last few years, ideas from quantum entanglement theory have been applied to describe such systems and to achieve a better understanding of the structure of the underlying states. It has been shown that the Density Matrix Renormalization Group (DMRG) [17, 18], which performs extremely well in numerically describing the ground states of one-dimensional systems, can be interpreted as a variational method over a class of states called Matrix Product States (MPS) [19]. These states have in turn a very intuitive description from the quantum information perspective, namely as maximally entangled pairs which are projected into a subspace [20]. This gave rise to improved algorithms [20–22] as well as to new classes of states, in particular Matrix Product Density Operators (MPDOs) and Projected Entangled Pair States (PEPS), which were successfully applied to numerically simulate thermal states [23] and states of two-dimensional systems [24], respectively.

This thesis deals with different aspects of entanglement and other quantum resources. In Chapter 2, we investigate how the notion of entanglement is changed if additional constraints beyond LOCC are imposed, and develop a theory of

entanglement in the presence of superselection rules, as e.g. particle number or charge conservation. We show that the additional constraints give rise to a new nonlocal resource, the *superselection induced variance* (SiV), which for pure bipartite states can be quantified by a single number in addition to the entropy of entanglement. We also investigate the mixed state case, where we derive several results on the quantification and transformation of the new resource, in analogy to the existing results on mixed state entanglement. We then illustrate that SiV is indeed a resource, as it can be used to overcome the restrictions imposed by the superselection rule. Finally, we show that the new resource also gives rise to new applications, and demonstrate how perfect data hiding protocols can be implemented.

We then turn towards bosonic systems, as given e.g. by light modes. For such systems, a new resource naturally arises: typically, particle-number preserving (passive) operations, i.e., phase shifters and beam splitters, are easy to realize, while operations which do not preserve particle number (squeezing operations) are usually hard to implement. The resource which arises from this restriction are *squeezed states*, i.e., states with a reduced uncertainty in one quadrature. In fact, squeezing is closely related to entanglement, since every entangled state with Gaussian Wigner function is also squeezed [25]. In Chapter 3, we investigate how squeezing can be generated in an optimal way from a noisy squeezing device, given free access to noiseless passive operations. We show how to maximize the squeezing obtained with one use of the device, and then consider the case of multiple iterations, where we prove a surprising result: In order to generate as much squeezing as quickly as possible, the best strategy is to maximize the amount of squeezing created in each iteration. This is indeed unexpected, as this is a global optimization problem which a priori need not be solvable locally. Finally, we investigate the corresponding question of optimal entanglement generation, where we show that for certain cases analogous results hold.

The remaining part of the thesis deals with the application of quantum information concepts, in particular entanglement theory, to the description of quantum many-body systems. We start in Chapter 4 by discussing ground state properties of lattices of harmonic oscillators subject to a quadratic Hamiltonian. These so-called Gaussian systems have a particularly simple description in terms of second moments, and therefore allow to derive analytic results for otherwise hardly tractable problems. We derive explicit expressions for the ground state and the excitation spectrum, and use these results to discuss the decay of correlations as a function of the spectral properties. For gapped systems, we show that the correlations decay exponentially; for one-dimensional systems, we explicitly compute the correlation length and find that it scales with the inverse of the energy gap, which is in accordance to what is known for finite-dimensional systems. We then consider critical systems, i.e., systems with a vanishing energy gap, and show that this implies a power law decay of the correlations, where the power depends on the dimension.

Matrix Product States (MPS) provide an efficient description of quantum states on lattices and perform extremely well in approximating ground states of local Hamiltonians. It is generally believed that the good performance of MPS in approximating ground states is due to the fact that both ground states of gapped local Hamiltonians and MPS obey a so-called area law, i.e. the entropy of any block of spins is bounded by a constant. In Chapter 5, we look more closely at the relation between the scaling of block entropies and the approximability by MPS. While an at most logarithmic scaling of the block entropy for Rényi entropies with $\alpha < 1$ (but not any more for the von Neumann entropy, $\alpha = 1$) indeed implies approximability by MPS [26], we find that conversely a faster than logarithmic increase for $\alpha > 1$, as well as a linear increase of the von Neumann entropy, implies non-approximability. For all other cases, we show that the scaling of entropies does not allow for conclusions about approximability. This includes the case of bounded von Neumann entropy, thus demonstrating that the reason for the approximability of ground states by MPS is not simply the fact that they obey an area law. We then apply the obtained results to show that MPS cannot be used to simulate the time evolution of one-dimensional quantum systems, not even for the case of translational invariant and time indepentent evolutions, starting with a translational invariant initial state.

A main reason for the performance of MPS-based numerical methods lies in the fact that expectation values of local observables can be efficiently evaluated classically on these states. Analogously, it has been shown that any MPS can be generated efficiently by a sequential "hen and egg" scheme, as e.g. given by atoms passing through a cavity [27]. Projected Entangled Pair States (PEPS) provide a natural extension of MPS to higher dimensions and have successfully been applied for numerical simulations [24]. However, there is evidence that those states can neither be created nor simulated efficiently: it has been demonstrated that thermal states of classical spin systems can be mapped to PEPS, and since finding ground states of spin glasses in two dimensions is $\mathsf{NP}$-hard, this poses lower bounds on their complexity [28]. In Chapter 6, we therefore investigate two questions: First, what is the power of creating PEPS, and second, what is the complexity of simulating them? We exactly determine the computational complexity for the two cases, which is given by the complexity classes $\mathsf{PP}$ and $\#\mathsf{P}$, respectively. Our central tool is a duality between PEPS and postselection, which allows to use well-established tools from quantum complexity theory. The result for the simulation of PEPS can be extended to the contraction of arbitary tensor networks, thus giving a quantum proof for a completely classical problem. We also find that all the complexity of PEPS is already contained in the two-dimensional case, which makes it an even more interesting subject to investigations. At the end of the chapter, we discuss the implications of these results for the approximation of ground states by PEPS and give evidence why creating ground states of gapped Hamiltonians might be an easier task than creating arbitrary PEPS.

Given the the success of MPS in the description of systems with a finite Hilbert space dimension per site, it is natural to look for extensions to continuous variables. In the last chapter, we introduce Gaussian Matrix Product States (GMPS), i.e., MPS for bosonic systems with a Gaussian phase space distribution. We prove that GMPS form a complete family, and in particular that any translational invariant state has a translational invariant GMPS representation. We then show that correlation functions of GMPS decay exponentially and that the correlation length can be directly computed from the GMPS representation, and finally that every GMPS is a ground state of a local Hamiltonian.

The results presented in this thesis have been published in [29–36].

# Chapter 2

# Entanglement in the presence of superselection rules

## 2.1 Introduction

One of the central results in quantum information theory has been the discovery that the amount of nonlocality contained in a bipartite quantum system can be quantified by a single number, the entropy of entanglement (EoE). Asymptotically, multiple copies of any two states can be converted into each other and thus into singlets provided that the total EoE is conserved [10]. On the other hand, entanglement is the key resource for some of the most interesting tasks in quantum information, as teleportation [3] and dense coding [4].

Entanglement has its origin in the restriction to those transformations which can be implemented by local operations and classical communication (LOCC) [37, 38]. In the same way, any additional restriction should lead to another nonlocal quantity and thus to new effects and applications. It has been noted by Popescu [39] that in many physical systems of interest such a restriction is given by a superselection rule (SSR). In the following, we will consider the superselection rule given by particle number or charge conservation; this is motivated, e.g., by recent quantum optical experiments on cold atomic gases. Indeed, the notion of entanglement is affected by the additional restrictions [40, 41], and new protocols arise, e.g., perfect data hiding [42] becomes possible [40]. On the other hand, it has been shown [40, 43] that the extra resource of a shared reference frame (i.e., a nonlocal state) allows to overcome the restrictions imposed by the SSR; conversely, private reference frames restrict the possible operations of an eavesdropper and can thus be employed for cryptographic tasks [44].

In this chapter, we quantify the nonlocal resource induced by particle number conservation both for pure and mixed states. We start by discussing the pure state case, where the main result is that the nonlocality contained in a bipartite pure state subject to SSR can be quantified by only one additional number,

the *superselection induced variance* (SiV): any two states can be interconverted asymptotically as long as the total EoE *and* SiV are conserved. Therefore, we discuss how the majorization criterion [45] which governs the conversion of quantum states has to be changed when SSR are present, and show that it asymptotically converges to the conservation of EoE (as it is the case without SSR) and SiV. We prove this result in detail for arbitrary states and show that it motivates the definition of two different types of standard forms for SiV which carry a linear and logarithmic amount of EoE, respectively.

While there exist pure states which carry only EoE, there are no pure states which contain solely SiV. On the other hand, it has been demonstrated [40] that there exist separable but nonlocal mixed states, i.e., states which have a separable decomposition and thus do not contain EoE, but are still nonlocal as all these decompositions violate the SSR, and therefore should contain SiV. In order to make these statements quantitative, we extend the concepts of EoE and SiV to mixed states subject to SSR. One natural way to do this is to consider the amount of pure states resources needed to create the state [11]; we show that this extension can be done in a meaningful way and that there indeed exist states which contain SiV but no EoE. The converse way is to ask whether it is possible to distill pure state resources from some mixed state [12]; we provide ways to distill both EoE and SiV, and we show that it is even possible to distill the SiV contained in separable states.

EoE is a *resource*—it allows to overcome the LOCC restrictions by teleportation. It is reasonable to assume that any restriction leads to a nonlocal quantity which in turn allows to overcome this restriction. Indeed, we give evidence that SiV can be used as a resource which allows to overcome the additional restrictions imposed by the SSR in a bipartite setting. Therefore, we will use two tasks: distinguishing locally undistinguishable quantum states and teleporting states with nonconstant local particle number [40]. We will show that not only pure states can be used as share reference frames for these tasks, but that there even exist separable states which together with one ebit of entanglement allow to perfectly teleport one qubit and thus to overcome all restrictions. Still, we find that there is a fundamental difference between EoE and SiV as a resource, as a finite amount of nonlocality does not allow to perfectly overcome the restrictions, which is due to the structure of the underlying Hilbert space.

The chapter is organized as follows. In Section 2.2, we introduce the concept of a superselection rule and show how it restricts the operations which can be implemented in a bipartite setting. In Section 2.3, we consider the conversion of pure states. We start with the conversion of single copies, which motivates the definition of SiV as a nonlocal monotone; then, we prove that asymptotically all states can be converted given that both SiV and EoE are conserved. Section 2.4 is devoted to mixed state nonlocality. First, we discuss formation of mixed states; beyond other results, we provide explicit formulas for the case of qubits. Second, we give different methods for the distillation of both EoE and SiV independently

as well as simultaneously. Finally, Section 2.5 discusses SiV as a resource; there, we quantify how well states with SiV can be used as shared reference frames which allow to overcome the new restrictions, and we demonstrate that one ebit of entanglement is still sufficient for teleportation.

## 2.2 Particle number conservation as a superselection rule

We will focus on particle number conservation as a SSR, but the results also apply to charge and other discrete quantities. In this case, the Hilbert space of the system $\mathcal{H}$ can be decomposed into a direct sum $\mathcal{H} = \bigoplus_{N=0}^{\infty} \mathcal{H}_N$ of the eigenspaces of the particle number operator $\hat{N}$, and the SSR imposes that for any operator $\mathcal{O}$, $[\mathcal{O}, \hat{N}] = 0$ must hold; thus, any operator can be written as a sum of operators $\mathcal{O}_N$ which have support on $\mathcal{H}_N$ only, $\mathcal{O} = \bigoplus_{N=0}^{\infty} \mathcal{O}_N$, and thus

$$\mathcal{O} = \sum_N P_N \mathcal{O} P_N \ , \tag{2.1}$$

where $P_N$ projects onto $\mathcal{H}_N$. As the same restriction holds for the admissible density operators, all states can be converted into each other, and no interesting new effects can be found.

Therefore, we consider SSR in a bipartite setting. Then, we have local particle number operators $\hat{N}_A$ and $\hat{N}_B$, and the total particle number operator is given by

$$\hat{N}_{AB} = \hat{N}_A \otimes \mathbb{1}_B + \mathbb{1}_A \otimes \hat{N}_B \ . \tag{2.2}$$

While the admissible states have to commute with the global particle number operator $\hat{N}_{AB}$, the local operations have to commute with the local particle number operators $\hat{N}_A$ and $\hat{N}_B$. This restriction is stronger than the one given by the bipartite setting alone and should therefore lead to a new nonlocal resource. More precisely, the operations on subspaces with fixed total particle number $N = N_{AB}$ are given by

$$\mathcal{O}_N^{AB} = \bigoplus_{N_A+N_B=N} (\mathcal{O}_{N_A}^A \otimes \mathbb{1}_{N_B}^B) \tag{2.3}$$

(and vice versa)—in addition to the restriction to products $\mathcal{O}^A \otimes \mathbb{1}$ imposed by the bipartite setting, a direct sum structure arises from the SSR. This product vs. sum structure will be present throughout the chapter, and is the reason for some fundamental differences between EoE (arising for the product structure) and SiV (arising from the direct sum).

The restriction to block-diagonal operations, Eq. (2.1), can be relaxed by adding ancilla modes with $m_0$ particles, performing a block-diagonal unitary $U$, and measuring resp. tracing out the ancillas. Then, the admissible (POVM/Kraus)

operators are given by $\mathcal{O} = P_m^{\text{anc}} U P_{m_0}^{\text{anc}}$; by applying (2.1) to $U$, this leads to $\mathcal{O} = \sum_N P_{N+\Delta} \mathcal{O} P_N$ (resp. $[\hat{N}, \mathcal{O}] = \Delta\mathcal{O}$, $\Delta$ might differ for each $\mathcal{O}$): $\mathcal{O}$ can shift the particle number by some $\Delta$. (Note that $\mathcal{O}^\dagger\mathcal{O}$ remains block-diagonal). As most of the results are only affected marginally by including ancillas, we will usually neglect them and just briefly comment on their effect as appropriate.

At the end of this section, let us introduce a few notational conventions. Logarithms are taken to the basis 2. A ket $|N\rangle$ denotes a state with $N$ particles. We will use this notation even if the underlying eigenspace is degenerate, unless the nonlocal properties under consideration depend on this degeneracy.

The restrictions imposed by the SSR on the allowed operations can be easily overcome by defining a new computational basis $|\hat{0}\rangle \equiv |01\rangle$, $|\hat{1}\rangle \equiv |10\rangle$ in which all states have the same particle number [40]. This motivates the definition of *two different* types of maximally entangled two-qubit states,

$$|\text{V-EPR}\rangle = |0\rangle_A|1\rangle_B + |1\rangle_A|0\rangle_B$$

(a "variance-EPR", as there is some variance in the local particle number), and

$$|\text{E-EPR}\rangle = |01\rangle_A|10\rangle_B + |10\rangle_A|01\rangle_B \equiv |\hat{0}\rangle_A|\hat{1}\rangle_B + |\hat{1}\rangle_A|\hat{0}\rangle_B$$

(an "entanglement-EPR", which is defined within an unrestricted subspace and only carries entanglement). The very difference between these two states will be a central issue in what follows.

## 2.3 Characterization of pure states

In this section, we characterize pure states in a bipartite setting, i.e., we determine the possible conversions by LOCC and thus quantify the nonlocality contained in a bipartite state. Without superselection rules, the majorization criterion determines whether the conversion between two bipartite pure states is possible [45, 46]. The conversion of multiple copies is governed by a much simpler criterion: it has been shown [10] that multiple copies of any two states can be interconverted reversibly. The conversion ratio is determined by only one quantity which fully characterizes the nonlocal properties of a bipartite state, the *entropy of entanglement* (EoE).

As we have seen in the preceding section, in addition to the tensor product structure induced by the bipartite setting the operators have to obey a direct sum structure. In this section, we show that these two structures lead to two complementary resources: while the tensor product again induces the majorization criterion and (asymptotically) EoE as a nonlocal resource, the direct sum gives rise to additional restrictions on the conversions of states and in turn leads to an own nonlocal resource.

### 2.3.1 The single copy case

Let us consider the following problem: given pure bipartite states $\phi$ and $\psi$, is it possible to convert $\phi$ to $\psi$ by LOCC? This task can be generalized naturally to a set of outcomes $\{(p_i, \psi_i)\}$, where each outcome $\psi_i$ is obtained with probability $p_i$.

Let us first see how this can be solved without SSR [45]. Therefore, let $\boldsymbol{\lambda} = (\lambda_k)$ and $\boldsymbol{\mu}^i = (\mu_k^i)$ be the Schmidt coefficients of $\phi$ and $\psi_i$, respectively, which completely characterize the states up to local unitaries. Without loss of generality, the Schmidt vectors may be taken decreasing ($\lambda_k \geq \lambda_{k+1}$) and of equal dimension (by appending zeros). Following [45], an LOCC strategy for the conversion

$$\phi \longrightarrow \{(p_i, \psi_i)\}$$

exists if and only if

$$\boldsymbol{\lambda} \prec \sum_i p_i \boldsymbol{\mu}^i \ .$$

Here, for two ordered vectors $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$, we say that $\boldsymbol{\lambda}$ *is majorized by* $\boldsymbol{\mu}$, $\boldsymbol{\lambda} \prec \boldsymbol{\mu}$, if $\sum_{k=1}^d \lambda_k \leq \sum_{k=1}^d \mu_k$ for all $1 \leq d < \dim \boldsymbol{\lambda}$, where equality holds for $d = \dim \boldsymbol{\lambda}$.

As an example, consider the states

$$\begin{aligned} |\phi\rangle &= \sqrt{\tfrac{1}{2}}|0\rangle_A|1\rangle_B + \sqrt{\tfrac{1}{2}}|1\rangle_A|0\rangle_B \ \text{ and} \\ |\psi\rangle &= \sqrt{\tfrac{1}{3}}\,|0\rangle_A|1\rangle_B + \sqrt{\tfrac{2}{3}}\,|1\rangle_A|0\rangle_B \ , \end{aligned}$$

which have the ordered Schmidt vectors $\boldsymbol{\lambda} = (1/2, 1/2)$ and $\boldsymbol{\mu} = (2/3, 1/3)$, respectively. Since $\boldsymbol{\lambda} \prec \boldsymbol{\mu}$, it is possible to convert $\phi \to \psi$; for instance, Alice might start with the POVM measurement given by $M_1 = \sqrt{1/3}|0\rangle\langle 0| + \sqrt{2/3}|1\rangle\langle 1|$ and $M_2 = \sqrt{1/3}|0\rangle\langle 0| + \sqrt{2/3}|1\rangle\langle 1|$ which yields the two states

$$\begin{aligned} |\psi_1\rangle &= \sqrt{\tfrac{1}{3}}\,|0\rangle_A|1\rangle_B + \sqrt{\tfrac{2}{3}}\,|1\rangle_A|0\rangle_B \ \text{ and} \\ |\psi_2\rangle &= \sqrt{\tfrac{2}{3}}\,|0\rangle_A|1\rangle_B + \sqrt{\tfrac{1}{3}}\,|1\rangle_A|0\rangle_B \ . \end{aligned}$$

with equal probabilities: $\psi_1$ is already equal to $\psi$, and $\psi_2$ can be converted to $\psi$ by a bilateral NOT operation.

Let us now see what is different when SSR apply: while the POVM measurement $\{M_1, M_2\}$ is compatible with the superselection rule, the local application of NOT operations is not; indeed, it is not possible at all to carry out $\phi \to \psi$ deterministically in the presence of SSR. In order to see this, define block-diagonal POVM operators $M_i = \bigoplus_n M_n^i$ on one local system. Then, the completeness relation $\sum M_i^\dagger M_i = \mathbb{1}$ yields $\sum_i M_n^{i\dagger} M_n^i = \mathbb{1}$ for all $n$. Therefore, any POVM operator is simply a direct sum of POVM operators acting within the subspaces with constant local particle number, i.e., the usual conditions for convertibility

have to hold for each subspace separately. Particularly, this implies that for pure states the *average weight* of each subspace with constant local particle number *cannot be changed* by local operations.

The impossibility to change the average weight of a subspace with fixed local particle number can even be proven at a much more fundamental level. Take multiple copies of some state $|\phi\rangle$ with nonconstant local particle number, and assume there is a way for Alice to change her local particle number distribution on average. As the total particle number is constant, this implies that the average particle number distribution of Bob's system is changed the other way round. Therewith, Alice can change Bob's density matrix remotely which would allow for supraluminal communication and therefore has to be ruled out. Classical communication between Alice and Bob, on the other hand, will increase Bob's knowledge of the *actual* particle number distribution, but it cannot influence the *average* distribution obtained.

In order to formulate this result precisely, note that any bipartite state $\phi \in \mathcal{H}_N$ subject to SSR can be written as $\phi = \phi^0 \oplus \cdots \oplus \phi^N$ with $\phi^n \in \mathcal{H}_n^A \otimes \mathcal{H}_{N-n}^B$, i.e., as a direct sum of unnormalized pure states with constant local particle number. Call the (ordered) *unnormalized* Schmidt coefficients of $\phi^n$ $\boldsymbol{\lambda}^n$. Then, $\phi$ is characterized up to local (SSR-compatible) unitaries by its *SSR-ordered Schmidt vector* $\boldsymbol{\lambda} = (\boldsymbol{\lambda}^0, \ldots, \boldsymbol{\lambda}^N)$.

**Theorem 2.1.** *Let $\phi$, $\psi_i$ be pure states and $\boldsymbol{\lambda}$, $\boldsymbol{\mu}_i$ their SSR-ordered Schmidt vectors. Then,*[1]

$$\phi \xrightarrow{\text{SSR}} \{(p_i, \psi_i)\} \tag{2.4}$$

*(i.e., there exists a SSR-compatible conversion strategy) if and only if*

$$\boldsymbol{\lambda}^n \prec \sum_i p_i \boldsymbol{\mu}_i^n \quad \forall\, n = 0, \ldots, N. \tag{2.5}$$

In order to see the connection to the conversions within the subspaces, let us re-express (2.5) by normalizing the SSR-ordered Schmidt vectors,

$$\hat{\boldsymbol{\lambda}}^n \prec \sum_i \underbrace{p_i \frac{\|\boldsymbol{\mu}_i\|}{\|\boldsymbol{\lambda}\|}}_{=:p_i'} \hat{\boldsymbol{\mu}}_i^n \quad \forall\, n = 0, \ldots, N\,,$$

where in the following a hat $\hat{\cdot}$ denotes the normalized vector. According to the usual majorization result, this holds iff we can convert

$$\hat{\phi}^n \longrightarrow \{p_i', \hat{\psi}_i^n\} \quad \forall n = 0, \ldots, N\ . \tag{2.6}$$

[1] If one includes ancillas, the Theorem only holds up to shifts in the local particle number performed on the output states, i.e., Eq. (2.5) has to be replaced by $\exists \nu_i \in \mathbb{N}\ \forall n = 0, \ldots, N$ : $\boldsymbol{\lambda}^n \prec \sum_i p_i \boldsymbol{\mu}_i^{n+\nu_i}$. Formally, up to local unitaries each state is then described by an *equivalence class* of SSR-ordered Schmidt vectors which are equal up to a shift in the particle number, and there have to exist vectors in these equivalence classes which satisfy Eq. (2.5). The proof transfers directly if one replaces $\hat{\psi}^n$ by $\hat{\psi}^{n+\nu_i}$.

Here, $\phi = \phi^0 \oplus \cdots \oplus \phi^N$ and $\psi_i = \psi_i^0 \oplus \cdots \oplus \psi_i^N$.

**Proof.** Exactly as without SSR, the most general strategy consists of Alice performing a generalized measurement and communicating the result to Bob, who then applies a unitary operation depending on the measurement outcome; the proof [47] can be directly transferred.

We show (2.4)⇔(2.6). The proof can be restricted to the case where each conversion $\phi \to (p, \psi)$ in (2.4) resp. (2.6) can be accomplished by a single POVM operator $M$, i.e., $M\phi = \sqrt{p}\psi$—otherwise, we can split $\phi \to (p, \psi)$ into $\phi \to (p_k, \psi)$, $\sum p_k = p$, where each conversion is the result of *one* of the POVM operators. This can be done as well for the system of conversions (2.6), where we have to split all subspaces simultaneously (this can be always done by additionally splitting single POVM operators into copies of itself).

First, assume that (2.4) holds. Then there exist POVM operators $M_i = \bigoplus_n M_i^n$ on Alices side for which $M_i\phi \cong_B \sqrt{p_i}\psi$ (i.e., up to a unitary on Bob's side). Decomposing this into the subspaces in the direct sum, one obtains $M_i^n\phi^n \cong_B \sqrt{p_i}\psi_i^n$ and thus

$$M_i^n \hat{\phi}^n \cong_B \underbrace{\sqrt{p_i \frac{\langle\psi_i|\psi_i\rangle}{\langle\phi|\phi\rangle}}}_{\equiv\sqrt{p_i'}} \hat{\psi}_i^n$$

for all $N$, i.e., the $M_i^n$ accomplish the set of conversions given by Eq. (2.6). Especially, as

$$\mathbb{1} = \sum_i \left(\bigoplus_n M_i^n\right)^\dagger \left(\bigoplus_n M_i^n\right) = \bigoplus_n \left(\sum_i M_i^{n\dagger} M_i^n\right) ,$$

the $M_i^n$ obey the completeness relation for POVM operators. As all arguments hold in both directions, this completes the proof. □

### 2.3.2 Variance as a nonlocal monotone

Let us now formulate an asymptotic version of the previous theorem. It is known that without SSR for a large number of copies the majorization criterion converges to the entropic criterion, i.e., the conservation of the total EoE [10]. With SSR, the probability distribution associated to the variation of the local particle number, $p_n = \sum_i p_i^n$, has to be conserved as well. Asymptotically, this distribution converges to a Gaussian which is completely characterized by its mean (which can be shifted using ancillas) and its variance. Therefore we define

**Definition 2.1.** *For a bipartite pure state $\phi$ shared by A and B, define the superselection induced variance (SiV)*

$$V(\phi) := 4\left[\langle\phi|\hat{N}_A^2|\phi\rangle - \langle\phi|\hat{N}_A|\phi\rangle^2\right] ,$$

*where $N_A$ is the particle number operator for Alice.*[2] *(One could equally well take $\hat{N}_B$, as $\hat{N}_A + \hat{N}_B = N = \text{const.}$)*

Let us now show that SiV is really an entanglement monotone [48] when SSR are present, namely that it cannot be increased on average by SSR-LOCC and vanishes on separable states. (On the contrary, note that $V(\phi) = 0$ does *not* imply that $\phi$ is separable—this is due to the fact that there exist two different nonlocal quantities when SSR are present.) Moreover, SiV is symmetric under interchange of $A$ and $B$ and additive: given two subsystems 1 and 2 shared by $A$ and $B$, $V(\phi_1 \otimes \phi_2) = V(\phi_1) + V(\phi_2)$, as can be readily seen by applying Eq. (2.2) to the two subsystems 1 and 2, $\hat{N}_{A1\,A2} = \hat{N}_{A1} \otimes \mathbb{1}_{A2} + \mathbb{1}_{A1} \otimes \hat{N}_{A2}$.

To show the monotonicity of SiV under SSR-LOCC, consider a POVM measurement $\{M_i^A\}$ on Alice's side. Then, the average SiV after the application of $\{M_i^A\}$ is given by

$$\bar{V}_M(\phi) = \sum_i \langle\phi|M_i^{A\dagger}\hat{N}_A^2 M_i^A|\phi\rangle - \sum_i \frac{\langle\phi|M_i^{A\dagger}\hat{N}_A M_i^A|\phi\rangle^2}{\langle\phi|M_i^{A\dagger}M_i^A|\phi\rangle} \ .$$

The first part reduces to $\langle\phi|\hat{N}_A^2|\phi\rangle$ (using $[\hat{N}_A, M_i^A] = 0$ and $\sum_i M_i^{A\dagger}M_i^A = \mathbb{1}$), while for the second part

$$\begin{aligned} \sum_i \frac{\langle\phi|M_i^{A\dagger}\hat{N}_A M_i^A|\phi\rangle^2}{\langle\phi|M_i^{A\dagger}M_i^A|\phi\rangle} &\overset{(*)}{\geq} \left(\sum_i \langle\phi|M_i^{A\dagger}\hat{N}_A M_i^A|\phi\rangle\right)^2 \\ &= \langle\phi|\hat{N}_A|\phi\rangle^2 \ . \end{aligned}$$

Here, $(*)$ has been derived using the Cauchy-Schwarz inequality

$$\left(\sum_i y_i\right)^2 = \left(\sum_i \sqrt{p_i}\frac{y_i}{\sqrt{p_i}}\right)^2 \leq \sum_i \frac{y_i^2}{p_i}\sum_{i'} p_{i'} \ . \tag{2.7}$$

Ancillas leave the result unaffected, as the extra contributions in $\bar{V}_M(\phi)$ originating from $[\hat{N}, \mathcal{O}] = \nu\mathcal{O}$ cancel out.

### 2.3.3 Reversible conversion of multiple copies

The introduction of SiV as a nonlocal monotone was motivated by the conversion of multiple copies, as it characterizes the joint particle number distribution. In the following, we will show that asymptotically SiV and EoE quantify the two complementary resources which completely characterize biparitite states up to SSR-LOCC.

[2] The factor 4 in the definition normalizes the SiV: $V(|\text{V-EPR}\rangle) = 1$.

**Theorem 2.2.** *In the presence of SSR, there exists an asymtotically reversible conversion*

$$|\phi\rangle^{\otimes N} \otimes |\hat{0}\rangle^{\otimes E(\phi)N} \longleftrightarrow \sum_n c_n |n\rangle|N-n\rangle \otimes |\text{E-EPR}\rangle^{\otimes E(\phi)N},$$

*where the coefficients $c_n$ are distributed Gaussian with SiV $NV(|\phi\rangle)$.*

Note that on the left hand side we have added ancilla states in the unrestricted "hat"–basis (cf. Sec. 2.2). The conversion transfers the entanglement contained in $|\phi\rangle^{\otimes N}$ to this second register as "accessible" entanglement in the form of |E-EPR⟩s, while the SiV stays in the first register.

**Proof.** First, we restrict ourself to the case of qubits, where $|\phi\rangle = \sqrt{p_0}|0\rangle|1\rangle + \sqrt{p_1}|1\rangle|0\rangle$. We will generalize the result in two steps: in a first step, we consider qu-$d$-its, where the local basis is $\{|0\rangle, \ldots, |d-1\rangle\}$, while in a second step we allow for arbitrary bipartite states, i.e., the local bases might contain several states with the same particle number.

For the beginning, let us only look at the first register. Taking $N$ copies of $|\phi\rangle$, we have

$$|\phi\rangle^{\otimes N} = \sum_{\mathbf{x}} \sqrt{p_0^{n_0} p_1^{n_1}} |\mathbf{x}\rangle|\neg\mathbf{x}\rangle \ ,$$

where the sum is taken over all possible $N$-bit strings $\mathbf{x}$. Here, $n_0$ and $n_1$ are the numbers of zeroes and ones in $\mathbf{x}$, respectively, and $\neg\mathbf{x}$ denotes the bitwise NOT of $\mathbf{x}$. This state can be grouped naturally as

$$|\phi\rangle^{\otimes N} = \sum_{n_0} \sqrt{p_0^{n_0} p_1^{N-n_0} \binom{N}{n_0}} |\chi_{N-n_0,n_0}\rangle \ , \tag{2.8}$$

where the state $|\chi_{N-n_0,n_0}\rangle \in \mathcal{H}^A_{N-n_0} \otimes \mathcal{H}^B_{n_0}$ is a maximally entangled state with Schmidt number $\binom{N}{n_0}$.

In the following, we show how to transfer the entanglement of $|\phi\rangle^{\otimes N}$ to the second register. Therefore, we have to break the tensor product structure $|\phi\rangle^{\otimes N}$ of the first register and create a new tensor product structure by properly transferring the entanglement to the second register. To this end, let us introduce the concept of typical subspaces [49]. An $\epsilon$-typical subspace of our Hilbert space is defined as $\mathcal{H}_\epsilon = \bigoplus_{n_0 \in \mathcal{S}_\epsilon} \mathcal{H}^A_{N-n_0} \otimes \mathcal{H}^B_{n_0}$, where the $\epsilon$-typical $n_0$ are those lying in $\mathcal{S}_\epsilon = \{n_0 : |n_0/N - p_0| < \epsilon\}$. It can be shown [49, 50] that projecting $|\phi\rangle^{\otimes N}$ onto $\mathcal{H}_\epsilon$ gives an error which vanishes for $N \to \infty$ such that we can restrict the sum in (2.8) to $n_0 \in \mathcal{S}_\epsilon$. Then,

$$\binom{N}{n_0} \geq \frac{1}{(N+1)^2} 2^{NH\left(\frac{n_0}{N}\right)} \geq 2^{N[H(p_0) - K\epsilon]} \tag{2.9}$$

with some $K > 0$ holds for all $n_0 \in \mathcal{S}_\epsilon$ ($\epsilon \ll 1$ and $N \gg 1$) [49]; here $H(p) = H(p, 1-p)$ is the Shannon entropy of the probability distribution $(p, 1-p)$. According to Theorem 2.1, we can transform

$$|\chi_{N-n_0,n_0}\rangle \to \frac{1}{\sqrt{E}} \sum_{i=1}^{E} |i_{N-n_0}\rangle_A |i'_{n_0}\rangle_B \; ; \; E = H(p_0) - K\epsilon$$

coherently in all subspaces in the restricted sum, where $|i_n\rangle$ are orthogonal states with $n$ particles. Then by local maps $|i_n\rangle|\hat{0}\rangle \mapsto |n\rangle|\hat{i}\rangle$, where $|\hat{i}\rangle$ are orthogonal and $|n\rangle = |1\cdots 10\cdots 0\rangle$, the entanglement $H(p_0) - K\epsilon$ can be transferred to the second register which gives

$$\sum_{n_0 \in \mathcal{S}_\epsilon} c_{n_0} |N-n_0\rangle|n_0\rangle \otimes \Big[|01\rangle|10\rangle + |10\rangle|01\rangle\Big]^{\otimes N[H(p_0)-K\epsilon]} , \tag{2.10}$$

where

$$c_{n_0} = \sqrt{p_0^{n_0} p_1^{N-n_0} \binom{N}{n_0}} .$$

The sum can be extended to all $n_0$ with high fidelity, and the $|c_{n_0}|^2$ approach a Gaussian distribution with variance $Np_0(1-p_0) = V(\phi)/4$. This is the only parameter characterizing the state (2.10), since the mean can be shifted by locally adding ancillas. As $H(p_0)$ is just $E(\phi)$, this completes the distillation direction of the proof.

The dilution direction can be proven using the converse of (2.9),

$$\binom{N}{n_0} \leq 2^{NH\left(\frac{n_0}{N}\right)} \leq 2^{N[H(p_0)+K\epsilon]} ,$$

in an $\epsilon$-typical subspace. Starting from

$$\sum_{n_0 \in \mathcal{S}_\epsilon} c_{n_0} |N-n_0\rangle|n_0\rangle \otimes \Big[|01\rangle|10\rangle + |10\rangle|01\rangle\Big]^{\otimes N[H(p_0)+K\epsilon]} ,$$

we can transfer the entanglement to the first register and then (again by Theorem 2.1) reduce the Schmidt number of each subspace to $\binom{N}{n_0}$, obtaining the projection of $|\phi\rangle^{\otimes N}$ onto the $\epsilon$-typical subspace, so that the dilution works as well. This completes the proof for qubits.

In a first step, we generalize the proof from qubits to $(I+1)$-level systems,

$$|\phi\rangle = \sum_{i=0}^{I} \sqrt{p_i} |i\rangle|I-i\rangle . \tag{2.11}$$

(Note that the coefficients can be made positive by local operations.) Again, for $N$ copies of $|\phi\rangle$, an $\epsilon$-typical subspace can be defined by restricting the number

$n_i$ of occurences of the state $|i\rangle|I-i\rangle$ in the product by $|n_i/N - p_i| < \epsilon$ for all $i$. Projecting the state onto an $\epsilon$-typical subspace again only yields a vanishingly small error, and the Schmidt number of the states with fixed numbers $(n_0, \ldots, n_I)$ is given by the multinomial coefficient $\binom{N}{n_0 \cdots n_I}$ and obeys the bounds [49]

$$2^{N[E(\phi)-K\epsilon]} \leq \binom{N}{n_0 \cdots n_I} \leq 2^{N[E(\phi)+K\epsilon]} .$$

Thus, it is possible to extract the entanglement $E(\phi)$ reversibly. Yet, there are several possible configurations $(n_0, \ldots, n_I)$ which yield the same local particle number $n = \sum_i i n_i$ such that there is still some entanglement left in each subspace. But as for $N$ copies of $|\phi\rangle$ the number of these configurations is bounded by $N^I$, this entanglement is logarithmic in $N$ and can be removed reversibly. Therefore, we can reversibly transform $|\phi\rangle^{\otimes N} \otimes |\hat{0}\rangle^{\otimes NE(\phi)}$ into

$$\sum c_n |n\rangle|IN-n\rangle \otimes \Big[|01\rangle|10\rangle + |10\rangle|01\rangle\Big]^{\otimes NE(\phi)} , \tag{2.12}$$

where the $c_n$ are given by the sum over all coefficients for which the particle number on Alice's side is $n$,

$$c_n = \sqrt{\sum_{\substack{\sum_i i n_i = n \\ \sum_i n_i = N}} p_0^{n_0} \cdots p_I^{n_I} \binom{N}{n_0 \cdots n_I}} .$$

It remains to be shown that the $|c_n|^2$ approach a Gaussian distribution. As long as all $p_i \neq 0$, this can be shown by expanding each $n_i$ within the typical subspace as $n_i = N(p_i + \delta_i)$ with $\delta_i < \epsilon$. This will work fine whenever $Np_i \gg 1$ and $\epsilon \ll p_i$ for all $i$. Yet, this condition cannot be satisfied if $p_i = 0$ for some $i$. This might (but need not!) lead to a periodic gap in the distribution of the $|c_n|^2$, e.g., for $I = 2$, $p_0 = p_2 = 1/2$. In that case, $|c_n|^2 = 0$ for all odd $n$.

In principle, such a gap has to be considered as a third nonlocal characteristic of a bipartite state. Still, it can be removed easily. In the example given above the gap is readily removed by adding *one* $|\text{V-EPR}\rangle$, such that the fraction of $|\text{V-EPR}\rangle$ per copies of $|\phi\rangle$ vanishes. By further adding an $|\text{E-EPR}\rangle$ (those are obtained anyway in the distillation) the $|\text{V-EPR}\rangle$ can be re-obtained—it therefore merely acts as a catalyst, "freeing" the subspaces with odd particle number.

The generalization to an arbitrary state is straightforward. Take

$$|\phi\rangle = \sum_{i=0}^{I} \sqrt{p_i} |\psi_{i,I-i}\rangle , \tag{2.13}$$

where $|\psi_{i,I-i}\rangle \in \mathcal{H}_i^A \otimes \mathcal{H}_{I-i}^B$ might themselves be entangled states. Applying the concept of typical subspaces to Eq. (2.13), we find that the number of occurences

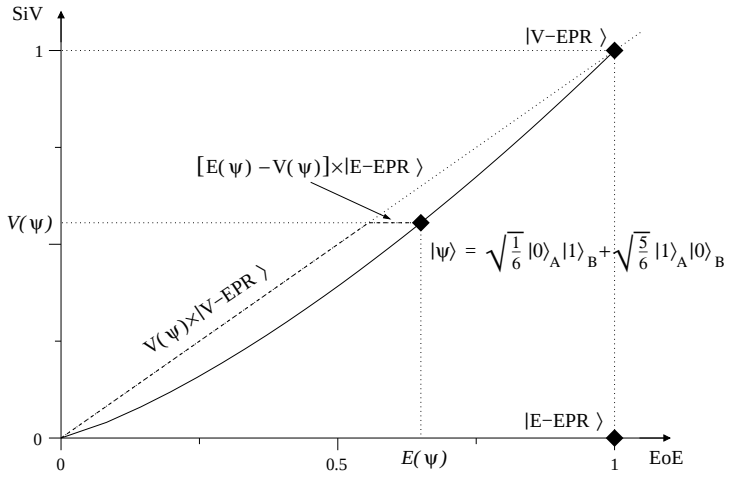


**Figure 2.1:** Characterization of pure qubit states in an $E$-$V$ diagram. All the states reside on the solid curve; asymptotically, any state can be converted into $V(\psi)$ copies of a $|\text{V-EPR}\rangle$ and $E(\psi) - V(\psi)$ of an $|\text{E-EPR}\rangle$.

of each $|\psi_{i,I-i}\rangle$ in the typical subspace is bounded by $(p_i \pm \epsilon)N$, and thus the entanglement $E(\psi_{i,I-i})$ contained in these states—which is already "accessible entanglement"—can be extracted reversibly. The remaining state is of the type of Eq. (2.11) (with the same coefficients $p_i$) and thus can be transformed reversibly into a Gaussian distributed state with width $NV(\phi)$ and $NH(p_0, \ldots, p_I)$ ebits of entanglement. It can be checked easily that the total number of Bell pairs is $NE(\phi)$. □

For qubits, Theorem 2.2 can be re-expressed.

**Corollary 2.3.** *For bipartite qubit states* $|\phi\rangle$*,*

$$|\phi\rangle \longleftrightarrow |\text{E-EPR}\rangle^{\otimes[E(\phi)-V(\phi)]}|\text{V-EPR}\rangle^{\otimes V(\phi)}$$

*in the asymptotic limit.*

This can be shown by applying Theorem 2.2 twice, together with $E(\phi) \geq V(\phi)$ (which only holds for qubits).

Fig. 2.1 illustrates this characterization of states in the $E$-$V$ diagram. Fig. 2.2 shows the $E$-$V$ diagram for qutrits, which is considerably more complex. The bounds are given by the states with highest variance $\alpha|0\rangle_A|2\rangle_B + \beta|2\rangle_A|0\rangle_B$ and the states with highest entanglement $\alpha|0\rangle_A|2\rangle_B + \beta|1\rangle_A|1\rangle_A + \alpha|0\rangle_A|2\rangle_B$. A decomposition as in the corollary is still possible if one replaces the $|\text{V-EPR}\rangle$ by $|0\rangle_A|2\rangle_B + |2\rangle_A|0\rangle_B$ which has maximal variance.

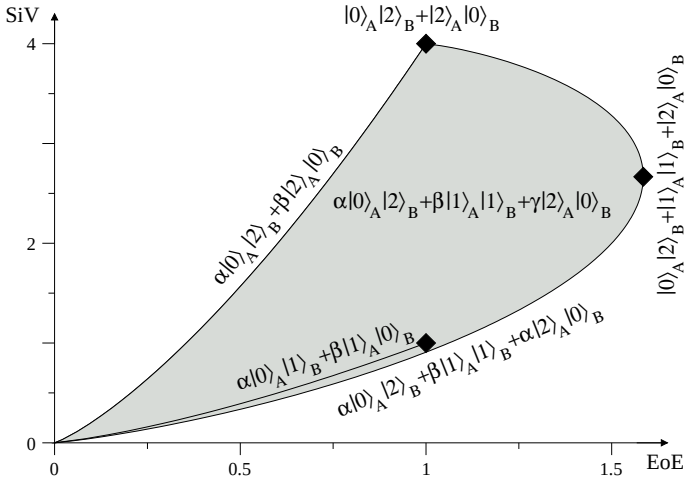


**Figure 2.2:** *E-V* diagram for qutrits, where the boundary states and the extremal states are given. The possible states reside in the gray area, the solid line within this area is the subset realizable by qubits.

## 2.4 Mixed states in the presence of superselection rules

### 2.4.1 Introduction

In the following section, we consider mixed states. We will show how the concepts of EoE and SiV as two complementary resources can be extended to mixed states, and discuss the connection with normal (SSR-free) entanglement measures.

Let us start by introducing a particularly interesting mixed state,

$$\rho_{\text{sep}} = \frac{1}{4}\begin{pmatrix} 1 & & & \\ & 1 & 1 & \\ & 1 & 1 & \\ & & & 1 \end{pmatrix} \tag{2.14}$$

in the basis $\{|0\rangle_A|0\rangle_B, |0\rangle_A|1\rangle_B, |1\rangle_A|0\rangle_B, |1\rangle_A|1\rangle_B\}$. This state has first been considered in [40], where it was shown that it is separable but nonlocal. Namely, it can be obtained by mixing $(|0\rangle_A+\omega|1\rangle_A)(|0\rangle_B+\omega|1\rangle_B)$ for $\omega \in \{1, i, -1, -i\}$ with equal probabilities, and therefore does not contain EoE. On the other hand, it is easy to see that there is no decomposition of $\rho_{\text{sep}}$ which is separable *and* compatible with the superselection rule, i.e., it cannot be created locally. Clearly, this can not happen with pure states.

Considering the results of the preceding section, it is natural to assume that $\rho_{\text{sep}}$ contains SiV but no EoE. In order to give quantitative meaning to such statements, we discuss two genuine extensions of nonlocal quantities to mixed states, defined by the asymptotic amount of pure state resources which are needed to create them and which can be extracted again.

### 2.4.2 Formation of mixed states

Let us start with the creation of mixed state in the presence of SSR. Similar to the normal case [11], we define:

**Definition 2.2.** *The entanglement of formation and the variance of formation in the presence of superselection rules are defined as*

$$E_F^{\mathrm{SSR}}(\rho) = \min_{\{p_i,\psi_i\}} \sum_i p_i E(\psi_i)$$

*and*

$$V_F^{\mathrm{SSR}}(\rho) = \min_{\{p_i,\psi_i\}} \sum_i p_i V(\psi_i) \ ,$$

*respectively. The minimum is taken over all possible decompositions of $\rho$, where the $\psi_i$ have to obey the SSR (i.e., they all have constant particle number).*

*The entanglement cost [51] and the variance cost in the presence of superselection rules are accordingly defined as the regularized versions of $E_F^{\mathrm{SSR}}$ and $V_F^{\mathrm{SSR}}$,*

$$E_c^{\mathrm{SSR}}(\rho) = \lim_{N\to\infty} \frac{E_F^{\mathrm{SSR}}(\rho^{\otimes N})}{N}$$

*and*

$$V_c^{\mathrm{SSR}}(\rho) = \lim_{N\to\infty} \frac{E_F^{\mathrm{SSR}}(\rho^{\otimes N})}{N} \ .$$

These definitions make sense, as they quantify the nonlocal resources we need at least to prepare the state $\rho$ with SSR [51].

As shown at the beginning of the section there exist states which do not contain any entanglement yet are nonlocal, as $\rho_{\mathrm{sep}}$ [Eq. (2.14)]. One easily finds that $E_F^{\mathrm{SSR}}(\rho_{\mathrm{sep}}) = 1/2$, $V_F^{\mathrm{SSR}}(\rho_{\mathrm{sep}}) = 1/2$, as each of the subblocks in $\rho_{\mathrm{sep}}$ has to be created separately. On the other hand, it seems reasonable to assume that $\rho_{\mathrm{sep}}$ can be prepared asymptotically without using entanglement. In the following, we prove an even stronger result: asymptotically, the entanglement needed to create any state $\rho$ is just the entanglement needed without SSR.

**Theorem 2.4.** *For any $\rho$ with bounded maximal particle number,*

$$E_c^{\mathrm{SSR}}(\rho) = E_c(\rho) \ ,$$

*i.e., the entanglement cost with SSR is the entanglement cost without SSR.*

**Proof.** Consider a mixed state $\sigma$ compatible with the SSR and let $\sum_i p_i |\psi_i\rangle\langle\psi_i| = \sigma$ be the optimal decomposition without SSR, i.e., $E_F(\rho) = \sum_i p_i E(\psi_i)$. Clearly, this decomposition need not obey the SSR, but we can use it to constuct a compatible decomposition with vanishing overhead. From (2.1), $\sigma = \sum_{n=0}^{N} P_n \sigma P_n$,

where $P_n$ is the projector onto the subspace with *totally* $n$ particles and $N$ the maximum total particle number in $\sigma$; therefore,

$$\sigma = \sum_{n,i} p_i p_{i,n} \frac{P_n|\psi_i\rangle\langle\psi_i|P_n}{p_{i,n}}$$

with $p_{i,n} = \langle\psi_i|P_n|\psi_i\rangle$ is a decomposition of $\sigma$ which is compatible with the SSR. For any $|\psi\rangle$ with at most $N$ particles, it holds that the measurement of the total particle number creates at most $\log(N+1)$ entanglement on average,

$$\sum_n \langle\psi|P_n|\psi\rangle E\left(\frac{P_n|\psi\rangle}{\sqrt{\langle\psi|P_n|\psi\rangle}}\right) \leq E(|\psi\rangle) + \log(N+1) \tag{2.15}$$

and with $\sigma = \rho^{\otimes M}$, the claim follows.

In order to see why (2.15) holds, we need the inequality $S(p_i\rho_i) \leq p_i S(\rho_i) + H(p_i)$ (see, e.g., [50] for a proof), and that $P_n = \sum_{i=0}^n P_i^A \otimes P_{n-i}^B$. Toghether, this gives the estimate

$$\begin{aligned} E\left(\frac{P_n|\psi\rangle}{\sqrt{\langle\psi|P_n|\psi\rangle}}\right) &= S\left(\frac{\mathrm{tr}_B \sum_{i=0}^n P_i^A \otimes P_{n-i}^B|\psi\rangle\langle\psi|P_i^A \otimes P_{n-i}^B}{\langle\psi|P_n|\psi\rangle}\right) \leq \\ &\leq \sum_{i=0}^n \frac{\langle\psi|P_i^A \otimes P_{n-i}^B|\psi\rangle}{\langle\psi|P_n|\psi\rangle}\ S\left(\frac{\mathrm{tr}_B P_i^A \otimes P_{n-i}^B|\psi\rangle\langle\psi|P_i^A \otimes P_{n-i}^B}{\langle\psi|P_i^A \otimes P_{n-i}^B|\psi\rangle}\right) + \\ &\quad + H\left(\left\{\frac{\langle\psi|P_i^A \otimes P_{n-i}^B|\psi\rangle}{\langle\psi|P_n|\psi\rangle}\right\}_{i=0}^n\right)\ . \end{aligned}$$

Clearly, the Shannon entropy $H$ is bounded by $\log(n+1) \leq \log(N+1)$, and thus the l.h.s. of Eq. (2.15), i.e., the entanglement averaged over $n$, is bounded by

$$\sum_{n=0}^N \sum_{i=0}^n \langle\psi|P_i^A \otimes P_{n-i}^B|\psi\rangle\ S\left(\frac{\mathrm{tr}_B P_i^A \otimes P_{n-i}^B|\psi\rangle\langle\psi|P_i^A \otimes P_{n-i}^B}{\langle\psi|P_i^A \otimes P_{n-i}^B|\psi\rangle}\right) + \log[N+1]\ .$$

The sum can be extended to $i = 0, \ldots, N$, $n - i = 0, \ldots, N$ as $|\psi\rangle$ has at most $N$ particles, and by the concavity of the von Neumann entropy, Eq. (2.15) follows. □

Note that this also implies that $E_F^{\mathrm{SSR}}$ is not additive [41]; $E_F^{\mathrm{SSR}}(\rho_{\mathrm{sep}}^{\otimes N})$, e.g., grows at most logarithmically.

Let us now consider $V_F^{\mathrm{SSR}}$ and $V_c^{\mathrm{SSR}}$. As expected, the entanglement cost of $\rho_{\mathrm{sep}}$ vanishes. But as $\rho_{\mathrm{sep}}$ still contains some kind of nonlocality, it is natural to assume that its variance cost is strictly nonzero. In the following, we prove a more general result, namely that $V_F^{\mathrm{SSR}}$ is additive on all states which are a direct sum of pure states (i.e., $\rho$ is block-diagonal and each block is a pure state); this holds, e.g., for $\rho_{\mathrm{sep}}$.

**Theorem 2.5.** *Let* $\rho = \bigoplus_i p_i|\phi_i\rangle\langle\phi_i|$, $\sigma = \bigoplus_j q_j|\psi_j\rangle\langle\psi_j|$, *where* $\sum_i p_i = \sum_j q_j = 1$. *Then*

$$V_F^{\mathrm{SSR}}(\rho\otimes\sigma) = V_F^{\mathrm{SSR}}(\rho) + V_F^{\mathrm{SSR}}(\sigma)\ .$$

**Proof.**

$$\begin{aligned} V_F^{\mathrm{SSR}}(\rho\otimes\sigma) &= V_F^{\mathrm{SSR}}\left(\bigoplus_{i,j} p_i q_j |\phi_i\rangle\langle\phi_i|\otimes|\psi_j\rangle\langle\psi_j|\right) \\ &\overset{(*)}{=} \sum_{i,j} p_i q_j V(\phi_i\otimes\psi_j) \\ &= \sum_i p_i V(\phi_i) + \sum_j q_j V(\psi_j) \\ &\overset{(*)}{=} V_F^{\mathrm{SSR}}(\rho) + V_F^{\mathrm{SSR}}(\sigma)\ , \end{aligned}$$

where in $(*)$ we used that $V_F^{\mathrm{SSR}}(\bigoplus_i r_i|\chi_i\rangle\langle\chi_i|) = \sum_i r_i V(\chi_i)$, with $\sum_i r_i = 1$. As subadditivity is clear from the convexity of $V_F^{\mathrm{SSR}}$, we only have to show superadditivity. For an arbitrary decomposition $|\zeta_j\rangle = \sum_i u_{ji}\sqrt{r_i}|\chi_i\rangle$ of $\bigoplus_i r_i|\chi_i\rangle\langle\chi_i| = \sum_j |\zeta_j\rangle\langle\zeta_j|$ [with an isometry $(u_{ji})$], this follows from

$$\begin{aligned} \sum_j \frac{\langle\zeta_j|\hat{N}_A|\zeta_j\rangle^2}{\langle\zeta_j|\zeta_j\rangle} &\overset{\text{(a)}}{=} \sum_j \frac{(\sum_i u_{ji}^* u_{ji} p_i \langle\chi_i|\hat{N}_A|\chi_i\rangle)^2}{\sum_i u_{ji}^* u_{ji} p_i} \\ &\overset{(2.7)}{\leq} \sum_{i,j} \frac{(u_{ji}^* u_{ji} p_i \langle\chi_i|\hat{N}_A|\chi_i\rangle)^2}{u_{ji}^* u_{ji} p_i} \overset{\text{(b)}}{=} \sum_i p_i \langle\chi_i|\hat{N}_A|\chi_i\rangle^2\ . \end{aligned}$$

Here, we used that (a) $\langle\chi_i|\chi_{i'}\rangle = \delta_{ii'}$, $\langle\chi_i|\hat{N}_A|\chi_{i'}\rangle \propto \delta_{ii'}$; and (b) $\sum_j u_{ji}^* u_{ji} = 1$. □

While it seems plausible that $V_F^{\mathrm{SSR}}$ is additive on all states and we did not find any counterexamples, this is apparently hard to prove. Let us note that unlike for $E_F$, the additivity of $V_F^{\mathrm{SSR}}$ is probably not related to its superadditivity. A counterexample for the superadditivity of $V_F^{\mathrm{SSR}}$ can easily be found,[3] and the direct equivalence proof of Pomeransky [52] cannot be transferred to SiV due to the different structure of the nonlinearity.

[3] E.g, for class of $2\times 2$ qubit states given by $|\phi\rangle = \alpha|00\rangle_A|11\rangle_B + \beta|01\rangle_A|10\rangle_B + \beta|10\rangle_A|01\rangle_B + \alpha|11\rangle_A|00\rangle_B$ with $\alpha = \sqrt{p/2}$, $\beta = \sqrt{(1-p)/2}$, and $p < 1/2$, $V_F^{\mathrm{SSR}}$ is not superadditive with respect to the two subsystems, since $V(|\phi\rangle) = 4p$ and $V_F^{\mathrm{SSR}}(\mathrm{tr}_i[|\phi\rangle\langle\phi|]) = 4p(1-p)$.

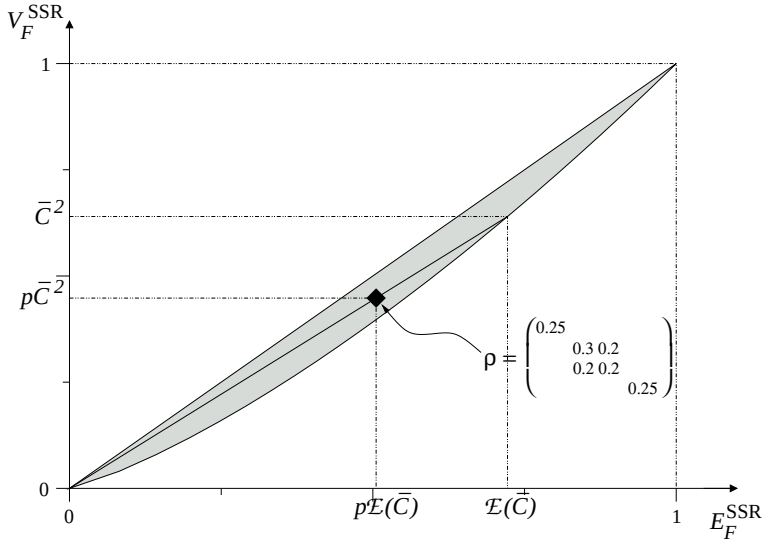


**Figure 2.3:** Relation of $p$, $\bar{C}$, $E_F^{\mathrm{SSR}}$ and $V_F^{\mathrm{SSR}}$ (see Section 2.4.3). The gray area gives the allowed range of $E_F^{\mathrm{SSR}}$ and $V_F^{\mathrm{SSR}}$ for qubits. The lower bound is obtained by plotting $\mathcal{E}(\bar{C})$ vs. $\bar{C}^2$. The point characterizing a mixed state $\rho$ can be found by dividing the line between the origin and the point $(\mathcal{E}(\bar{C}), \bar{C}^2)$ located on the boundary at the ratio of $p : 1-p$.

### 2.4.3 Formation of qubits

In the following, we compute explicit formulas for $E_F^{\mathrm{SSR}}$ and $V_F^{\mathrm{SSR}}$ of qubits. A general bipartite two-qubit state subject to SSR is given by

$$\rho = \begin{pmatrix} w_{00} & & & \\ & w_{01} & \gamma & \\ & \gamma & w_{10} & \\ & & & w_{11} \end{pmatrix} ,$$

where $\gamma \geq 0$ (this can be achieved by local unitaries). Using the results of Wootters [53], we find $E_F(\rho) = \mathcal{E}(C)$, where $\mathcal{E}(C) = H(1/2 + \sqrt{1-C^2}/2)$, $H$ is the binary entropy, and the concurrence $C \equiv C(\rho)$ is here given by

$$C = \max(0, 2\gamma - 2\sqrt{w_{00}w_{11}}) .$$

With SSR, $\rho$ has to be built subspace by subspace, where the one-particle subspace $\rho_1$ is the only one which might be entangled. The concurrence for $\rho_1/\mathrm{tr}[\rho_1]$ is

$$\bar{C} = 2\gamma/p$$

with $p = w_{01} + w_{10} = \mathrm{tr}[\rho_1]$, and thus

$$E_F^{\mathrm{SSR}}(\rho) = p\mathcal{E}(\bar{C}) .$$

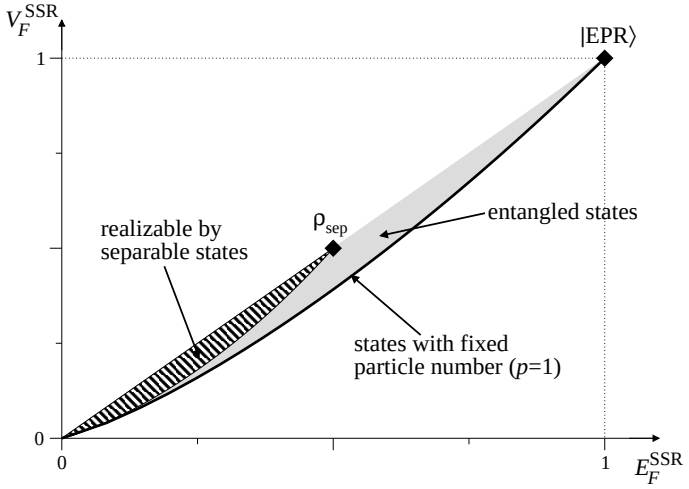


**Figure 2.4:** Different regions of mixed states in the $E$-$V$ diagram. The solid line corresponds to the states with fixed total particle number, i.e., $p = 1$. Separable states have to stay within the dashed area (although there exist entangled states as well). Note that $\rho_{\mathrm{sep}}$ [Eq. (2.14)] is the "extremal separable state".

The relation between the normal concurrence $C$ and the SSR-concurrence $\bar{C}$ is given by the bounds $p\bar{C}-(1-p) \leq C \leq p\bar{C}$, i.e., $E_F$ and $E_F^{\mathrm{SSR}}$ are not completely independent. As $\mathcal{E}$ is concave, $E_F \leq E_F^{\mathrm{SSR}}$, as necessary.

An optimal decomposition of $\rho_1$ can be found as follows. Define $s$ as a root of $\bar{C}/2 = \sqrt{s(1-s)}$. Then, $\rho_1$ can be written as a mixture of $\sqrt{s}|01\rangle + \sqrt{1-s}|10\rangle$ and $\sqrt{1-s}|01\rangle + \sqrt{s}|10\rangle$, and both have the desired EoE.

The same decomposition gives the optimal variance as well (note that this only holds for qubits). Therefore, observe that both states have SiV $4s(1-s) = \bar{C}^2$, i.e., $\bar{C}^2$ is an upper bound for $V_F^{\mathrm{SSR}}(\rho)$, and for pure states equality holds. On the other hand, $\bar{C}^2$ is convex: for any one-particle subblock $\rho_1 = p\sigma + (1-p)\sigma'$ with off-diagonal elements $v = pw + (1-p)w'$ it holds that $v^2 \leq pw^2 + (1-p)w'^2$. Therefore equality holds, and

$$V_F^{\mathrm{SSR}}(\rho) = p\bar{C}^2 \ . \tag{2.16}$$

Thus, with respect to formation $1 \times 1$ qubit states are characterized by two parameters: the weight of $\rho_1$, $p$, and the concurrence of $\rho_1$, $\bar{C}$. It can be checked easily that $0 \leq p, \bar{C} \leq 1$ in order for $\rho$ to be positive. A necessary condition for separable states is $\bar{C} \leq (1-p)/p$ (this is tight, but $p$, $\bar{C}$ do not tell everything about separability, cf. the inequality relating $C$, $\bar{C}$ given above).

Fig. 2.3 shows how for a particular state $p$ and $\bar{C}$ can be determined from the $E$-$V$ diagram, and Fig. 2.4 gives a "phase diagram" for mixed states.

### 2.4.4 Distillation of nonlocal resources

In the following, we consider the problem complementary to formation: given a mixed state, is it possible to *distill* the nonlocal resources contained in this state? This distilled state could then be used to perform some nonlocal task as teleportation with high fidelity. Naturally, there exist two types of distillation

protocols: the first one aims to increase the fidelity of the states being distilled with the final state, while the second returns the target state *itself* with some finite yield in the asymptotic limit.

In the following, we will focus on qubits. Without SSR, it has been shown for both types of distillation how they can be implemented: in the so-called recurrence protocol [12], both parties apply an XOR operation (a bilateral XOR or BXOR) to two copies of the state and then measure one of them; thus, on average they increase their knowledge about the second. Hashing protocols [11], on the contrary, are aimed to asymptotically return a finite yield of pure states: by subsequent application of BXOR operations partial information about the states can be collected in a subset which is then measured; by the law of large numbers, this partial information asymptotically fully determines the remaining states.

The presence of superselection rules inposes severe restrictions on distillation procedures. It has been shown, e.g., that the entanglement contained in *one* copy of a $|\text{V-EPR}\rangle$ cannot be accessed, while for multiple copies of $|\text{V-EPR}\rangle$, all entanglement up to a logarithmic amount can be used [41], as also follows from Theorem 2.2. The central problem in distilling states containing SiV is that the BXOR operation is ruled out by the SSR, and there is no adequate replacement. One way to overcome this problem is to use a third copy of the state as an (imperfect) shared reference frame and to construct a three-copy protocol which probabilistically implements BXOR. Indeed, we will show that one needs three-copy protocols to distill both EoE and SiV. Unfortunately, this is of no use for the implementation of hashing protocols, as the errors of the BXOR-approximation accumulate, and each BXOR uses up the reference frame copy of the state whereas hashing would require $O(N^2)$ BXOR operations.

The existence of two distinct resources makes the field of distillation much more rich: there will occur trade-offs between the two resources in distillation, and one might even think of *spending* one resource to distill the other. For instance, we will show that it is possible to distill all separable but nonlocal states towards $\rho_{\text{sep}}$ [Eq. (2.14)], and in turn, if one adds some entanglement, all the SiV contained in $\rho_{\text{sep}}$ can be converted to a $|\text{V-EPR}\rangle$.

### Reduction to standard states

To simplify analysis, in [11] the distillation of qubits has been considered for a standard form, namely Bell-diagonal states; any state can be made Bell-diagonal by LOCC. Yet, these operations are ruled out by the SSR, so that we have to introduce a different normal form. Therefore, consider a general bipartite qubit state with SSR

$$\rho = \begin{pmatrix} w_{00} & & & \\ & w_{01} & \gamma & \\ & \gamma^* & w_{10} & \\ & & & w_{11} \end{pmatrix} , \tag{2.17}$$

where $w_{ij} \geq 0$ and $\gamma \geq 0$ (the latter can be accomplished by local unitaries). By local filtering operations [54]

$$F_A \propto \sqrt[4]{w_{10}w_{11}}|0\rangle\langle 0| + \sqrt[4]{w_{00}w_{01}}|1\rangle\langle 1| \ ,$$
$$F_B \propto \sqrt[4]{w_{01}w_{11}}|0\rangle\langle 0| + \sqrt[4]{w_{00}w_{10}}|1\rangle\langle 1| \ ,$$

this can be transformed probabilistically to

$$\tilde{\rho} = \frac{1}{2(1+w)} \begin{pmatrix} w & & & \\ & 1 & v & \\ & v & 1 & \\ & & & w \end{pmatrix} \ . \tag{2.18}$$

Here,

$$w = \sqrt{\frac{w_{00}w_{11}}{w_{01}w_{10}}} \ , \quad v = \frac{|\gamma|}{\sqrt{w_{01}w_{10}}} \ .$$

In the following, we will call (2.18) the *standard form* $\tilde{\rho}$ of a two-qubit state $\rho$ and only consider states of this type. The standard form is Bell-diagonal and unique for each $\rho$, and by the reverse POVM $F'_A \propto F_A^{-1}$, $F'_B \propto F_B^{-1}$, $\tilde{\rho}$ is converted back to $\rho$. Thus, any state can be transformed probabilistically to its standard form and back by LOCC, and therefore the standard form of states containing EoE and SiV still contains EoE and SiV.[4]

Note that the two parameters $(w, v)$ describing the standard form $\tilde{\rho}$ are directly related to $(p, \bar{C})$ used to characterize $E_F^{\mathrm{SSR}}(\tilde{\rho})$ and $V_F^{\mathrm{SSR}}(\tilde{\rho})$ in Section 2.4.3: $v = \bar{C}$ and $w = 1/p - 1$.

### Distilling entanglement

Let us first demonstrate that it is possible to distill all entangled qubit states, as it is the case without SSR. Therefore, take two copies of an arbitrary state $\rho$ in its standard form $\tilde{\rho}$ [Eq. (2.18)] and project locally onto the one-particle subspaces. The resulting state in the $\{|\hat{0}\rangle, |\hat{1}\rangle\}$-basis is

$$\hat{\rho} = \begin{pmatrix} w^2 & & & \\ & 1 & v^2 & \\ & v^2 & 1 & \\ & & & w^2 \end{pmatrix} \ .$$

[4] Some special cases for $\rho$ have to be considered separately. Note that if $v = 0$, $\rho$ can be created by LOCC, and $v = 0$ follows if $w_{01} = 0$ or $w_{10} = 0$. If $w_{00} = w_{11} = 0$, filtering still works in both directions if $w_{00}$ and $w_{11}$ are deleted from the filtering operators. In case $w_{00} = 0$, $w_{11} > 0$ or vice versa, the situation gets more complicated. In order to distill $\rho$, one adds $|00\rangle\langle 00|$ with some weight which does not destroy the entanglement; then, $\rho$ can be distilled to $|\text{V-EPR}\rangle$, from where it can be easily reconstructed.

Obviously, $\hat{\rho}$ is entangled iff $\tilde{\rho}$ is entangled iff $\rho$ is entangled, and as $\hat{\rho}$ has constant local particle number, it can be distilled as usual [11, 12]. Therefore, it is possible to distill any entangled two-qubit state if we do not care about its SiV. Even more, if we have an infinite amount of SiV available, we can distill with the same rate as without SSR by using the SiV as a perfect reference frame.

**Distilling separable states**

Clearly, the SiV contained in separable but nonlocal states cannot be distilled, as pure states with SiV always contain entanglement. One solution to this problem is to distill towards $\rho_{\text{sep}}$ [Eq. (2.14)]; we will show how this can be done (and why $\rho_{\text{sep}}$ is a good choice) in the next subsection. Alternatively, one might try to add entanglement (e.g., $|\text{E-EPR}\rangle$s) and then distill the SiV of separable states to $|\text{V-EPR}\rangle$s.

In the following, we show how $\rho_{\text{sep}} \otimes |\text{E-EPR}\rangle\langle\text{E-EPR}|$ can be transformed to a $|\text{V-EPR}\rangle$ with probability 1/2, thereby distilling all the SiV contained in $\rho_{\text{sep}}$ to a pure state. Clearly, $|\text{V-EPR}\rangle$ can be obtained from an $|\text{E-EPR}\rangle = |01\rangle|10\rangle + |10\rangle|01\rangle$ by applying a BXOR operation, but this is ruled out by the SSR. The idea in the following is to use the SiV contained in $\rho_{\text{sep}}$ as a shared reference frame in order to carry out the BXOR operation probabilistically. In order to see how this works, write $\rho_{\text{sep}}$ as a mixture of $(|0\rangle + \omega|1\rangle)_A(|0\rangle + \omega|1\rangle)_B$ over all $\omega = e^{i\phi}$. If we manage to project the total state onto subspaces where $\omega$ simply gives a global phase, we can make use of the SiV of $\rho_{\text{sep}}$. Therefore, start with the state $|\text{E-EPR}\rangle\langle\text{E-EPR}| \otimes \rho_{\text{sep}}$ which can be written as a mixture of the states

$$\begin{aligned} |\psi_0\rangle &\propto |010\rangle|100\rangle + |100\rangle|010\rangle \ , \\ |\psi_1\rangle &\propto |010\rangle|101\rangle + |100\rangle|011\rangle + |011\rangle|100\rangle + |101\rangle|010\rangle \ , \\ |\psi_2\rangle &\propto |011\rangle|101\rangle + |101\rangle|011\rangle \end{aligned}$$

with probabilities 1/4, 1/2, and 1/4. Clearly, there is no measurement which tells us which $|\psi_i\rangle$ we actually have without either destroying the entanglement contained in $|\psi_0\rangle$ and/or $|\psi_2\rangle$ or the variance contained in $|\psi_1\rangle$. As we want to extract the variance, we have to sacrifice the EoE of $|\psi_{0,2}\rangle$: both parties do a projective measurement onto the subspaces spanned by $\{|010\rangle, |101\rangle\}$ and $\{|100\rangle, |011\rangle\}$. If the measurement outcomes match, Alice and Bob share a known state with EoE and SiV 1 which can be converted to a $|\text{V-EPR}\rangle$; otherwise, the entanglement is lost. Both cases are equally likely, and thus the average yield of SiV is $1/2 = V_F^{\text{SSR}}(\rho_{\text{sep}})$ which is optimal. On the other hand, we had to sacrifice half of the entanglement—there is a trade-off between the two resources.

The procedure described above can be generalized to arbitrary states, where it allows to distill the one-particle subblock. Note that if $\rho$ is entangled, the required $|\text{E-EPR}\rangle$'s can be distilled from $\rho$ itself.

### Recurrence protocols

In the following, we will look for protocols which allow to distill EoE *and* SiV. Particularly, we would like to have a protocol which allows to concentrate the SiV contained in separable states. As already mentioned at the beginning of the section, the usual recurrence protocols cannot be applied as BXOR cannot be implemented. (In fact, it is not even possible to find an operation doing a comparable job, i.e., computing the parity, only for $|0\rangle|1\rangle \pm |1\rangle|0\rangle$.) Yet, similar to the preceding subsection we can use a third copy as a shared reference frame which allows to implement the desired recurrence operation in a probabilistic way. Indeed, we will see that three-copy protocols suffice for all distillation tasks.

General $N$-copy recurrence protocols can be represented by local POVM operators which act on $N$ qubits and leave one qubit (i.e., $2 \times 2^N$ matrices). These operators must be realizable by SSR-compatible operations, i.e., by an $N$-qubit POVM, followed by a measurement of all but one qubit (omitting the measurement decreases our information about the state and thus does not help). Therefore, the POVM operators must have the shape of two adjacent rows of SSR-compatible $N$-qubit operations [Eq. (2.1)]; except normalization, this is the only condition.

Possible protocols are illustrated in Fig. 2.5. Any state can be brought to standard form Eq. (2.18) by local filtering operations and can be parametrized by a tuple $(v, w)$, $0 \leq v \leq 1$, $0 \leq w$; the states with $v > w$ are entangled (Fig. 2.5a).

Given a *single* copy of $\tilde{\rho}(v, w)$, Alice and Bob can either increase $w$ (by adding $|00\rangle\langle 00| + |11\rangle\langle 11|$) or decrease $v$ and $w$ by the same fraction (by adding $|01\rangle\langle 01| + |10\rangle\langle 10|$), or anything inbetween, as is illustrated in Fig. 2.5b. Obviously, $\rho_{\text{sep}}$ can be transformed to any other separable state deterministically—therefore, it is indeed *the* standard separable but nonlocal state, as an EPR is for entanglement.

Let us turn our attention to two-copy protocols. As the output state will not necessarily have standard form, we have to include filtering in the local POVM operators which restricts their degrees of freedom to one complex number each, so that it is easy to check that the best protocols are given by

$$M_A = M_B \propto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

and

$$M'_A \propto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \quad M'_B \propto \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The resulting transformations are

$$(v, w) \mapsto (v, \sqrt{\frac{1 + v^2 + w^2}{2}})$$

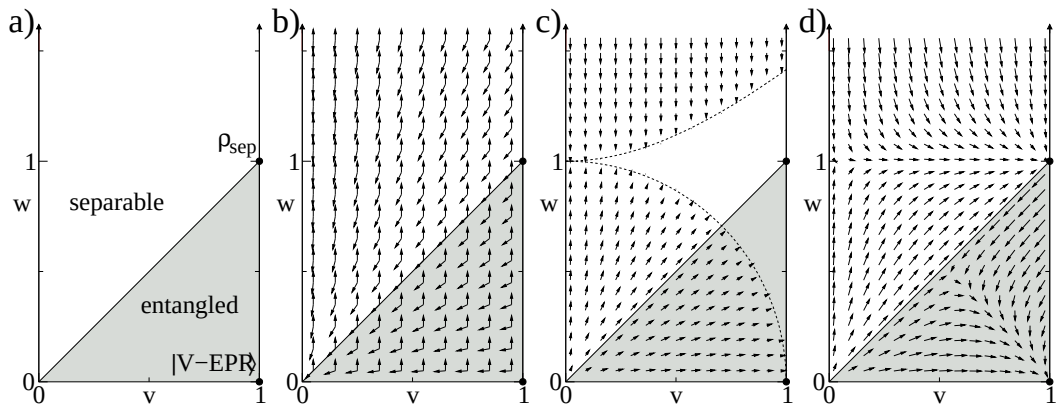


**Figure 2.5:** **a)** Diagram characterizing mixed states according to their standard form. The entangled states are exactly those in the gray area. **b)** Transformations possible by one-copy operations: $w$ can be increased, or $w$ and $v$ can be decreased simultaneously. Thereby, the $|\text{V-EPR}\rangle$ can be transformed to any state, while $\rho_{\text{sep}}$ can generate any separable state. **c)** Additional transformations realizable by two-copy recurrence protocols. Thereby, it is not possible to reach $\rho_{\text{sep}}$ or $|\text{V-EPR}\rangle$. **d)** Three-copy protocols allow to distill all separable states towards $\rho_{\text{sep}}$ and all entangled ones towards $|\text{V-EPR}\rangle$.

and

$$(v, w) \mapsto \sqrt{\frac{2}{1 + v^2 + w^2}}\,(v, w) \ ,$$

respectively. Fig. 2.5c shows where this gives an advantage over the one-copy protocol Fig. 2.5b. Obviously, two-copy protocols do neither allow to distill separable state to $\rho_{\text{sep}}$ nor do they allow to distill entangled states to $|\text{V-EPR}\rangle$.

For three copies, though, the following two pairs of POVM operators provide a way to distill all states:

$$M_A = M_B \propto \begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{pmatrix}$$

distills all separable states to $\rho_{\text{sep}}$ by virtue of

$$(v, w) \mapsto \left( v + \frac{v - v^3}{1 + 2v^2 + 2w^2}, \frac{w(2 + 2v^2 + w^2)}{1 + 2v^2 + 2w^2} \right)$$

whereas

$$M'_A \propto \begin{pmatrix} 0 & 1 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{pmatrix} \ , \quad M'_B \propto \begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 & 1 & 0 \end{pmatrix}$$

distills entangled states towards a $|\text{V-EPR}\rangle$, and

$$(v, w) \mapsto \left( \frac{v(6 + 3v^2 - 2w^2)}{3 + 6v^2 + 6w^2}, \frac{w(6 - 2v^2 + 3w^2)}{3 + 6v^2 + 6w^2} \right) \ .$$

This is illustrated in Fig. 2.5d.

## 2.5 SiV as a resource

### 2.5.1 Introduction

In its standard form, i.e., as a singlet, EoE can be used to teleport quantum bits and thus allows to overcome the restriction to LOCC. In this section, we will show that SiV is a resource in very much the same way, namely it allows to overcome the restrictions imposed by SSR in a bipartite setting. Despite the similarities, there are some major differences. Firstly, while for EoE there only exists one standard form (the maximally entangled state depending on the dimension of the system), for SiV there exist two different standard states: singlets $|0\rangle|N\rangle + |N\rangle|0\rangle$ with SiV $N^2$ (as in Corollary 2.3) and the Gaussian distributed states with large variance (as in Theorem 2.2). Second, there are no pure states which carry only SiV—SiV as a resource which is independent of EoE only exists for mixed states where resources are difficult to quantify. Finally, we will find that we need an infinite amount of nonlocality in order to completely overcome the restrictions imposed by the SSR—this is fundamentally different from EoE where one ebit of entanglement is sufficient to perfectly teleport one quantum bit.

In order to demonstrate (and partly quantify) that SiV is useful to overcome the restrictions imposed by the SSR, we will use the tasks of distinguishing and teleportation. It has been shown that with SSR there exists a perfect data hiding protocol [42] which allows to encode a classical bit in a bipartite state such that it cannot be revealed by LOCC [40]. This protocol can be extended to a protocol hiding $\log N$ bits in the Fourier states

$$|\zeta_k^N\rangle = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} e^{2\pi i k n/N} |n\rangle_A |N-1-n\rangle_B \ . \tag{2.19}$$

These states can be distinguished perfectly by LOCC if no SSR are present (therefore, both parties measure in the Fourier basis and compare their outcomes), but with SSR, they become totally indistinguishable (see Theorem 2.6 below).

The second task we use to show that SiV is a resource is teleportation of a state with nonconstant local particle number: Alice holds one half of a state $|\phi\rangle_{AC}$ which she wants to teleport to Bob using an in general mixed helper state $\rho_{AB}$. Clearly, one ebit of EoE is necessary for this task, but if $V(|\phi\rangle) > 0$, also SiV is needed [40]. We will show that one ebit of EoE is still sufficient, but the amount of SiV has to grow superlinearly with $V(|\phi\rangle)$ and is infinite for perfect teleportation.

### 2.5.2 A general protocol

Let us first quote a Theorem from [40] which will be very useful in the following.

**Theorem 2.6** ([40])**.** *In the presence of superselection rules, the states $\rho$ and $\mathcal{N}_{A|B}(\rho)$ cannot be distinguished by LOCC. Here, $\mathcal{N}_{A|B}$ is the "dephasing map"*

$$\mathcal{N}_{A|B}(\rho) = \sum_{n_A,n_B} P^A_{n_A} \otimes P^B_{n_B} \rho P^A_{n_A} \otimes P^B_{n_B} \ ,$$

*with $P^X_{n_X}$ the local projector onto the subspace with $n_X$ particles.*

By this theorem, we can highly restrict the class of allowed protocols. Let us show this for the task of distinguishing, where Alice and Bob initially share $\rho = |\zeta^N_k\rangle\langle\zeta^N_k| \otimes \sigma$ [cf. Eq. (2.19)], and they have to determine $k$. At the end of the protocol, Alice and Bob get an answer $k'$ according to a probability distribution $\{p_{k'}\}$. But if they had started with $\mathcal{N}_{A|B}(\rho)$ instead, Theorem 2.6 tells us that the probability distribution of their outcomes would have been just the same. Therefore, Alice and Bob can start their protocol by measuring their particle number operators $\hat{N}_A$ and $\hat{N}_B$—if they discard their outcomes, they just implemented $\mathcal{N}_{A|B}$, and their knowledge of $N_A$ and $N_B$ will not affect the *average* probability distribution which is solely relevant unless the figure of merit is nonlinear. The same holds for the teleportation scenario with respect to the partition $A|BC$, i.e., in this case only Alice is allowed to measure her particle number (this map is actually weaker than $\mathcal{N}_{A|BC}$). Note that for a pure state, measuring $\hat{N}_A$ also determines $N_{BC}$ (and thus implements $\mathcal{N}_{A|BC}$), which is different in the mixed state case and closely connected to the fact that mixed states alone are not sufficient for teleportation.

### 2.5.3 Pure states

Assume Alice wants to teleport her share of the state

$$|\phi\rangle = \sum_n \alpha_n |n\rangle_A |-n\rangle_C$$

to Bob using

$$|\psi\rangle = \sum_m \beta_m |m\rangle_A |-m\rangle_B \ .$$

Here, we use a simplified notation, where $-\infty < n, m < \infty$, $\sum |\alpha_n|^2 = \sum |\beta_m|^2 = 1$, and the support of the $\alpha_n$, $\beta_m$ is bounded below such that the particle number can be made positive by adding ancillas. As shown before, Alice can start any protocol by measuring $N_A = K$, yielding

$$\sqrt{p_K}|\chi_K\rangle_{ABC} = \sum_{n+m=K} \alpha_n \beta_m |n,m\rangle_A |-n\rangle_C |-m\rangle_B \tag{2.20}$$

with included probability $p_K$. If Alice now measures in the Fourier basis and communicates her result, the originally tripartite state can be reconstructed by

Bob and Charlie; this strategy is optimal as no information is lost and the state gets less nonlocal. Up to shifts in the particle number, the state is then

$$\sqrt{p_K}|\chi_K\rangle_{BC} = \sum_n \alpha_n \beta_{K-n} |n\rangle_B |-n\rangle \ .$$

We will use the average entanglement fidelity as the figure of merit,

$$\bar{F} = \left\langle \sum_K p_K \left|\langle\phi|\chi_K\rangle\right|^2 \right\rangle = \sum_\Delta \Pi(\Delta) C(\Delta)$$

where $\Pi(\Delta) = \sum_n \langle p_n p_{n+\Delta}\rangle$ ($p_n = |\alpha_n|^2$) and $C(\Delta) = \sum_m \beta_m^* \beta_{m-\Delta}$. The average $\langle\cdot\rangle$ is taken over all states $\phi$, where for teleportation we assume a unitarily invariant distribution. It is straightforward to check that local filtering operations cannot increase $\bar{F}$. Also, for the task of distinguishing it can be shown that $\bar{F}$ gives the optimal success probability for the inconclusive case [55]. For distinguishing, $\Pi_D(\Delta) = \max(N - |\Delta|, 0)/N^2$, while for teleportation, $\Pi_T(\Delta) = [\max(N-|\Delta|, 0)+\delta_{\Delta,0}]/N(N+1)$ [56]. In both cases, $\alpha_n \neq 0$ for $n = 0, \ldots, N-1$.

We will analyze two natural types of helper states: states with constant distribution $\beta_m = 1/\sqrt{M}$, $m = 0, \ldots, M-1$, and states with Gaussian distribution with variance $V(\psi)$. One finds $C_C(\Delta) = \max(M - |\Delta|, 0)/M$, and $C_G(\Delta) = \exp[-\Delta^2/2V] \approx 1 - \Delta^2/2V$. The resulting error probabilities for all four cases are given in Table 2.1. Note that for the Gaussian distributed $|\psi\rangle$, in both cases

$$p_{\text{err}} = \frac{\langle V(\phi)\rangle}{4V(\psi)}$$

holds, i.e., the error probability is given by the *ratio* of the variances. (Actually, this even holds without taking the average over $\phi$.)

In all cases, the error vanishes only if the size of the helper state grows superlinearly with the size of the unknown state; thus, the scaling of SiV as a resource is unfavorable compared to the behaviour of EoE. This is a direct consequence of the direct sum structure in (2.3) which is opposed to the tensor product structure leading to EoE: while with a tensor product structure, $N$ particles generate a $2^N$-dimensional Hilbert space, for the direct sum structure the underlying space only has dimension $N+1$. This also holds for mixed states, where this is the size of the largest coherent subspace. For the same reason, the data hiding scheme Eq. (2.19) is optimal in the sense that the *available* Hilbert space has only dimension $N$.

### 2.5.4 Mixed states

Let us now demonstrate that separable mixed states with SiV can also be used as a shared reference frame [40]. First, we demonstrate how Alice and Bob can use the state $\rho_{\text{sep}}$ to distinguish the states $|01\rangle \pm |10\rangle$. By very much the same

| helper | task | |
|---|---|---|
| | distinguish | teleport |
| constant | $p_{\text{err}} = \frac{(N+1)(N-1)}{3MN}$ | $p_{\text{err}} = \frac{N}{3M}$ |
| Gaussian | $p_{\text{err}} = \frac{(N+1)(N-1)}{12V(\psi)}$ | $p_{\text{err}} = \frac{N(N-1)}{12V(\psi)}$ |

**Table 2.1:** Error probabilities for distinguishing and teleportation where the helper state is either a maximally entangled state with Schmidt number $N$ or a Gaussian distributed state with variance $V(\psi)$.

argument as before Alice and Bob can start their protocol by measuring their local particle number. By *adding* their outcomes, Alice and Bob immediately know whether they are dealing with the $|0\rangle|0\rangle/|1\rangle|1\rangle$ part of $\rho_{\text{sep}}$ or with the $|\text{V-EPR}\rangle$. In the first case all information is lost while in the second case they can just proceed as if they had started with $|\text{V-EPR}\rangle$ itself. This case occurs with probability 1/2, i.e., all the SiV contained in $\rho_{\text{sep}}$ can be used. Clearly, this protocol can not be used for teleportation as $\hat{N}_{BC}$ cannot be implemented locally.

Let us now show that separable but nonlocal states can be used to overcome locality constraints arbitrarily well, i.e., they can serve as arbitrarily precise reference frames. Therefore, we use the separable state [40, 57]

$$\rho_{\text{coh}}(\alpha) = \int \frac{\mathrm{d}\phi}{2\pi} |\alpha e^{i\phi}\rangle\langle\alpha e^{i\phi}| \otimes |\alpha e^{i\phi}\rangle\langle\alpha e^{i\phi}|$$

where for $\alpha > 0$,

$$|\alpha e^{i\phi}\rangle = e^{-\alpha^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} e^{in\phi}$$

is a coherent state with amplitude $\alpha e^{i\phi}$. It has been shown [40] that for $\alpha \to \infty$, $\rho_{\text{coh}}(\alpha)$ can be used to distinguish $|01\rangle \pm |01\rangle$ with arbitrary precision. In the following, we will show that this state together with one $|\text{E-EPR}\rangle$ can be used to perfectly teleport a state with nonconstant local particle number and therefore may serve as an arbitrarily precise reference frame.

First, let us use Theorem 2.5 to show that this state has indeed infinite SiV for $\alpha \to \infty$. Therefore, it is enough to note that

$$\begin{aligned}
\rho_{\text{coh}}(\alpha) &= \sum_{N=0}^{\infty} p_N |\theta_N\rangle\langle\theta_N| \quad \text{with} \\
p_N &= e^{-2\alpha^2} \frac{(2\alpha^2)^N}{N!} \ , \\
|\theta_N\rangle &= \frac{1}{\sqrt{2^N}} \sum_{n=0}^{N} \binom{N}{n}^{1/2} |n, N-n\rangle \ ,
\end{aligned}$$

and thus $V_F^{\mathrm{SSR}}(\rho_{\mathrm{coh}}(\alpha)) = V_c^{\mathrm{SSR}}(\rho_{\mathrm{coh}}(\alpha)) = \alpha^2/2 \to \infty$ for $\alpha \to \infty$. In is interesting to note that each of the $|\theta_N\rangle$ approximates a state with Gaussian distribution such that $\rho_{\mathrm{coh}}(\alpha)$ might be considered as the EoE-free mixed state version of Gaussian distributed states.

In order to see how a mixed state can be used to teleport, let Alice and Charlie initially share $|\phi\rangle = \alpha|01\rangle + \beta|10\rangle$ (the proof is completely analogous for qu-$d$-its), and assume Alice wants to teleport her share to Bob. Therefore, Alice and Bob are provided with an $|\text{E-EPR}\rangle_{AB}$ and with some mixed state

$$\rho = \sum_{\substack{n,m,n',m' \\ n+m=n'+m'}} \rho_{n,m}^{n',m'} |n\rangle_A\langle n'| \otimes |m\rangle_B\langle m'| \ , \tag{2.21}$$

where the condition $n+m = n'+m'$ comes from the SSR, Eq. (2.1). For simplicity, let us assume that all $\rho_{n,m}^{n',m'}$ are nonnegative. Alice once more starts by measuring her local particle number operator on $|\phi\rangle\langle\phi| \otimes \rho$. (In this step, we do not have to care about the $|\text{E-EPR}\rangle$ which has constant local particle number.) For a measurement outcome $n$, the resulting state (probability included) is

$$\begin{aligned}\sum_m \ \Big[ & |\alpha|^2 \rho_{n,m-1}^{n,m-1} |0,n\rangle_A\langle 0,n| \otimes |m-1\rangle_B\langle m-1| \otimes |1\rangle_C\langle 1| \\ & + |\beta|^2 \rho_{n-1,m}^{n-1,m} |1,n-1\rangle_A\langle 1,n-1| \otimes |m\rangle_B\langle m| \otimes |0\rangle_C\langle 0| \\ & + \alpha\beta^* \rho_{n,m-1}^{n-1,m} |0,n\rangle_A\langle 1,n-1| \otimes |m-1\rangle_B\langle m| \otimes |1\rangle_C\langle 0| \\ & + \alpha^*\beta \rho_{n-1,m}^{n,m-1} |1,n-1\rangle_A\langle 0,n| \otimes |m\rangle_B\langle m-1| \otimes |0\rangle_C\langle 1| \Big]\end{aligned}$$

As Alice's share now has constant particle number and lies within a two-dimensional subspace, she can use the $|\text{E-EPR}\rangle_{AB}$ to teleport her share to Bob. If we label the two teleported basis states $|\hat{a}\rangle = |0,n\rangle$, $|\hat{b}\rangle = |1,n-1\rangle$, Bob and Charlie then share the state

$$\begin{aligned}\sum_m \Big[ & |\alpha|^2 \rho_{n,m-1}^{n,m-1} |\hat{a},m-1\rangle_B\langle \hat{a},m-1| \otimes |1\rangle_C\langle 1| \\ & + |\beta|^2 \rho_{n-1,m}^{n-1,m} |\hat{b},m\rangle_B\langle \hat{b},m| \otimes |0\rangle_C\langle 0| \\ & + \alpha\beta^* \rho_{n,m-1}^{n-1,m} |\hat{a},m-1\rangle_B\langle \hat{b},m| \otimes |1\rangle_C\langle 0| \\ & + \alpha^*\beta \rho_{n-1,m}^{n,m-1} |\hat{b},m\rangle_B\langle \hat{a},m-1| \otimes |0\rangle_C\langle 1| \Big]\end{aligned}$$

Bob now projects onto the subspaces spanned by the pairs of states $|0_m\rangle \equiv |\hat{a},m-1\rangle$ and $|1_m\rangle \equiv |\hat{b},m\rangle$ and obtains

$$\begin{aligned} & |\alpha|^2 \rho_{n,m-1}^{n,m-1} |0\rangle_B\langle 0| \otimes |1\rangle_C\langle 1| \\ & + |\beta|^2 \rho_{n-1,m}^{n-1,m} |1\rangle_B\langle 1| \otimes |0\rangle_C\langle 0| \\ & + \alpha\beta^* \rho_{n,m-1}^{n-1,m} |0\rangle_B\langle 1| \otimes |1\rangle_C\langle 0| \\ & + \alpha^*\beta \rho_{n-1,m}^{n,m-1} |1\rangle_B\langle 0| \otimes |0\rangle_C\langle 1| \end{aligned} \tag{2.22}$$

(where we omitted the subscript $m$). By looking at the average fidelity with the original state, we find that the error vanishes iff

$$\sum_{n,m} \rho_{n,m-1}^{n,m-1} = \sum_{n,m} \rho_{n-1,m}^{n-1,m} = \sum_{n,m} \rho_{n-1,m}^{n,m-1} \ . \tag{2.23}$$

Since $\rho$ is positive this implies that $\rho_{n,m-1}^{n,m-1} \approx \rho_{n-1,m}^{n-1,m} \approx \rho_{n-1,m}^{n,m-1}$ for most $n, m$, as one would expect from Eq. (2.22). It is straightforward to check that Eq. (2.23) holds for $\rho_{\mathrm{coh}}(\alpha)$ for $\alpha \to \infty$, and that the $2 \times 2$ subblocks of the density matrix really approximate pure states.

One might expect that $N \to \infty$ copies of $\rho_{\mathrm{sep}}$ could be used just the same way, but the situation is quite different: filtering operations which bring $\rho_{\mathrm{sep}}^{\otimes N}$ into a form (2.21) destroy the off-diagonal elements of the density matrix with high probability so that (2.23) cannot be satisfied; therefore it is questionable whether multiple copies of $\rho_{\mathrm{sep}}$ can be used as an arbitrarily precise reference frame. On the other hand, this is not so much different from the pure state scenario: while multiple copies of a $|\text{V-EPR}\rangle$ might indeed be used as a perfect reference frame, these states carry an amount of entanglement which grows *linearly* with the precision of the reference frame, whereas a single Gaussian distributed state with large SiV only has *logarithmic*—and thus in some sense vanishing—entanglement and is therefore much closer to the case of separable reference frames.

Let us note that the teleportation scenario can be altered by joining $B$ and $C$. This is no longer teleportation, of course, and can be accomplished by LOCC without SSR. On the other hand, it is still an impossible task if SSR are present and is thus suitable to characterize mixed states as reference frames without the need for additional entanglement.

### 2.5.5 Hiding quantum states

Let us close by showing that the data hiding protocol given in [40] resp. its extension Eq. (2.19) can be used to construct a mixed state scheme to hide quantum data as well. At a first glance, one might try to encode the two degrees of freedom of a qubit in the phases of the state $|02\rangle + e^{i\phi}|11\rangle + e^{i\phi'}|20\rangle$, but this cannot be accomplished by a linear map. Therefore, we encode the qubit $\alpha|01\rangle + \beta|10\rangle$ in one of the states $|\phi_0\rangle = \alpha|01\rangle + \beta|10\rangle$, $|\phi_1\rangle = \beta|01\rangle + \alpha|10\rangle$ with equal probabilities which is then distributed between Alice and Bob. Additionally, Alice and Bob are provided with a state $|\psi_{0/1}\rangle = |02\rangle \pm |20\rangle$ which encodes the state Alice and Bob actually share. Thus, Alice and Bob share a state which they cannot distinguish from the totally mixed state by LOCC (Theorem 2.6), but they can perfectly recover the original qubit if they join. This scheme can be extended to hide $N$-level quantum states using one of the states

$$|\phi_k\rangle = \sum_{n=0}^{N-1} \alpha_{n+k \bmod N} |n, N-1-n\rangle \ ; \quad k = 0, \ldots, N-1 \ .$$

Together with the state encoding $k$, $N^2-1$ particles are needed, and the associated Hilbert space dimension is $N^2$.

## 2.6 Conclusions

Adding restrictions to the operations permissible on a quantum system gives rise to a new resource which in turn allows to overcome this restriction. The restriction to LOCC, for example, leads to EoE as a nonlocal resource. Adding SSR to a bipartite system leads to an additional resource, the superselection induced variance SiV. We could show that SiV and EoE together completely characterize bipartite states in the asymptotic limit. Thereby, two different kind of standard forms arise, namely singlets and Gaussian distributed states with logarithmic EoE.

The search for states which only carry SiV led us to mixed states, where we considered entanglement and variance of formation. We could show that the concept of entanglement does not have to be changed and thus there exist states which carry SiV but no EoE, and we provided explicit formulas for the case of qubits. As to distillation, we could show that both EoE and SiV can be distilled, and we provided various ways to do that. Thereby, we found that there exist mixed standard states for SiV which do not carry EoE. While it is possible to extend recurrence protocols such that they work with SSR by using a third copy as a reference frame, it is unlikely that protocols with asymptotic yield work.

Finally, we showed that SiV is a resource which allows to overcome the restrictions imposed by the SSR, but we also saw that there are fundamental differences to EoE as the size of the reference frame has to grow superlinearly with the problem size, which is due to direct sum structure of the underlying Hilbert space.

# Chapter 3

# Optimal generation of squeezing and entanglement

## 3.1 Introduction

Squeezed states are a valuable resource for different fields of physics. They can increase the resolution of precision measurements, as exploited in gravitational wave detectors [58], improve spectroscopic sensitivity [59, 60], and enhance signal-to-noise ratios [61, 62], e.g., in optical communication. Moreover, squeezing acts as a basic building block for the generation of continuous variable entanglement [63–65], which in turn is a cornerstone for quantum information purposes. Unfortunately, squeezing is an expensive resource as well: squeezed states are hard to create and the involved operations are subject to losses and noise inevitably restricting the attainable amount of squeezing. On the other hand, passive operations—in quantum optical setups implemented by beam-splitters and phase shifters—can often be performed at low cost and they are, compared to the squeezers, relatively noiseless.

In this chapter, we address the following question: How can we exploit a given noisy squeezing device in an optimal way, when supplemented by arbitrary noiseless passive operations? We derive the optimal strategy for single and repeated use of the squeezing device, calculate the achievable squeezing, and finally discuss what this implies for the optimal generation of entanglement. To this end, we will use a black box model for the physical squeezing device: This gives us the possibility to derive optimality results which are equally applicable to a wide range of physical implementations.

The argumentation will make use of the covariance matrix formalism, which was mainly developed in the context of continuous variable states having a Gaussian Wigner distribution—so called Gaussian states [66]. The latter naturally appear in quantum optical settings (the field of a light mode) as well as in atomic ensembles (collective spin degrees) and ion traps (vibrational modes). We restrict

ourselves to the natural class of Gaussian operations, i.e., operations preserving the Gaussian character of a state [67–69]. This includes all time evolutions governed by operators quadratic in bosonic creation and annihilation operators. All the presented results hold for an arbitrary number of modes, and although it might be reasonable to think in terms of Gaussian states, we do not have to restrict the input states to be Gaussian.

## 3.2 Gaussian states and operations

We will begin with introducing the notation and some basic results [66–70]. Consider a system of $N$ bosonic modes with the respective canonical operators $(Q_1, P_1, \ldots, Q_N, P_N) = \vec{R}$. These are related to the annihilation operators via $a_j = (Q_j + iP_j)/\sqrt{2}$ and satisfy the canonical commutation relations $[R_k, R_l] = i\sigma_{kl}\mathbb{1}$ governed by the symplectic matrix

$$\sigma = \bigoplus_{i=1}^{N} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} .$$

The displacement $\vec{d}$ in phase space and the covariance matrix (CM) $\gamma$ corresponding to a state $\rho$ are then given by

$$\begin{aligned} d_i &= \mathrm{tr}[\rho R_i] \ , \\ \gamma_{ij} &= \mathrm{tr}\big[\rho\{(R_i - d_i), (R_j - d_j)\}_+\big] \ , \end{aligned}$$

where $\{\cdot,\cdot\}_+$ denotes the anti-commutator.

While for coherent states $\gamma = \mathbb{1}$, a state is called squeezed if its uncertainty in some direction in phase space is below the uncertainty of the coherent state, i.e., if $s(\gamma) \equiv \lambda_{\min}(\gamma) < 1$, where $s(\gamma)$ is the *squeezing* of $\gamma$ measured by its smallest eigenvalue [71]. Note that by this definition, *more* squeezing means a *smaller* $s$. As the squeezing is independent of the displacement $\vec{d}$, we omit the latter in the following.

Let us now focus on Gaussian operations. Unitary Gaussian operations are precisely those realizable by quadratic Hamiltonians, so that they naturally appear in many physical systems. In phase space they act as symplectic operations $S \in \mathrm{Sp}(2N)$ on the covariance matrix $\gamma \mapsto S^T\gamma S$. Symplectic operations preserve the canonical commutation relations and are thus given by the group $\mathrm{Sp}(2N) = \{S \in \mathbb{R}^{2N\times 2N} \ : \ S^T\sigma S = \sigma\}$. An important subgroup is given by the group of orthogonal symplectic transformations $\mathrm{K}(2N) = \mathrm{O}(2N) \cap \mathrm{Sp}(2N)$ [71]. Physically, these correspond to passive operations, which can, in quantum optical setups, be implemented by beam-splitters and phase shifters [72]. Obviously, passive transformations cannot change the squeezing, since elements from $\mathrm{K}(2N)$ preserve the spectrum and in particular the smallest eigenvalue of the CM.

We will now introduce the model we use to describe the squeezing device. In general, a noisy operation $\mathcal{E}$ can be regarded as a noiseless map on a larger system including the environment, which is discarded afterwards. If the overall time-evolution is governed by a quadratic Hamiltonian and the environment is in a Gaussian state (e.g., a thermal reservoir), it can be shown that these operations are exactly the ones which act as

$$\gamma \mapsto \mathcal{E}(\gamma) = X^T \gamma X + Y \ , \qquad X^T i\sigma X + Y \geq i\sigma \tag{3.1}$$

with $X, Y \in \mathbb{R}^{2N \times 2N}$ [67]. Here, the equation on the right hand side ensures complete positivity, i.e., guarantees that the operation is physically reasonable. While the $X^T \gamma X$ part of $\mathcal{E}$ represents a joint rotation and distortion of the input $\gamma$, the $Y$ contribution is a noise term which may consist of quantum as well as classical noise. In the following, we will consider squeezing devices of the type in Eq. (3.1), as these are the ones which naturally appear in many experiments. We will however show later on that the results even hold for arbitrary Gaussian operations (which may include measurements and conditional feedback operations).

The question we are going to investigate is the following: Given some noisy device $\mathcal{E}$ and the set of passive operations $\mathrm{K}(2N)$, how can we generate squeeezing as efficiently as possible from a given input state? Naturally, this general question can be asked in several specific ways. First of all, one might ask how much squeezing can be generated by a *single application* of the device given a certain initial state, as in Fig. 3.1a.

The much more interesting question, of course, relates to an iterative scenario, i.e., how can we generate squeezing as efficiently as possible by *repeated application* of $\mathcal{E}$ with passive operations $K_i$ inbetween (Fig. 3.1b)? In this scenario we may either allow for different $K_i$ or choose them identically as it is for instance the case in a ring cavity setup (Fig. 3.1c).

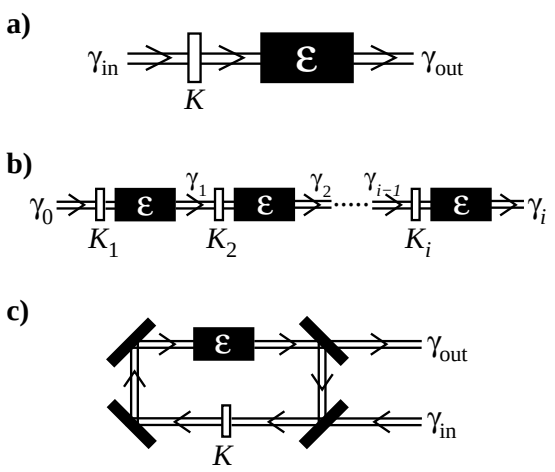


**Figure 3.1:** Various scenarios for the optimization of squeezing: **a)** *single iteration case*: from a given input $\gamma_{\text{in}}$, we want to generate as much squeezing as possible by properly choosing $K$ and using the noisy device $\mathcal{E}$; **b)** *multiple iteration case*: the device can be applied repeatedly, and we have to determine the $K_i$'s such as to maximize the finally obtained squeezing; **c)** *circular setup*: with identical $K_i = K$.

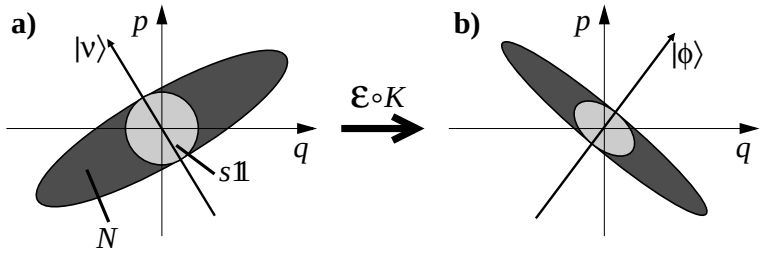


**Figure 3.2:** Illustration of the single iteration optimality proof. Left: the input state is decomposed into a coherent kernel $s\mathbb{1}$ and into some extra noise $N$ with a null eigenvector $|\nu\rangle$. Right: after the application of the operation $\mathcal{E}$, the coherent kernel is squeezed in the direction $|\phi\rangle$. By choice of $K$, one can achieve $K(X|\phi\rangle) \propto |\nu\rangle$, i.e., the noise $N$ leads to no contribution in the most squeezed direction.

## 3.3 Single iteration case

In order to prepare for the more complicated scenarios, let us first have a look at the case of a single iteration (Fig. 3.1a), starting from a given input CM $\gamma$. This is the basic building block for all the iterative protocols. We now use a formal trick in order to facilitate the derivation: we split the CM into two parts (Fig. 3.2a),

$$\gamma = s\mathbb{1} + N \ , \tag{3.2}$$

where $s = s(\gamma)$ is the squeezing of $\gamma$. The first part can be regarded as a "coherent kernel" of $\gamma$. It may have sub-Heisenberg variance (if $s < 1$) and it is invariant under passive operations. The second part is a "noise term" $N$ which ensures that $\gamma$ is a physical state. As $s$ is the smallest eigenvalue of $\gamma$, $N \geq 0$ has a null space.

Let us now see what happens if we rotate $\gamma = s\mathbb{1} + N$ by some passive operation $K$ and then send it through $\mathcal{E}$: the coherent kernel is invariant under $K$, and thus

$$\gamma' \equiv \mathcal{E}(K^T \gamma K) = \mathcal{E}(s\mathbb{1}) + X^T K^T N K X \ ,$$

i.e., the action of the squeezing device on the "coherent kernel" plus the "noise" part which has been rotated and squeezed by $X$. Note that the first part no more depends on the choice of the passive operation; furthermore, as $N \geq 0$ and thus $X^T K^T N K X \geq 0$, the smallest eigenvalue of $\mathcal{E}(s\mathbb{1})$ gives a bound to the squeezing of the output. In the following, we show that this bound can be achieved. Therefore, let $s_0 = s(\mathcal{E}(s\mathbb{1})) = \lambda_{\min}[sX^T X + Y]$ be the squeezing obtained from the input $s\mathbb{1}$ with corresponding eigenvector $|\phi\rangle$.[1] On the one

[1] Throughout this chapter, we use bra-kets to denote ordinary vectors. This is *not* a quantum state.

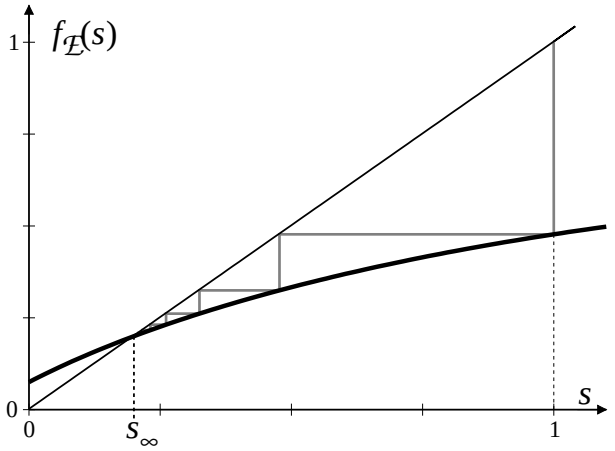


**Figure 3.3:** The function $s' = f_{\mathcal{E}}(s)$ describes the optimal output squeezing $s'$ which can be obtained from input squeezing $s$ by an operation $\mathcal{E}$. The straight line is the identity, and the gray zig-zag line describes the evolution of $f$ when using the optimal strategy for multiple iterations, cf. Fig. 3.1b.

hand, $s' \equiv s(\gamma') \geq s_0$, and on the other hand,

$$\begin{aligned} s' \leq \langle\phi|\mathcal{E}(K^T\gamma K)|\phi\rangle &= \langle\phi|\mathcal{E}(s\mathbb{1}) + X^T K^T N K X|\phi\rangle \\ &= s_0 + \langle\phi|X^T K^T N K X|\phi\rangle \ . \end{aligned}$$

Recall that by definition (3.2), $N$ has a null eigenvector which we denote by $|\nu\rangle$, $N|\nu\rangle = 0$. By choosing $K$ such that $K(X|\phi\rangle) \propto |\nu\rangle$ (which can be always done with passive $K$, cf. [71]), the second term vanishes, and we indeed find that $s' = s_0 = s(\mathcal{E}(s\mathbb{1}))$.

The proof is also illustrated in Fig. 3.2: Choosing $K$ appropriately ensures that the noise $N$ does not contribute in the most squeezed direction. Note that for a given $\mathcal{E}$ it is now straight forward to derive an optimal $K$ and by exploiting the results of [72] to decompose it into an array of beam-splitters and phase shifters.

For a single iteration of $\mathcal{E}$ this shows that the optimally achievable squeezing $s'$ at the output is given by the squeezing obtained from the non-physical input $s\mathbb{1}$:

$$s' = f_{\mathcal{E}}(s) := \lambda_{\min}[sX^T X + Y] \ . \tag{3.3}$$

It is highly interesting to note that, therefore, the optimal value of the final squeezing does *only* depend on the initial squeezing (and on the properties of $\mathcal{E}$), but on no other property of the initial state, irrespective of the number of modes considered. It can be easily checked that the respective function $f_{\mathcal{E}}$ in Eq. (3.3) is concave, monotonously increasing, and $f_{\mathcal{E}}(0) \geq 0$ (Fig. 3.3).

## 3.4 Multiple iteration case

The fact that the optimal output squeezing does only depend on the input squeezing immediately implies that in the case of multiple iterations the passive operations $K_i$ can be optimized successively in order to obtain the global optimum. This is a remarkable result, as in general problems of this kind require optimization over all parameters (i.e., over all $K_i$) simultaneously. Graphically, the squeezing in each iteration step moves along a zig-zag line between $f_{\mathcal{E}}(s)$ and the identity, as shown in Fig. 3.3. The optimal output squeezing $s_\infty$ (for number of iterations $\to \infty$) is determined by $f_{\mathcal{E}}(s_\infty) = s_\infty$, $s_\infty \geq 0$. By inserting the definition (3.3) of $f_{\mathcal{E}}$ and solving for $s_\infty$ we obtain

$$s_\infty = \frac{-1}{\lambda_{\min}[(X^T X - \mathbb{1})Y^{-1}]} \ .$$

The convergence of $s(\gamma_i)$ to the optimal value is exponentially fast and bounded from above and below by the slope of $f_{\mathcal{E}}$ at $s_\infty$ and the slope from $(s_\infty, f_{\mathcal{E}}(s_\infty))$ to the starting point $(s_0, f_{\mathcal{E}}(s_0))$, respectively. Note that for $s(\gamma_{\text{in}}) < s_\infty$, however, $s(\gamma_{\text{out}}) > s(\gamma_{\text{in}})$, so that squeezing is destroyed by applying the operation $\mathcal{E}$.

In realistic physical scenarios, it might be difficult to tune the passive operations independently and it is more likely that the *same* physical device will be passed again and again, e.g., in a ring cavity (cf. Fig. 3.1c), and thus only one fixed passive operation $K \equiv K_i \ \forall i$ can be implemented. In the following, we demonstrate that this is already sufficient to reach the optimal squeezing $s_\infty$. The proper $K$ is the one which preserves the squeezing at the optimality point, corresponding to a zig-zag line along the tangent of $f_{\mathcal{E}}$ at $s_\infty$. The convergence is thus still exponentially fast.

In order to see how this works, consider the non-physical output

$$\tilde{\gamma} = \mathcal{E}(s_\infty \mathbb{1}) = s_\infty X^T X + Y$$

obtained from the input $s_\infty \mathbb{1}$. By the properties of $s_\infty$, it is clear that $\lambda_{\min}(\tilde{\gamma}) = s_\infty$ with a corresponding *normalized* eigenvector $|\psi_\infty\rangle$. Now choose $K_\infty$ such that

$$|\psi_\infty\rangle \propto K_\infty X |\psi_\infty\rangle \ . \tag{3.4}$$

This is exactly the $K$ which preserves the optimality point, as $|\psi_\infty\rangle$ is the null eigenvector of $\tilde{\gamma} - s_\infty \mathbb{1}$. For any initial state $\gamma$ with $s(\gamma) > s_\infty$, choose the decomposition $\gamma_0 = s_\infty \mathbb{1} + P$ , where $P > 0$ and $\lambda_{\min}(P) + s_\infty = s_0 \equiv s(\gamma_0)$. After one iteration $\gamma_1 = \mathcal{E}(K_\infty^T \gamma_0 K_\infty)$, we have

$$\begin{aligned} s_1 &\leq \langle \psi_\infty | \gamma_1 | \psi_\infty \rangle = s_\infty + \langle \psi_\infty | X^T K_\infty^T P K_\infty X | \psi_\infty \rangle \\ &\overset{(3.4)}{=} s_\infty + \langle \psi_\infty | P | \psi_\infty \rangle \ \underbrace{\langle \psi_\infty | X^T X | \psi_\infty \rangle}_{\alpha} \ . \end{aligned}$$

As we will show in a moment, $\alpha < 1$, and for the squeezing $s_n \equiv s(\gamma_n)$ after $n$ iterations it holds by recursion that $s_n \leq s_\infty + \alpha^n \langle\psi_\infty|P|\psi_\infty\rangle$, which converges exponentially to $s_\infty$. From

$$s_\infty \langle\psi_\infty|X^T X|\psi_\infty\rangle + \langle\psi_\infty|Y|\psi_\infty\rangle = s_\infty$$

it follows that

$$\alpha \equiv \langle\psi_\infty|X^T X|\psi_\infty\rangle = \frac{s_\infty - \langle\psi_\infty|Y|\psi_\infty\rangle}{s_\infty}$$

which is positive and strictly smaller than 1 as long as $\langle\psi_\infty|Y|\psi_\infty\rangle > 0$, which is generically the case.[2]

## 3.5 General Gaussian maps

In the following, we show that the results just obtained for channels of the type (3.1) also hold for the most general type of Gaussian channels which may include measurements and postprocessing. Channels of this kind appear, e.g., in the creation of spin squeezing using quantum nondemolition measurements with feedback [73]. The most general memoryless operation on $N$ modes can described by a $2N$ mode covariance matrix $\Gamma$ via the Jamiolkowski isomorphism [69] as

$$\mathcal{E}(\gamma) = A - C(B+\gamma)^{-1}C^T\,, \text{ where } \tilde{\Gamma} = \begin{pmatrix} A & C \\ C^T & B \end{pmatrix}\,, \tag{3.5}$$

and $\tilde{\Gamma}$ is the partial transpose of $\Gamma$. Again, $s \equiv s(\gamma)$, $|\psi\rangle$ is the eigenvector corresponding to the smallest eigenvalue of $\mathcal{E}(s\mathbb{1})$, and we need to show that $K$ can be choosen such that $s(\mathcal{E}(s\mathbb{1})) = s(\mathcal{E}(K\gamma K^T))$, i.e., that $\langle\psi|C(B+s\mathbb{1})^{-1}C^T|\psi\rangle = \langle\psi|C(B+K\gamma K^T)^{-1}C^T|\psi\rangle$. This means that for $|\chi\rangle \equiv C^T|\psi\rangle$, $P \equiv (B+s\mathbb{1}) > 0$, and $N \equiv \gamma - s\mathbb{1} \geq 0$ with a null eigenspace we have to find a $K$ such that

$$\begin{aligned} &\langle\chi|P^{-1} - (P+KNK^T)^{-1}|\chi\rangle = \\ &\langle\chi|P^{-1/2}\underbrace{\left[\mathbb{1} - \left[\mathbb{1} + P^{-1/2}KNK^TP^{-1/2}\right]^{-1}\right]}_{(*)}P^{-1/2}|\chi\rangle \end{aligned} \tag{3.6}$$

vanishes. Since $(*)$ has the same null eigenspace as $P^{-1/2}KNK^TP^{-1/2}$, the expression (3.6) can be indeed made zero by choosing $K$ such that $|\nu\rangle \propto K^TP^{-1}|\chi\rangle$ (where $N|\nu\rangle = 0$). This proves that for any Gaussian operation the optimal output squeezing can be computed on $s\mathbb{1}$, thus generalizing the previous results.

[2]For singular $Y$, the same result follows after a straightforward but tedious discussion.

## 3.6 An example

Let us now consider an example which illustrates how the representation of the operation in Eq. (3.1) is related to the master equation $\dot{\rho} = i[\rho, \mathcal{H}] + \mathcal{L}[\rho]$ of a system. From it, one obtains a master equation for the evolution of the covariance matrix,

$$\dot{\gamma} = A\gamma + \gamma A^T + N \ . \tag{3.7}$$

For the case of photon losses to a vacuum reservoir, $\mathcal{L}[\rho] = \nu(2a\rho a^\dagger - a^\dagger a\rho - \rho a^\dagger a)$, one obtains $A = 2\sigma H - \nu\mathbb{1}$ and $N = 2\nu\mathbb{1}$, where $H$ is the Hamiltonian matrix, i.e., $\mathcal{H} = (Q, P)H(Q, P)^\dagger$. By integration, one finds that applying (3.7) for a time $t$ leads to a map $\mathcal{E}_t : \gamma \mapsto X^T\gamma X + Y$ with $X = e^{-\nu t}\exp[-2H\sigma t]$ and

$$Y = X^T \left[2\nu \int_0^t e^{2\nu\tau} e^{-2\sigma H\tau} e^{2H\sigma\tau} \mathrm{d}\tau\right] X \ .$$

At this point, the previously derived results can be applied. Fig. 3.4 shows a one mode example with $\mathcal{H} = \frac{3}{4}a^\dagger a - \frac{1}{4}a^\dagger a^\dagger + \text{h.c.}$ and the noise level $\nu = 0.1$. Note that additional passive operations enhance the attainable squeezing although the noise $N$ is rotationally invariant.

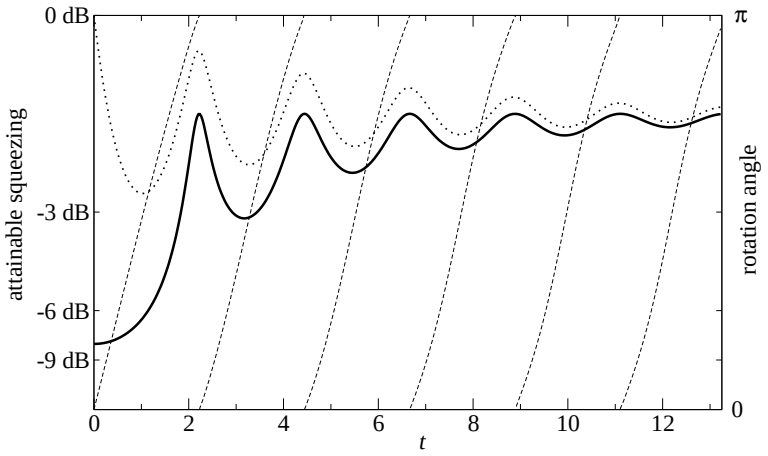


**Figure 3.4:** Optimization results for a physical device $\mathcal{E}_t$ which is given by applying the master equation for a time $t$, with $H = \frac{3}{4}a^\dagger a - \frac{1}{4}a^\dagger a^\dagger + \text{h.c.}$ and $\nu = 0.1$ (cf. text). The dotted line shows the squeezing obtained by simply applying $\mathcal{E}_t$ to a coherent state, while the solid line gives the asymptotically attainable optimum as derived. The dashed line shows the rotation angle which leads to the optimal value.

## 3.7 Optimal entanglement generation

Let us now consider the optimal generation of entanglement from noisy operations. In a practical scenario, this question corresponds to a situation where we are again given arbitrary passive operations, including such which couple the modes which we want to entangle. In this setting, the optimality result on squeezing generation has direct implications: It has been shown that the amount of entanglement which can be generated by passive operations starting from two squeezed Gaussian input states only depends on their squeezing irrespective of the number of modes [25]; for two modes, e.g., this is done by sending the states onto a beam splitter after rotating them into orthogonal directions. Using this result, we can immediately determine how much entanglement we can generate from Gaussian inputs using a black box $\mathcal{E}$ supplemented by passive operations, namely $E_{\mathcal{N},\mathrm{opt}} = \log(s_{\mathrm{opt}}^{-1})$, where $s_{\mathrm{opt}}$ is the maximal squeezing generated by $\mathcal{E}$ and the entanglement is measured by the logarithmic negativity [74]. Moreover, by combining the results it is straightforward to explicitly derive the optimal entangling protocol for any given Gaussian device.

While this answers the question of optimal entanglement generation from the practical point of view, where squeezing is typically expensive while entangling passive operations as beam splitters are cheap, there remains the quantum information theoretic perspective of entanglement as a resource. In that case, the natural setup would be the following: Given two parties, Alice and Bob, wo have access to some entangling channel plus free local operations, what is the best way to create as much entanglement between them as quickly as possible? In the following, we show for a specific instance of this problem that exactly as for the case of squeezing, the maximum amount of entanglement for multiple interations is obtained by maximizing the entanglement created in each step.

We restrict to channels $\mathcal{E}$ which act on $1+1$ modes and are symmetric, i.e., invariant under interchange of Alice and Bob. These channels are of the form

$$\mathcal{E} : \gamma \mapsto X^T \gamma X + Y$$

with

$$X = \begin{pmatrix} A & C \\ C & A \end{pmatrix} \text{ and } X = \begin{pmatrix} B & D \\ D & B \end{pmatrix}$$

(in Alice-Bob partition) where $B = B^T$, $D = D^T$.

In the following, we show that starting from a symmetric state $\gamma$, the maximal entanglement of

$$\gamma' = \mathcal{E}\left((S_A \oplus S_B)^T \gamma (S_A \oplus S_B)\right) \tag{3.8}$$

(where the maximization is over the local operations $S_A$, $S_B$) solely depends on the entanglement of $\gamma$. This in turn implies once more that in the iterative scenario where starting from a coherent state, (3.8) is repeated several times, the global optimum can be obtained by maximizing the entanglement generated in

each step. However, this result is based on the conjecture that the optimal output entanglement in (3.8) for a symmetric input $\gamma$ can obtained with a symmetric sympectic operation, $S_A = S_B$. We have tested this conjecture extensively for both the robustness and the Gaussian entanglement of formation (see below) and it seems to hold, even for the case of symmetric channels of the more general form (3.5). On the other hand, for finite-dimensional systems a counterexample can be easily found.[3]

We will measure the entanglement using the so-called "robustness" [69]

$$p(\gamma) := \sup\{0 \leq p \leq 1 | p(\gamma_A \oplus \gamma_B) \leq \gamma)\} \ , \tag{3.9}$$

where $\gamma_A, \gamma_B \geq i\sigma$. For $p = 1$, the state is separable, while for $p = 0$, the state is infinitely entangled; thus, strictly speaking $1 - p$ is an entanglement monotone. For symmetric $1 + 1$ mode states, it coincides with the negativity as well as with the entanglement of formation.

It holds that every symmetric state $\gamma$ can be written up to a sympectic transformation $S \oplus S$ as[4]

$$\gamma = p\mathbb{1} + \lambda P_q^+ + \mu P_p^- \ , \tag{3.10}$$

where $p \equiv p(\gamma)$, and $P_{q,p}^{\pm} = |\omega_{q,p}^{\pm}\rangle\langle\omega_{q,p}^{\pm}|$ are projectors onto $\langle\omega_q^{\pm}| = (1, 0, \pm 1, 0)$ and $\langle\omega_p^{\pm}| = (0, 1, 0, \pm 1)$. Note the close analogy to the decomposition (3.2) for squeezed states. We will use a linearized version of the robustness for symmetric states,

$$p(\gamma) = \frac{1}{2} \inf_{L_2} \mathrm{tr}\left[L_2^T \gamma L_2 (P_p^+ + P_q^-)\right] \ ,$$

where $L_2 = L \oplus L$ is a local and symmetric sympectic operation [77].

In order to simplify the following derivation, we now change to a "magic basis", which in this case is the basis spanned by $\left\{|\omega_q^+\rangle, |\omega_p^+\rangle, |\omega_q^-\rangle, |\omega_p^-\rangle\right\}$. Effectively, this transformation exchanges the role of "product" and "symmetric". In

[3] This counterexample is due to K. G. Vollbrecht [75]: Consider the channel

$$\rho \mapsto \mathcal{E}(\rho) = |\psi\rangle\langle\psi| \ \langle 01|\rho|01\rangle \ + \ \mathbb{F}|\psi\rangle\langle\psi|\mathbb{F} \ \langle 10|\rho|10\rangle \ + \ |00\rangle\langle 00| \ (\langle 00|\rho|00\rangle + \langle 11|\rho|11\rangle)$$

where $\mathbb{F}$ flips Alice's and Bob's system. This channel is clearly symmetric, and the maximum entanglement is obtained for $|01\rangle$ and $|10\rangle$ at the input—a state $|\psi\rangle$ for which the entanglement of formation strictly decreases when it is mixed with $\mathbb{F}|\psi\rangle$ or $|00\rangle$ can easily be found.

[4] This can be seen as follows: Take the symmetric $\gamma$ in the standard form $\left(\begin{smallmatrix} n & c \\ c & n \end{smallmatrix}\right)_q \oplus \left(\begin{smallmatrix} n & -d \\ -d & n \end{smallmatrix}\right)_p$ in $q$-$p$ partition, which can be always obtained by symmetric transformations [76]. By local symmetric squeezing, this can be transformed to $\left(\begin{smallmatrix} r & c \\ c & r \end{smallmatrix}\right)_q \oplus \left(\begin{smallmatrix} s & -d \\ -d & s \end{smallmatrix}\right)_p$, where $r - c = s - d$. The optimal $\gamma_A \oplus \gamma_B$ in (3.9) is then of the form $\left(\begin{smallmatrix} x & \\ & y \end{smallmatrix}\right)_A \oplus \left(\begin{smallmatrix} x & \\ & y \end{smallmatrix}\right)_B$ (if not, twirl with the transposition $\theta \oplus \theta$, $\theta = \left(\begin{smallmatrix} 1 & \\ & -1 \end{smallmatrix}\right)$, and with the flip). It is now easy to see that this implies that $x = y$ for the optimal $p$, and the representation (3.10) follows.

particular, the matrices involved transform as

$$S_2 = \begin{pmatrix} S & \\ & S \end{pmatrix} \quad \rightarrow \quad S_2 = \begin{pmatrix} S & \\ & S \end{pmatrix} ,$$

$$X = \begin{pmatrix} A & B \\ B & A \end{pmatrix} \quad \rightarrow \quad X = \begin{pmatrix} A+B & \\ & A-B \end{pmatrix} =: \begin{pmatrix} F_+ & \\ & F_- \end{pmatrix} ,$$

$$Y = \begin{pmatrix} C & D \\ D & C \end{pmatrix} \quad \rightarrow \quad Y = \begin{pmatrix} C+D & \\ & C-D \end{pmatrix} =: \begin{pmatrix} G_+ & \\ & G_- \end{pmatrix} .$$

Thus, the original non-local symmetric channel is replaced by non-symmetric product (i.e. local) channel. The linearized robustness in the magic basis reads

$$p(\gamma) = \frac{1}{2} \inf_L \operatorname{tr} \left[ (L \oplus L)^T \gamma (L \oplus L) \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{pmatrix} \right] . \tag{3.11}$$

In the following, we will show that a result analogous to the squeezing case holds, namely that one can determine the optimal output robustness by considering the non-physical input $p\mathbb{1}$ instead of $\gamma$, which in turn yields the corresponding result that the optimal entanglement solely depends on the entanglement of the input. Let $S \oplus S$ be the transformation which has to be applied to $p\mathbb{1}$ to achieve the maximum output robustness, and let $L$ be the optimal choice for that output in (3.11). Then, the optimal output robustness is

$$\begin{aligned} p_{\text{opt}}(\gamma_{\text{out}}) = \frac{1}{2} \big\{ & p \langle 1| L^T F_+^T S^T \mathbb{1} \underbrace{S F_+ L |1\rangle}_{=:|v_+\rangle} + \langle 1| L^T G_+ L |1\rangle + \\ & p \langle 0| L^T F_-^T S^T \mathbb{1} \underbrace{S F_- L |0\rangle}_{=:|v_-\rangle} + \langle 0| L^T G_- L |0\rangle \big\} . \end{aligned} \tag{3.12}$$

We will now show that adding $P_q^+$ and $P_p^-$ to the input [cf. Eq. (3.10)] does not affect the optimal robustness; therefore, we will use the up to now unused passive degree of freedom of $S$. As the symplectic operation $S$ is chosen optimally, any change of $S$—corresponding to an arbitrary change of the input CM $\mathbb{1}$ in (3.12)—will lead to an increase of the sum in (3.12). Thus, for any CM $\gamma$,

$$\langle v_+|\gamma|v_+\rangle + \langle v_-|\gamma|v_-\rangle \geq \langle v_+|\mathbb{1}|v_+\rangle + \langle v_-|\mathbb{1}|v_-\rangle , \tag{3.13}$$

with $|v_\pm\rangle$ defined in (3.12). This implies that

$$\langle v_+|v_+\rangle = \langle v_-|v_-\rangle \tag{3.14}$$

[otherwise, if w.l.o.g. $\langle v_+|v_+\rangle > \langle v_-|v_-\rangle$, a state properly squeezed in direction of $|v_+\rangle$ will violate (3.13)], and

$$\langle v_+|v_-\rangle = 0$$

[otherwise, again a state properly squeezed in direction of $|v_+\rangle$, together with (3.14), will violate (3.13)].

Thus, there exists a compact operation $K$ s.th. $K^T\binom{1}{0} \propto |v_-\rangle$, $K^T\binom{0}{1} \propto |v_+\rangle$, and we replace $S$ by $KS$ in (3.12). Then, one immediately sees that adding $\lambda P_q^+ + \mu P_p^-$ to $p\mathbb{1}$, which corresponds to adding $\lambda \left(\begin{smallmatrix} 1 & \\ & 0 \end{smallmatrix}\right) \oplus \mu \left(\begin{smallmatrix} 0 & \\ & 1 \end{smallmatrix}\right)$ in the magic basis, does not change the optimum in (3.12), since

$$\underbrace{\langle 1|L^T F_+^T S^T}_{\langle v_+|} K^T \left(\begin{smallmatrix} 1 & \\ & 0 \end{smallmatrix}\right) K \underbrace{SF_+L|1\rangle}_{|v_+\rangle} = 0$$

and

$$\underbrace{\langle 1|L^T F_-^T S^T}_{\langle v_-|} K^T \left(\begin{smallmatrix} 0 & \\ & 1 \end{smallmatrix}\right) K \underbrace{SF_-L|1\rangle}_{|v_-\rangle} = 0 \ .$$

# Chapter 4

# Gaussian states on harmonic lattices

## 4.1 Introduction

The importance of bosonic Gaussian states arises from two facts. First, they provide a very good description for accessible states of a large variety of physical systems. In fact, every ground and thermal state of a quadratic bosonic Hamiltonian is Gaussian and remains so under quadratic time evolutions. In this way quadratic approximations naturally lead to Gaussian states. Hence, they are ubiquitous in quantum optics as well as in the description of vibrational modes in solid states, ion traps or nanomechanical oscillators.

The second point for the relevance of Gaussian states is that they admit a powerful phase space description which enables us to solve quantum many-body problems which are otherwise (e.g., for spin systems) hardly tractable. In particular, the phase space dimension, and with it the complexity of many tasks, scales linearly rather than exponentially in the number of involved subsystems. For this reason quadratic Hamiltonians and the corresponding Gaussian states also play a paradigmatic role as they may serve as an exactly solvable toy model from which insight into other quantum systems may be gained.

Exploiting the symplectic tools of the phase space description, exact solutions have been found for various problems in quantum information theory as well as in quantum statistical mechanics. In fact, many recent works form a bridge between these two fields as they address entanglement questions for asymptotically large lattices of quadratically coupled harmonic oscillators: the entropic area law [78–80] has been investigated as well as entanglement statics [81–83], dynamics [84–86], and frustration [77, 87].

In this chapter, we analytically derive general properties of ground states of translationally invariant quadratic Hamiltonians on a cubic lattice. Related investigations of correlation functions were recently carried out in [88, 89] for

| interaction | non-critical | critical |
|---|---|---|
| local | $O^*\big(e^{-n/\xi}\big)$<br>$d=1$: $\xi \sim \frac{1}{\sqrt{\Delta m^*}} = \frac{v}{\Delta}$ | $d=1:\quad O^*\Big(\frac{1}{n^2}\Big)$<br>$d>1:\quad O\Big(\frac{\log n}{n^{d+1}}\Big)$ |
| $O\big(n^{-\infty}\big)$ | $O\big(n^{-\infty}\big)$ | |
| $O\Big(\frac{1}{n^\alpha}\Big)$<br>$\alpha > 2d+1$ | $O\Big(\frac{1}{n^{\nu-d}}\Big)$<br>$\alpha > \nu \in \mathbb{N}$ | |
| $\frac{c}{n^\alpha}$<br>$d=1$ | $\alpha \geq 2:\ \Theta\Big(\frac{1}{n^\alpha}\Big)$ | $\alpha = 3:\ \begin{cases} \Theta\big(\frac{1}{n^2}\big) & , c>0 \\ \Theta\big(\frac{\sqrt{\log n}}{n^2}\big) & , c<0 \end{cases}$<br>$\alpha > 3:\quad \Theta\big(\frac{1}{n^2}\big)$ |

**Table 4.1:** Summary of the bounds derived on the asymptotic scaling of ground state correlations, depending on the scaling of the interaction (left column). Here, $n$ is the distance between two points (harmonic oscillators) on a cubic lattice of dimension $d$. $O$ denotes upper bounds, $O^*$ tight upper bounds, and $\Theta$ the exact asyptotics. The table shows the results for generic interactions—special cases are discussed in the text.

finite dimensional spin systems and in [78, 90] for generic harmonic lattices with non-critical finite range interactions.

We start by giving an outlook and a non-technical summary of the main results. The results on the asymptotic scaling of ground state correlations are summarized in Table 4.1.

In Section 4.2, we introduce some basic results on quadratic Hamiltonians together with the used notation. We then turn towards translationally invariant systems in Section 4.3. First, we show that every pure translational invariant Gaussian state is point symmetric: This implies that the spectral gap of the symmetrized rather than the original Hamiltonian determines the characteristic properties of the ground state. We provide a general formula for the latter and express its covariance matrix in terms of a product of the inverse of the Fourier transformed spectral function and the Hamiltonian matrix.

We then investigate the behavior of non-critical systems (Section 4.4) and show that if the Hamiltonian is gapped, the correlations decay according to the interaction: a (super) polynomial decay of the interaction leads to the same (super) polynomial decay for the correlations, and (following Ref. [78]) finite

range interactions lead to exponentially decaying correlations. In Section 4.5, we particularly focus on the relation of correlation length and gap. We derive an explicit formula for the correlation length for gapped 1D-Hamiltonians with finite range interactions. The correlation length $\xi$ is expressed in terms of the dominating zero of the complex spectral function, which close to a critical point is in turn determined by the spectral gap $\Delta$ and the effective mass $m^*$ at the band gap via $\xi \sim (m^*\Delta)^{-1/2}$. When the change in the Hamiltonian is given by a global scaling of the interactions, this proves the folk theorem $\xi \sim 1/\Delta$.

Finally, in Section 4.6 we turn towards critical systems. We show that for generic $d$-dimensional critical systems the correlations decay as $1/n^{d+1}$, where $n$ is the distance between two points on the lattice. Whereas for sufficiently fast decreasing interactions in $d = 1$ the asymptotic bound is exactly polynomial, it contains an additional logarithmic correction for $d \geq 2$. Similarly for $d = 1$ a logarithmic deviation is found if the interaction decays exactly like $-1/n^3$.

## 4.2 Quadratic Hamiltonians and their ground states

Consider a system of $N$ bosonic modes which are characterized by $N$ pairs of canonical operators $(Q_1, P_1, \ldots, Q_N, P_N) =: R$. The canonical commutation relations (CCR) are governed by the symplectic matrix $\sigma$ via

$$\big[R_k, R_l\big] = i\sigma_{kl} \ , \quad \sigma = \bigoplus_{n=1}^{N} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \ ,$$

and the system may be equivalently described in terms of bosonic creation and annihilation operators $a_l = (Q_l + iP_l)/\sqrt{2}$. Quadratic Hamiltonians are of the form

$$\mathcal{H} = \frac{1}{2} \sum_{kl} H_{kl} R_k R_l \ ,$$

where the Hamiltonian matrix $H$ is real and positive semidefinite due to the Hermiticity and lower semi-boundedness of the Hamiltonian $\mathcal{H}$. Without loss of generality we neglect linear and constant terms since they can easily be incorporated by a displacement of the canonical operators and a change of the energy offset. Before we discuss the general case we mention some important special instances of quadratic Hamiltonians: a well studied 1D example of this class is the case of nearest neighbor interactions in the position operators of harmonic oscillators on a chain with periodic boundary conditions

$$\mathcal{H}_\kappa = \frac{1}{2} \sum_{i=1}^{N} Q_i^2 + P_i^2 - \kappa \, Q_i Q_{i+1} \ , \quad \kappa \in [-1, 1] \ . \tag{4.1}$$

This kind of spring-like interaction was studied in the context of information transfer [84], entanglement statics [81–83] and entanglement dynamics [86]. Moreover, it can be considered as the discretization of a massive bosonic continuum theory given by the Klein-Gordon Hamiltonian

$$\mathcal{H}_{\mathrm{KG}} = \frac{1}{2}\int_{-L/2}^{L/2}\Big[\dot{\phi}(x)^2 + \big(\nabla\phi(x)\big)^2 + m^2\phi(x)^2\Big]dx\ ,$$

where the coupling $\kappa$ is related to the mass $m$ by $\kappa^{-1} = 1 + \frac{1}{2}\big(\frac{mL}{N}\big)^2$ [82]. Other finite range quadratic Hamiltonians appear as limiting cases of finite range spin Hamiltonians via the Holstein–Primakoff approximation [91]. In this way the $xy$-spin model with transverse magnetic field can for instance be mapped onto a quadratic bosonic Hamiltonian in the limit of strong polarization where $a \simeq (\sigma_x + i\sigma_y)/2$. Longer range interactions appear naturally for instance in 1D systems of trapped ions. These can either be implemented as Coulomb crystals in Paul traps or in arrays of ion microtraps. When expanding around the equilibrium positions, the interaction between two ions at position $i$ and $j \neq i$ is—in harmonic approximation—of the form $\frac{c\,Q_iQ_j}{|i-j|^3}$, where $c > 0$ ($c < 0$) if $Q_i$, $Q_j$ are position operators in radial (axial) direction [92].

Let us now return to the general case and briefly recall the normal mode decomposition [93]: every Hamiltonian matrix can be brought to a diagonal normal form by a congruence transformation with a symplectic matrix $S \in \mathrm{Sp}(2N,\mathbb{R}) = \{S|S\sigma S^T = \sigma\}$:[1]

$$SHS^T = \bigoplus_{i=1}^{I}\begin{pmatrix}\varepsilon_i & 0\\ 0 & \varepsilon_i\end{pmatrix} \oplus \bigoplus_{j=1}^{J}\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}, \quad \varepsilon_i > 0\ , \tag{4.2}$$

where the *symplectic eigenvalues* $\varepsilon_i$ are the square roots of the duplicate nonzero eigenvalues of $\sigma H\sigma^T H$. The diagonalizing symplectic transformation $S$ has a unitary representation $U_S$ on Hilbert space which transforms the Hamiltonian according to

$$U_S\mathcal{H}U_S^\dagger = \tfrac{1}{2}\sum_{i=1}^{I}\big(Q_i^2 + P_i^2\big)\,\varepsilon_i + \tfrac{1}{2}\sum_{j=1}^{J}P_j^2 = \sum_{i=1}^{I}\big(a_i^\dagger a_i + \tfrac{1}{2}\big)\varepsilon_i + \tfrac{1}{2}\sum_{j=1}^{J}P_j^2\ . \tag{4.3}$$

Hence, by Eq. (4.3) the ground state energy $E_0$ and the energy gap $\Delta$ can easily be expressed in terms of the symplectic eigenvalues of the Hamiltonian matrix:

$$E_0 = \tfrac{1}{2}\sum_{i=1}^{I}\varepsilon_i\ , \qquad \Delta = \begin{cases}\min_i \varepsilon_i\ , & J = 0\\ 0\ , & J > 0\end{cases}\ . \tag{4.4}$$

[1]Note that we disregard systems where the Hamiltonian contains irrelevant normal modes.

The case of a vanishing energy gap $\Delta = 0$ is called *critical* and the respective ground states are often qualitatively different from those of non-critical Hamiltonians. For the Hamiltonian $\mathcal{H}_\kappa$, Eq. (4.1), this happens in the strong coupling limit $|\kappa| = 1 - \Delta^2 \to 1$, and in the case of 1D Coulomb crystals a vanishing energy gap in the radial modes can be considered as the origin of a *structural phase transition* where the linear alignment of the ions becomes unstable and changes to a zig-zag configuration [94–96]. Needless to say that these phase transitions appear as well in higher dimensions and for various different configurations [97].

Ground and thermal states of quadratic Hamiltonians are *Gaussian states*, i.e, states having a Gaussian Wigner distribution in phase space. In the mathematical physics literature they are known as bosonic *quasi-free states* [98, 99]. These states are completely characterized by their first moments $d_k = \mathrm{tr}\big[\rho R_k\big]$ (which are w.l.o.g. set to zero in our case) and their *covariance matrix* (CM)

$$\gamma_{kl} = \mathrm{tr}\Big[\rho\big\{R_k - d_k, R_l - d_l\big\}_+\Big] \,, \tag{4.5}$$

where $\{\cdot,\cdot\}_+$ is the anticommutator. The CM satisfies $\gamma \geq i\sigma$, which expresses Heisenberg's uncertainty relation and is equivalent to the positivity of the corresponding density operator $\rho \geq 0$. In order to find the ground state of a quadratic Hamiltonian, observe that

$$\tfrac{1}{2}\sum_i \varepsilon_i \overset{(4.4)}{=} E_0 = \inf_\rho \mathrm{tr}[\rho\mathcal{H}] \overset{(4.5)}{=} \tfrac{1}{4}\inf_\gamma \mathrm{tr}[\gamma H] \,. \tag{4.6}$$

By virtue of Eqs. (4.2,4.3) the infimum is attained for the ground state covariance matrix

$$\gamma = \lim_{s\to\infty} S^T \left[\bigoplus_{i=1}^{I} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \oplus \bigoplus_{j=1}^{J} \begin{pmatrix} s & 0 \\ 0 & s^{-1} \end{pmatrix}\right] S \,, \tag{4.7}$$

which reduces to $\gamma = S^T S$ in the non-critical case. Note that the ground state is unique as long as $H$ does not contain irrelevant normal modes [which we have neglected from the very beginning in Eq. (4.2)].

In many cases it is convenient to change the order of the canonical operators such that $R = (Q_1, \ldots, Q_N, P_1, \ldots, P_N)$. Then the covariance matrix as well as the Hamiltonian matrix can be written in block form

$$H = \begin{pmatrix} H_Q & H_{QP} \\ H_{QP}^T & H_P \end{pmatrix} \,.$$

In this representation a quadratic Hamiltonian is particle number preserving iff $H_Q = H_P$ and $H_{QP} = -H_{QP}^T$, that is, the Hamiltonian contains only terms of the kind $a_i^\dagger a_j + a_j^\dagger a_i$. In quantum optics terms of the form $a_i^\dagger a_j^\dagger$, which are not number preserving, are neglected within the framework of the *rotating wave approximation*. The resulting Hamiltonians have particular simple ground states:

**Theorem 4.1a.** *The ground state of any particle number preserving Hamiltonian is the vacuum with $\gamma = \mathbb{1}$, and the corresponding ground state energy is given by $E_0 = \frac{1}{4}\mathrm{tr}H$.*

**Proof.** Number preserving Hamiltonians are most easily expressed in terms of creation and annihilation operators. For this reason we change to the respective complex representation via the transformation

$$H \;\mapsto\; \Omega H \Omega^T = \begin{pmatrix} 0 & X \\ \bar{X} & 0 \end{pmatrix} , \quad \Omega = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1} & -i\mathbb{1} \\ \mathbb{1} & i\mathbb{1} \end{pmatrix} .$$

In this basis $H$ is transformed to normal form via a block diagonal unitary transformation $U \oplus \bar{U}$ which in turn corresponds to an element of the orthogonal subgroup of the symplectic group $\mathrm{Sp}(2N, \mathbb{R}) \cap \mathrm{SO}(2N) \simeq \mathrm{U}(N)$ [100]. Hence, the diagonalizing $S$ in Eqs. (4.2,4.7) is orthogonal and since $J = 0$ due to particle number conservation, we have $\gamma = S^T S = \mathbb{1}$. $E_0$ follows then immediately from Eq. (4.6). □

Another important class of quadratic Hamiltonians for which the ground state CM takes on a particular simple form corresponds to the case $H_{QP} = 0$ where there is no coupling between the momentum and position operators:

**Theorem 4.1b.** *For a quadratic Hamiltonian with Hamiltonian matrix $H = H_Q \oplus H_P$ the ground state energy and the ground state CM are given by*

$$E_0 = \tfrac{1}{2}\mathrm{tr}\big[\sqrt{H_Q}\sqrt{H_P}\big],\ \gamma = X \oplus X^{-1},\ X = H_Q^{-1/2}\sqrt{H_Q^{1/2} H_P H_Q^{1/2}} H_Q^{-1/2}. \quad (4.8)$$

**Proof.** Since $\sigma H \sigma^T H = H_P H_Q \oplus H_Q H_P$, the symplectic eigenvalues of $H$ are given by the eigenvalues of $\sqrt{H_Q}\sqrt{H_P}$ and thus $E_0 = \frac{1}{2}\mathrm{tr}\big[\sqrt{H_Q}\sqrt{H_P}\big]$. Moreover, by the uniqueness of the ground state and the fact that $E_0 = \frac{1}{4}\mathrm{tr}[\gamma H]$ with $\gamma$ from Eq. (4.8) we know that $\gamma$ is the ground state CM (as it is an admissible pure state CM by construction). □

Finally we give a general formula for the ground state CM in cases where the blocks in the Hamiltonian matrix can be diagonalized simultaneously. This is of particular importance as it applies to all translational invariant Hamiltonians discussed in the following sections.

**Theorem 4.1c.** *Consider a quadratic Hamiltonian for which the blocks $H_Q$, $H_P$, $H_{QP}$ of the Hamiltonian matrix can be diagonalized simultaneously and in addition $H_{QP} = H_{QP}^T$. Then with*

$$\hat{\mathcal{E}} \;=\; \sqrt{H_Q H_P - H_{QP}^2} \qquad \textit{we have} \quad (4.9)$$

$$E_0 \;=\; \tfrac{1}{2}\mathrm{tr}[\hat{\mathcal{E}}]\ , \quad \Delta = \lambda_{\min}\big(\hat{\mathcal{E}}\big)\ , \quad \gamma = (\hat{\mathcal{E}} \oplus \hat{\mathcal{E}})^{-1} \sigma H \sigma^T\ . \quad (4.10)$$

**Proof.** Since $\sigma H \sigma^T H = \hat{\mathcal{E}}^2 \oplus \hat{\mathcal{E}}^2$ we have indeed $E_0 = \frac{1}{2}\mathrm{tr}[\hat{\mathcal{E}}]$ and $\Delta = \lambda_{\min}(\hat{\mathcal{E}})$. Positivity $\gamma \geq 0$ is implied by $H \geq 0$ such that we can safely talk about the symplectic eigenvalues of $\gamma$. The latter are, however, all equal to one due to $(\gamma\sigma)^2 = -\mathbb{1}$ so that $\gamma$ is an admissible pure state CM. Moreover it belongs to the ground state since $\frac{1}{4}\mathrm{tr}[H\gamma] = E_0$. □

## 4.3 Translationally invariant systems

Let us now turn towards translationally invariant systems. We consider cubic lattices in $d$ dimensions with periodic boundary conditions. For simplicity we assume that the size of the lattice is $N^d$. The system is again characterized by a Hamiltonian matrix $H_{kl}$, where the indices $k, l$, which correspond to two points (harmonic oscillators) on the lattice, are now $d$-component vectors in $\mathbb{Z}_N^d$. Translational invariance is then reflected by the fact that any matrix element $A_{kl}$, $A \in \{H_Q, H_P, H_{QP}\}$ depends only on the relative position $k - l$ of the two points on the lattice, and we will therefore often write $A_{k-l} = A_{kl}$. Note that due to the periodic boundary conditions $k - l$ is understood modulo $N$ in each component. Matrices of this type are called *circulant*, and they are all simultaneously diagonalized via the Fourier transform

$$\begin{aligned} \mathcal{F}_{\alpha\beta} &= \frac{1}{\sqrt{N}}\, e^{\frac{2\pi i}{N}\alpha\beta}\,, \quad \alpha, \beta \in \mathbb{Z}_N\,, \qquad \text{such that} \\ \hat{A} &:= \mathcal{F}^{\otimes d} A \mathcal{F}^{\dagger \otimes d} = \mathrm{diag}\left( \sum_{n \in \mathbb{Z}_N^d} A_n\, e^{-\frac{2\pi i}{N} m\, n} \right)_m\,, \end{aligned}$$

where $m\,n$ is the usual scalar product in $\mathbb{Z}_N^d$. It follows immediately that all circulant matrices mutually commute.

In the following, we will show that we can without loss of generality restrict ourselves to point-symmetric Hamiltonians, i.e., those for which $H_{QP} = H_{QP}^T$ (which means that $\mathcal{H}$ contains only pairs $Q_k P_l + Q_l P_k$). For dimension $d = 1$ this is often called reflection symmetry.

**Theorem 4.2.** *Any translationally invariant pure state CM $\Gamma$ is point symmetric.*

**Proof.** For the proof, we use that any pure state covariance matrix can be written as

$$\Gamma = \begin{pmatrix} \Gamma_Q & \Gamma_{QP} \\ \Gamma_{QP}^T & \Gamma_P \end{pmatrix} = \begin{pmatrix} X & XY \\ YX & X^{-1} + YXY \end{pmatrix}\,,$$

where $X \geq 0$ and $Y$ is real and symmetric [101]. From translational invariance, it follows that all blocks and thus $X$ and $Y$ have to be circulant and therefore commute. Hence, $\Gamma_{QP} = XY = YX = \Gamma_{QP}^T$, i.e., $\Gamma$ is point symmetric. □

Let $\mathcal{P} : \mathbb{Z}_N^d \to \mathbb{Z}_N^d$ be the reflection on the lattice and define the symmetrization operation $\mathcal{S}(A) = \frac{1}{2}(A + \mathcal{P}A\mathcal{P})$ such that by the above theorem $\mathcal{S}(\gamma) = \gamma$ for every translational invariant pure state CM. Then due to the cyclicity of the trace we have for any translational invariant Hamiltonian

$$\inf_\gamma \mathrm{tr}\big[H\gamma\big] = \inf_\gamma \mathrm{tr}\big[\mathcal{S}(H)\gamma\big] \ .$$

Hence, the point-symmetrized Hamiltonian $\mathcal{S}(H)$, which differs from $H$ by the off-diagonal block $\mathcal{S}(H_{QP}) = \frac{1}{2}(H_{QP} + H_{QP}^T)$ has both the same ground state energy and the same ground state as $H$. Together with Theorem 4.1c this leads us to the following:

**Theorem 4.3.** *Consider any translationally invariant quadratic Hamiltonian. With* $\hat{\mathcal{E}} = \big[H_Q H_P - \frac{1}{4}(H_{QP} + H_{QP}^T)^2\big]^{1/2}$ *the ground state CM and the corresponding ground state energy are given by*

$$E_0 = \tfrac{1}{2}\mathrm{tr}[\hat{\mathcal{E}}] \ , \quad \gamma = \big(\hat{\mathcal{E}} \oplus \hat{\mathcal{E}}\big)^{-1} \sigma \mathcal{S}(H) \sigma^T \ . \tag{4.11}$$

It is important to note that the energy gaps of $H$ and $\mathcal{S}(H)$ will in general be different. In particular $H$ might be gapless while $\mathcal{S}(H)$ is gapped. However, as we will see in the following sections, the properties of $\gamma$ depend on the gap $\Delta = \lambda_{\min}(\hat{\mathcal{E}})$ of the symmetrized Hamiltonian rather than on that of the original $H$. For this reason we will in the following for simplicity assume $H_{QP} = H_{QP}^T$. By Theorem 4.3 all results can then also be applied to the general case without point symmetry if one only keeps in mind that $\Delta$ is the gap corresponding to $\mathcal{S}(H)$.

Note that the eigenvalues of $\hat{\mathcal{E}}$ are the symplectic eigenvalues of $\mathcal{S}(H)$, i.e., $\mathcal{E} = \mathcal{F}^{\otimes d} \hat{\mathcal{E}} \mathcal{F}^{\dagger \otimes d}$ is the excitation spectrum of the Hamiltonian. This is the reason for the notation where $\mathcal{E}$ resides in Fourier space and $\hat{\mathcal{E}}$ in real space, which is differs from the normal usage of the hat.

### Correlation functions

According to Eqs. (4.9–4.11) we have to compute the entries of functions of matrices in order to learn about the entries of the covariance matrix. This is most conveniently done by a double Fourier transformation, where one uses that $\widehat{f(M)} = f(\hat{M})$, and we find

$$[f(M)]_{nm} = \frac{1}{N^d} \sum_{r,s} e^{-\frac{2\pi i}{N} nr} [f(\hat{M})]_{rs} e^{\frac{2\pi i}{N} sm} \ . \tag{4.12}$$

As we consider translationally invariant systems, $M$ is circulant and thus $\hat{M}$ is diagonal. We define the function

$$\hat{M}(\phi) = \sum_{n \in \mathbb{Z}_N^d} M_n \, e^{-in\phi} \tag{4.13}$$

such that $\hat{M}(2\pi r/N) = \hat{M}_{r,r}$. As $f(M)$ is solely determined by its first row, we can write

$$[f(M)]_n = \frac{1}{N^d} \sum_{r \in \mathbb{Z}_N^d} e^{2\pi i \, nr/N} f(\hat{M}(2\pi r/N)) \; . \tag{4.14}$$

In the following we will use the index $n \in \mathbb{Z}^d$ for the relative position of two points on the lattice. Their distance will be measured either by the $l_1, l_2$ or $l_\infty$ norm. Since we are considering finite dimensional lattices these are all equivalent for our purpose and we will simply write $\|n\|$. In the thermodynamic limit $N \to \infty$, the sum in Eq. (4.14) converges to the integral

$$[f(M)]_n = \frac{1}{(2\pi)^d} \int_{\mathcal{T}^d} \mathrm{d}\phi \, f(\hat{M}(\phi)) \, e^{in\phi} \quad \text{with} \quad \hat{M}(\phi) = \sum_{n \in \mathbb{Z}^d} M_n \, e^{-in\phi} \; , \tag{4.15}$$

where $\mathcal{T}^d$ is the $d$-dimensional torus, i.e., $[0, 2\pi]^d$ with periodic boundary conditions. The convergence holds as soon as $\sum |M_n| < \infty$ [which holds e.g. for $M_n = O(\|n\|^{-\alpha})$ with some $\alpha > d$] and $f$ is continuous on an open interval which contains the range of $\hat{M}$.

From the definition (4.15) of $\hat{M}$, it follows that $\hat{M} \in \mathscr{C}^k(\mathcal{T}^d)$ (the $n$ times continuously differentiable functions on $\mathcal{T}^d$) whenever the entries $M_n$ decay at least as fast as $\|n\|^{-\alpha}$ for some $\alpha > k + d$, since then the sum of the derivatives converges uniformly. Particularly, if the entries of $M$ decay faster than any polynomial, then $\hat{M} \in \mathscr{C}^\infty(\mathcal{T}^d)$. In the following the most important function of the type $f \circ \hat{M}$ will be the *spectral function*

$$\mathcal{E}(\phi) = \sqrt{\sum_{n \in \mathbb{Z}^d} e^{-in\phi} \Big( [H_Q H_P]_n - [H_{QP}^2]_n \Big)} \; . \tag{4.16}$$

**Asymptotic notation**

In this chapter, we investigate the asymptotic scaling of correlations, which we will describe using the Landau symbols $o$, $O$, and $\Theta$, as well as the symbol $O^*$ for tight bounds:

- $f(x) = o(g(x))$ means $\lim\limits_{x \to \infty} \frac{f(x)}{g(x)} = 0$, i.e., $f$ vanishes strictly faster than $g$ for $x \to \infty$;
- $f(x) = O(g(x))$, if $\limsup\limits_{x \to \infty} \left| \frac{f(x)}{g(x)} \right| < \infty$, i.e., $f$ vanishes at least as fast as $g$;
- $f(x) = \Theta(g(x))$, if $f(x) = O(g(x))$ and $g(x) = O(f(x))$ (i.e., exact asymptotics);

- $f(x) = O^*(g(x))$, if $f(x) = O(g(x))$ but $f(x) \neq o(g(x))$, i.e., $g$ is a tight bound on $f$.[2] If $f$ is taken from a set (e.g., those function consistent with the assumptions of a theorem) we will write $f = O^*(g)$ if $g$ is a tight bound for at least one $f$ (i.e., the best possible universal bound under the given assumptions).

If talking about Hamiltonians, the scaling is meant to hold for all blocks, e.g., if the interaction vanishes as $O(\|n\|^{-\alpha})$ for $n \to \infty$, this holds for all the blocks $H_Q$, $H_P$, and $H_{QP} = H_{PQ}^T$. The same holds for covariance matrices in the non-critical case. By the shorthand notation $f(n) = o(\|n\|^{-\infty})$, we mean that $f(n) = o(\|n\|^{-\alpha})\ \forall \alpha > 0$. Note finally that the Landau symbols are also used in (Taylor) expansions around a point $x_0$ where the considered limit is $x \to x_0$ rather than $x \to \infty$.

## 4.4 Non-critical systems

In this section, we analyze the ground state correlations of non-critical systems, i.e., those which exhibit an energy gap $\Delta > 0$ between the ground an the first excited state. Simply speaking, we will show that the decay of correlations reflects the decay of the interaction. While local (super-polynomially decaying) interactions imply exponentially (super-polynomially) decaying correlations, a polynomial decay of interactions will lead to the same polynomial law for the correlations.

According to Theorem 4.3, we will consider a translationally invariant system with a point-symmetric Hamiltonian ($H_{QP} = H_{QP}^T$). Following (4.10, 4.11), we have to determine the entries of $(\hat{\mathcal{E}}^{-1} \oplus \hat{\mathcal{E}}^{-1})\sigma H \sigma^T$, with $\hat{\mathcal{E}} = (H_Q H_P + H_{QP}^2)^{1/2}$. In Lemma 4.4 we will first show that it is possible to consider the two contributions independently, and as the asymptotics of $\sigma H \sigma^T$ is known, we only have to care about the entries of $\hat{\mathcal{E}}^{-1}$, i.e., we have to determine the asymptotic behavior of the integral

$$(\hat{\mathcal{E}}^{-1})_n = \frac{1}{(2\pi)^d} \int_{\mathcal{T}^d} \mathrm{d}\phi\, \mathcal{E}^{-1}(\phi) e^{in\phi} \ ,$$

where $\mathcal{E} = (\hat{H}_Q \hat{H}_P + \hat{H}_{QP}^2)^{1/2}$.

**Lemma 4.4.** *Given two asymptotic circulant matrices $A$, $B$ in $d$ dimensions with polynomially decaying entries, $A_n = O(\|n\|^{-\alpha})$, $B_n = O(\|n\|^{-\beta})$, $\alpha, \beta > d$. Then*

$$(AB)_n = O^*(\|n\|^{-\mu}) \ , \quad \mu := \min\{\alpha, \beta\} \ .$$

[2] In order to see the difference to $\Theta$, take an $f(x) = g(x)$ for even $x$, $f(x) = 0$ for odd $x$, $x \in \mathbb{N}$. Although $f$ does not bound $g$, thus $f(x) \neq \Theta(g(x))$, the bound $g$ is certainly tight. A situation like this is met, e.g., in Theorem 4.9, where the correlations oscillate within an exponentially decaying envelope.

**Proof.** With $Q_\eta(n) := \min\{1, \|n\|^{-\eta}\}$, we know that $|A_n| = O(Q_\alpha)$ and $|B_n| = O(Q_\beta)$, and

$$|(AB)_n| = \left|\sum_j A_{0,j} B_{j,n}\right| \leq \sum_j |A_j||B_{n-j}| = O\Big(\sum_j Q_\alpha(j) Q_\beta(n-j)\Big) \ . \quad (4.17)$$

We consider only one half space $\|j\| \leq \|n-j\|$, where we bound $Q_\beta(n-j) \leq Q_\beta(n/2)$. As $Q_\alpha(j)$ is summable, the contribution of this half-plane is $O\big(Q_\beta(n/2)\big)$. The other half-plane gives the same result with $\alpha$ and $\beta$ interchanged, which proves the bound, while tightness follows by taking all $A_n$, $B_n$ positive. □

We now determine the asymptotics of $(\hat{\mathcal{E}}^{-1})_n$ for different types of Hamiltonians.

**Lemma 4.5.** *For non-critical systems with rapidly decaying interactions, i.e., as $o(\|n\|^{-\infty})$, the entries of $\hat{\mathcal{E}}^{-1}$ decay rapidly as well. That is,*

$$\Delta > 0 \ \Rightarrow \ (\hat{\mathcal{E}}^{-1})_n = o(\|n\|^{-\infty}) \ .$$

**Proof.** As the interactions decay as $o(\|n\|^{-\infty})$, $\hat{H}_\bullet \in \mathscr{C}^\infty(\mathcal{T}^d)$ ($\bullet = Q, P, PQ$), and thus $\mathcal{E}^2 = \hat{H}_Q \hat{H}_P + \hat{H}_{QP}^2 \in \mathscr{C}^\infty(\mathcal{T}^d)$. Since the system is gapped, i.e., $\mathcal{E} \geq \Delta > 0$, it follows that also $g := \mathcal{E}^{-1} \in \mathscr{C}^\infty(\mathcal{T}^d)$. For the proof, we need to bound

$$(\hat{\mathcal{E}}^{-1})_n = \frac{1}{(2\pi)^d} \int_{\mathcal{T}^d} \mathrm{d}\phi \, g(\phi) e^{in\phi}$$

by $\|n\|^{-\kappa}$ for all $\kappa \in \mathbb{N}$. First, let us have a look at the one-dimensional case. By integration by parts, we get

$$(\hat{\mathcal{E}}^{-1})_n = \frac{1}{2\pi} \left[\frac{1}{in} g(\phi) e^{in\phi}\right]_{\phi=-\pi}^{\pi} - \frac{1}{2\pi i n} \int_{-\pi}^{\pi} \mathrm{d}\phi \, g'(\phi) e^{in\phi} \ ,$$

where the first part vanishes due to the periodicity of $g$. As $g \in \mathscr{C}^\infty(\mathcal{T}^1)$, the integration by parts can be iterated arbitrarily often and all the brackets vanish, such that after $\kappa$ iterations,

$$(\hat{\mathcal{E}}^{-1})_n = \frac{1}{2\pi (in)^\kappa} \int_{-\pi}^{\pi} \mathrm{d}\phi \, g^{(\kappa)}(\phi) e^{in\phi} \ .$$

As $g^{(\kappa)}(\phi)$ is continuous, the integral can be bounded by $\int |g^{(\kappa)}(\phi)| \mathrm{d}\phi =: C_\kappa < \infty$, such that finally

$$|(\hat{\mathcal{E}}^{-1})_n| \leq \frac{C_\kappa}{n^\kappa} \quad \forall \kappa \in \mathbb{N} \, ,$$

which completes the proof of the one-dimensional case.

The extension to higher dimensions is straightforward. For a given $n = (n_1, \ldots, n_d)$, integrate by parts with respect to the $\phi_i$ for which $|n_i| = \|n\|_\infty$; we assume $i = 1$ without loss of generality. As $g(\cdot, \phi_2, \ldots, \phi_d) \in \mathscr{C}^\infty(\mathcal{S}^1)$, the same arguments as in the 1D case show

$$|(\hat{\mathcal{E}}^{-1})_n| \leq \frac{1}{(2\pi)^d |n_1|^\kappa} \int_{\mathcal{T}^d} \left| \frac{\partial^\kappa}{\partial \phi_1^\kappa} g(\phi) \right| \mathrm{d}\phi = \frac{C_\kappa}{\|n\|_\infty^\kappa} \ .$$

□

For systems with *local* interactions, a stronger version of Lemma 4.5 can be obtained:

**Lemma 4.6.** *For a system with finite range interaction, the entries of $\hat{\mathcal{E}}^{-1}$ decay exponentially.*

This has been proven in [78] for Hamiltonians of the type $H = V \oplus \mathbb{1}$, exploiting a result on functions of banded matrices [102]. Following Eqs. (4.9,4.11) the generalization to arbitrary translational invariant Hamiltonians is straightforward by replacing $V$ with $H_Q H_P - H_{QP}^2$. In fact, it has been shown recently that the result even extends to non translational invariant Hamiltonians of the form in Theorem 4.1b [90].

Finally, we consider systems with polynomially decaying interaction.

**Lemma 4.7.** *For a 1D lattice with $H = V \oplus \mathbb{1} > 0$ and an exactly polynomially decaying interaction*

$$V_{ij} = \begin{cases} i = j & : \quad a \\ i \neq j & : \quad \frac{b}{|i-j|^\nu} \end{cases} \ , \ 2 \leq \nu \in \mathbb{N} \ ,$$

*$\hat{\mathcal{E}}^{-1}$ decays polynomially with the same exponent, $(\hat{\mathcal{E}}^{-1})_n = (V^{1/2})_n = \Theta(|n|^{-\nu})$.*

Hamiltonians of this type appear, e.g., for the vibrational degrees of freedom of ions in a linear trap, where $\nu = 3$.

**Proof.** We need to estimate $(\hat{\mathcal{E}}^{-1})_n \overset{(4.9)}{=} (V^{-1/2})_n = \frac{1}{2\pi} \int_0^{2\pi} \hat{V}^{-1/2}(\phi) e^{in\phi} \mathrm{d}\phi$ . Note that

$$\hat{V}(\phi) = a + 2b \sum_{n=1}^{\infty} \frac{\cos(n\phi)}{n^\nu} = a + 2b \,\mathrm{Re}\left[\mathrm{Li}_\nu(e^{i\phi})\right] > 0 \ , \tag{4.18}$$

where $\mathrm{Li}_\nu(z) = \sum_{n \geq 1} z^n / n^\nu$ is the polylogarithm. The polynomial decay of coefficients implies $\hat{V} \in \mathscr{C}^{\nu-2}(\mathcal{S}^1)$, and as the system is non-critical, $\hat{V}^{-1/2} \in \mathscr{C}^{\nu-2}(\mathcal{S}^1)$. As $\mathrm{Li}_\nu$ has an analytic continuation to $\mathbb{C} \backslash [1; \infty)$, $\hat{V} \in \mathscr{C}^\infty((0; 2\pi))$ and thus $\hat{V}^{-1/2} \in \mathscr{C}^\infty((0; 2\pi))$. We can therefore integrate by parts $\nu - 1$ times, and as all brackets vanish due to periodicity, we obtain

$$(\hat{\mathcal{E}}^{-1})_n = \frac{1}{2\pi (in)^{\nu-1}} \int_0^{2\pi} \left[ \frac{\mathrm{d}^{\nu-1}}{\mathrm{d}\phi^{\nu-1}} \hat{V}^{-1/2}(\phi) \right] e^{in\phi} \mathrm{d}\phi \ , \tag{4.19}$$

and

$$\frac{\mathrm{d}^{\nu-1}}{\mathrm{d}\phi^{\nu-1}}\hat{V}^{-1/2}(\phi) = -\frac{\hat{V}^{(\nu-1)}(\phi)}{2\hat{V}(\phi)^{3/2}} + \frac{3(\nu-2)\hat{V}^{(\nu-2)}(\phi)\hat{V}^{(1)}(\phi)}{4\hat{V}(\phi)^{5/2}} + g(\phi) \ . \qquad (4.20)$$

Note that the second term only appears if $\nu \geq 3$, and $g$ only if $\nu \geq 4$. As $g(\phi) \in \mathscr{C}^1(\mathcal{S}^1)$, its Fourier coefficients vanish as $O(n^{-1})$, as can be shown by integrating by parts. The second term can be integrated by parts as well, the bracket vanishes due to continuity, and we remain with

$$\frac{1}{in}\int_0^{2\pi}\left[\frac{3(\nu-2)\hat{V}^{(\nu-1)}(\phi)\hat{V}^{(1)}(\phi)}{4\hat{V}(\phi)^{5/2}} + h(\phi)\right]e^{in\phi}\mathrm{d}\phi \ ,$$

with $h \in \mathscr{C}(\mathcal{S}^1)$. [For $\nu = 3$, a factor 2 appears as $(\hat{V}^{(1)})' = \hat{V}^{(\nu-1)}$.] As we will show later, $\hat{V}^{(\nu-1)}$ is absolutely integrable, hence the integral exists, and thus the Fourier coefficients of the second term in Eq. (4.20) vanish as $O(n^{-1})$ as well. Finally, it remains to bound

$$\int_0^{2\pi}\frac{\hat{V}^{(\nu-1)}(\phi)}{2\hat{V}(\phi)^{3/2}}e^{in\phi}\mathrm{d}\phi \ . \qquad (4.21)$$

As $\mathrm{Li}'_\nu(x) = \mathrm{Li}_{\nu-1}(x)/x$, it follows from Eq. (4.18) that

$$V^{(\nu-1)}(\phi) \ = \ 2b\,\mathrm{Re}\left[i^{\nu-1}\mathrm{Li}_1(e^{i\phi})\right] \ = \ 2b\,\mathrm{Re}\left[-i^{\nu-1}\log(1-e^{i\phi})\right] \ ,$$

where the last step is from the definition of $\mathrm{Li}_1$.

We now distinguish two cases. First, assume that $\nu$ is even. Then,

$$V^{(\nu-1)}(\phi) \propto \mathrm{Im}\log(1-e^{i\phi}) = \arg(1-e^{i\phi}) = (\phi-\pi)/2$$

on $(0;2\pi)$, hence the integrand in Eq. (4.21) is bounded and has a bounded derivative, and by integration by parts, the integral Eq. (4.21) is $O(n^{-1})$. In case $\nu$ is odd we have

$$V^{(\nu-1)}(\phi) \ \propto \ \mathrm{Re}\log(1-e^{i\phi}) \ = \ \log\left|1-e^{i\phi}\right| = \log(2\sin(\phi/2))$$

on $(0;2\pi)$. With $h(\phi) := \hat{V}^{-3/2}(\phi)/2$, the integrand in Eq. (4.21) can be written as

$$\hat{V}^{(\nu-1)}(\phi)h(\phi) \propto \log(2\sin(\phi/2))\ h(0) + \log(2\sin(\phi/2))\ [h(\phi)-h(0)] \ . \qquad (4.22)$$

The first term gives a contribution proportional to

$$\int_0^{2\pi}\log(2\sin(\phi/2))\cos(n\phi)\mathrm{d}\phi = -\frac{1}{2n}$$

as it is the back-transform of $-\frac{1}{2}\sum_{n\geq 1}\cos(n\phi)/n$. For the second term, note that $h \in \mathscr{C}^1(\mathcal{S}^1)$ for $\nu \geq 3$ and thus $h(\phi) - h(0) = h'(0)\phi + o(\phi)$ by Taylor's theorem. Therefore, the log singularity vanishes, and we can once more integrate by parts. The derivative is

$$\frac{1}{2}\cot(\phi/2)\,[h(\phi) - h(0)] + \log(2\sin(\phi/2))\,h'(\phi)\ .$$

In the left part, the $1/\phi$ singularity of $\cot(\phi/2)$ is cancelled out by $h(\phi) - h(0) = O(\phi)$, and the second part is integrable as $h' \in \mathscr{C}(\mathcal{S}^1)$, so that the contribution of the integral (4.21) is $O(n^{-1})$ as well.

In order to show that $n^{-\nu}$ is also a lower bound on $(\hat{V}^{-1/2})_n$, one has to analyze the asymptotics more carefully. Using the Riemann-Lebesgue lemma—which says that the Fourier coefficients of absolutely integrable functions are $o(1)$—one finds that all terms in (4.19) vanish as $o(1/n^{\nu})$, except for the integral (4.21). Now for even $\nu$, (4.21) can be integrated by parts, and while the brackets give a $\Theta(n^{-\nu})$ term, the remaining integral is $o(n^{-\nu})$, which proves that $(\hat{V}^{-1/2})_n = \Theta(n^{-\nu})$. For odd $\nu$, on the other hand, the first part of (4.22) gives exactly a polynomial decay, while the contributions from the second part vanishes as $o(n^{-\nu})$, which proves $(\hat{V}^{-1/2})_n = \Theta(n^{-\nu})$ for odd $\nu$ as well. □

**Generalizations of Lemma 4.7**

The preceding lemma can be extended to non-integer exponents $\alpha \notin \mathbb{N}$: If $V_n \propto n^{-\alpha}$, $n \neq 0$, then $(\hat{\mathcal{E}}^{-1})_n = O(n^{-\alpha})$.

For the proof, define $\alpha = \nu + \varepsilon$, $\nu \in \mathbb{N}$, $0 < \varepsilon < 1$. Then $\hat{V} \in \mathscr{C}^{\nu-1}(\mathcal{S}^1)$, $\hat{V} \in \mathscr{C}^{\infty}((0;2\pi))$, and one can integrate by parts $\nu$ times, where all brackets vanish. What remains is to bound the Fourier integral of the $\nu$'th derivative of $\hat{V}^{-1/2}$ by $n^{-\varepsilon}$. An upper bound can be established by noting that $|\hat{V}^{(\nu)}(\phi)| \leq |\mathrm{Li}_{\varepsilon}(e^{i\phi})| = O(\phi^{\varepsilon-1})$ and $|\hat{V}^{(\nu+1)}(\phi)| = O(\phi^{\varepsilon-2})$. It follows that all contributions in the Fourier integral except the singularity from $\hat{V}^{(\nu)}$ lead to $o(1/n)$ contributions as can be shown by another integration by parts. In order to bound the Fourier integral of the $O(\phi^{\varepsilon-1})$ term, split the Fourier integral at $\frac{1}{n}$. The integral over $[0;\frac{1}{n}]$ can be directly bounded by $n^{-\varepsilon}$, while for $[\frac{1}{n};1]$, an equivalent bound can be established after integration by parts, using $\hat{V}^{(\nu+1)} = O(\phi^{\varepsilon-2})$. This method is discussed in more detail in the proof of Theorem 4.15, following Eq. (4.44).

The proof that $n^{-\varepsilon}$ is also a lower bound to $(\hat{\mathcal{E}}^{-1})_n$ is more involved. From a series expansion of $\hat{V}$ and its derivatives, it can be seen that it suffices to bound the sine and cosine Fourier coefficients of $\phi^{\varepsilon-1}$ from below. As in the proof of Theorem 4.14, this is accomplished by splitting the integral into single oscillations of the sine or cosine and bounding each part by the derivative of $\phi^{\varepsilon-1}$.

For polynomially bounded interactions $V_n = O(n^{-\alpha})$, $\alpha > 1$, not very much can be said without further knowledge. With $\nu < \alpha$, $\nu \in \mathbb{N}$ the largest integer strictly smaller than $\alpha$, we know that $\hat{V} \in \mathscr{C}^{\nu-1}(\mathcal{S}^1)$. Thus, one can integrate

by parts $\nu - 1$ times, the brackets vanish, and the remaining Fourier integral is $o(1)$ using the Riemann-Lebesgue lemma. It follows that $(\hat{\mathcal{E}}^{-1})_n = o\big(n^{-(\nu-1)}\big)$. In contrast to the case of an exactly polynomial decay, this can be extended to higher spatial dimensions $d > 1$ by replacing $\nu - 1$ with $\nu - d$, which yields $(\hat{\mathcal{E}}^{-1})_n = o\big(n^{-(\nu-d)}\big)$.

We now use the preceding lemmas about the entries of $\hat{\mathcal{E}}^{-1}$ (Lemma 4.5–4.7) to derive corresponding results on the correlations of ground states of non-critical systems.

**Theorem 4.8.** *For systems with $\Delta > 0$, the following holds:*

*(i) If the Hamiltonian $H$ has finite range, the ground state correlations decay exponentially.*

*(ii) If $H$ decays as $o(\|n\|^{-\infty})$, the ground state correlations decay as $o(\|n\|^{-\infty})$ as well.*

*(iii) For a 1D system with $H = V \oplus \mathbb{1}$ where $V$ decays with a power law $|n|^{-\nu}$, $\nu \geq 2$, the ground state correlations decay as $\Theta(|n|^{-\nu})$.*

**Proof.** In all cases, we have to find the scaling of the ground state $\gamma$ which is the product $\gamma = (\hat{\mathcal{E}}^{-1} \oplus \hat{\mathcal{E}}^{-1})\sigma H \sigma^T$, Eq. (4.10). Part *(i)* follows directly from Lemma 4.6, as multiplying with a finite-range $\sigma H \sigma^T$ doesn't change the exponential decay, while *(ii)* follows from Lemma 4.5, the $o(\|n\|^{-\infty})$ decay of $\sigma H \sigma^T$, and Lemma 4.4. To show *(iii)*, note that for $H = V \oplus \mathbb{1}$, the ground state is $\gamma = V^{-1/2} \oplus V^{1/2}$, and from Lemma 4.7, $O(n^{-\nu})$ follows. For $\hat{V}^{-1/2}$, Lemma 4.7 also includes that the bound is exact, while for $\hat{V}^{1/2}$, it can be shown by transferring the proof of the lemma one-to-one. □

Note that a simple converse of Theorem 4.8 always holds: for each translationally invariant pure state CM $\gamma$, there exists a Hamiltonian $H$ with the same asymptotic behavior as $\gamma$ such that $\gamma$ is the ground state of $H$. This can be trivially seen by choosing $H = \sigma\gamma\sigma^T$.

## 4.5 Correlation length and gap

In this section, we consider one-dimensional chains with local gapped Hamiltonians. We compute the correlation length for these systems and use this result to derive a relation between correlation length and gap.

**Theorem 4.9.** *Consider a non-critical 1D chain with a local Hamiltonian. Define the complex extension of the spectral function*

$$\mathcal{E}(\phi) = \Big[\sum_{n=0}^{L} c_n \cos(n\phi)\Big]^{1/2}$$

*in Eq. (4.16) as*

$$g(z) := \sum_{n=0}^{L} c_n \frac{z^n + z^{-n}}{2} \ ,$$

*such that*

$$g(e^{i\phi}) = \mathcal{E}^2(\phi) \overset{(4.9)}{=} \hat{H}_Q(\phi)\hat{H}_P(\phi) - \hat{H}_{QP}^2(\phi) \ ,$$

*and let $\tilde{z}$ be zero of $g$ with the largest magnitude smaller than one. Then, the correlation length*

$$\xi = -\frac{1}{\log|\tilde{z}|}$$

*determines the asymptotic scaling of the correlations which is given by*

- $O^*(e^{-n/\xi}/\sqrt{n})$*, if $\tilde{z}$ is a zero of order one,*
- $O^*(e^{-n/\xi})$*, if $\tilde{z}$ is a zero of even order,*
- $o(e^{-n/(\xi+\varepsilon)})$ *for all $\varepsilon > 0$, if $\tilde{z}$ is a zero of odd order larger than one.*

For the nearest neighbor interaction Hamiltonian $\mathcal{H}_\kappa$ from Eq. (4.1) one has for instance $\mathcal{E}(\phi) = \sqrt{1 - \kappa\cos(\phi)}$, so that $g$ has simple zeros at $z_0 = (1 \pm \sqrt{1-\kappa^2})/\kappa$. Therefore $\tilde{z} = (1 - \sqrt{1-\kappa^2})/\kappa$, and the correlations decay as $\Theta(e^{-n/\xi}/\sqrt{n})$ where $\xi = -1/\log|\tilde{z}|$.

**Proof.** For local Hamiltonians, the correlations decay as the matrix elements of $\hat{\mathcal{E}}^{-1}$ [Eq. (4.10)]. By Fourier transforming (4.9), we have that

$$\mathcal{E}(\phi) = \sqrt{g(e^{i\phi})} \ ,$$

with

$$g(e^{i\phi}) = \hat{H}_Q(\phi)\hat{H}_P(\phi) - \hat{H}_{QP}^2(\phi) = \sum_{n=0}^{L} c_n \cos(n\phi)$$

an even trigonometric polynomial (we assume $c_L \neq 0$ without loss of generality), and $\min(g(e^{i\phi})) = \Delta^2$. We have to compute

$$(\hat{\mathcal{E}}^{-1})_n \ = \ \frac{1}{2\pi}\int_0^{2\pi} \frac{1}{\mathcal{E}(\phi)} e^{in\phi} \mathrm{d}\phi \ = \ \frac{1}{2\pi i}\int_{\mathcal{S}^1} \frac{z^{n-1}}{\sqrt{g(z)}} \, \mathrm{d}z \ , \tag{4.23}$$

where $\mathcal{S}^1$ is the unit circle. The function $g(z)$ has a pole of order $L$ at zero and $2L$ zeros altogether. Since $\min(g(\phi)) = \Delta^2 > 0$, $g$ has no zeros on the unit circle. As $g(z) = g(1/z)$, the zeros come in pairs, and $L$ of them are inside the unit circle. Also, the conjugate of a zero is a zero as well. From each zero with odd multiplicity emerges a branch cut of $\sqrt{g(z)}$. We arrange all the branch cuts inside the unit circle such that they go straight to the middle where they annihilate with another cut. In case $L$ is odd, the last cut is annihilated by the

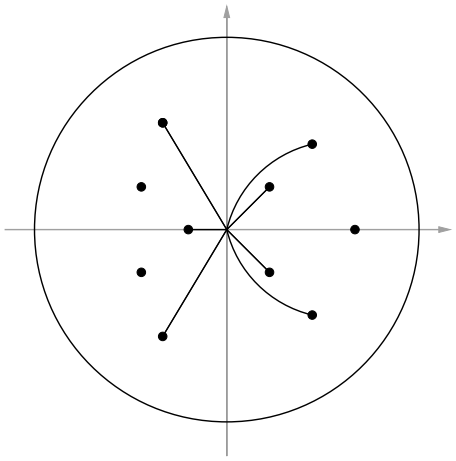

**Figure 4.1:** Sample arrangement of branch cuts and poles of $\sqrt{g}$ inside the unit circle. From each odd order zero of $g$, a branch cut emerges. All cuts go to 0 where they cancel with another cut. In case their number is odd, there is an additional branch point at 0 cancelling the last cut. In case two zeros are on a line to the origin, the cuts are chosen curved. The integral of $\sqrt{g}$ around the unit circle is equal to the integral around the cuts, plus integrals around the residues which originate from the even order zeros of $g$.

singularity of $\sqrt{g(z)}$ at 0. If two zeros lie on a line, one cut curves slightly. A sample arrangement is shown in Fig. 4.1.

Following Cauchy's theorem, the integral can be decomposed into integrals along the different branch cuts and around the residues of $1/\sqrt{g}$, and one has to estimate the contributions from the different types of zeros of $g$. The simplest case is given by zeros $z_0$ with even multiplicity $2m$. In that case, define $h(z) := g(z)/(z-z_0)^{2m}$ which has no zero around $z_0$. The contribution from $z_0$ to the correlations is then given by the residue at $z_0$ and is

$$\frac{1}{(m-1)!}\,\frac{\mathrm{d}^{m-1}}{\mathrm{d}z^{m-1}}\left(\frac{z^{n-1}}{\sqrt{h(z)}}\right)\bigg|_{z=z_0} \propto z_0^{n-(m-1)}$$

for $n-(m-1)>0$, i.e., it scales as $|z_0|^n$. Note that for $z_0 \notin \mathbb{R}$, the imaginary parts originating from $z_0$ and its conjugate $\bar{z}_0$ exactly cancel out, but the scaling is still given by $|z_0|^n = e^{n\log|z_0|}$, i.e., $\xi = -1/\log|z_0|$ is the corresponding correlation length.

If $z_0$ is a simple zero of $g(z)$, we have to integrate around the branch cut. Assume first that the cut goes to zero in a straight line, and consider a contour with distance $\varepsilon$ to the slit. Both the contribution from the $\varepsilon$ region around zero and the $\varepsilon$ semicircle at $z_0$ vanish as $\varepsilon \to 0$, and the total integral is therefore given by twice the integral along the cut,

$$\frac{1}{\pi i}\int_0^{z_0} \frac{z^{n-1}}{\sqrt{z-z_0}\sqrt{h(z)}}\mathrm{d}z \ ,$$

where again $h(z) = g(z)/(z-z_0)$. Intuitively, for growing $n$ the part of the integral close to $z_0$ becomes more and more dominating, i.e., the integral is well approximated by the modified integral where $h(z)$ has been replaced by $h(z_0)$.

After rotating it onto the real axis, this integral—up to a phase—reads

$$\frac{1}{\pi\sqrt{|h(z_0)|}}\int_0^{|z_0|}\frac{r^{n-1}}{\sqrt{|z_0|-r}}\mathrm{d}r = \frac{|z_0|^{n-1/2}\Gamma(n)}{\sqrt{\pi|h(z_0)|}\,\Gamma(n+\frac{1}{2})} \tag{4.24}$$

which for large $n$ is

$$\frac{1}{\sqrt{\pi|z_0 h(z_0)|}}\,\frac{|z_0|^n}{\sqrt{n}} + O\left(\frac{|z_0|^n}{n^{3/2}}\right)\ . \tag{4.25}$$

In order to justify the approximation $h(z) \rightsquigarrow h(z_0)$, consider the difference of the two respective integrals. It is bounded by

$$\left|\int_0^{z_0}\frac{|z|^{n-1}}{\sqrt{|z-z_0|}}\underbrace{\left|\frac{1}{\sqrt{h(z)}}-\frac{1}{\sqrt{h(z_0)}}\right|}_{(*)}\mathrm{d}z\right|\ .$$

On $[z_0/2, z_0]$, $h(z)$ is analytic and has no zeros, thus, $|h(z)^{-1/2} - h(z_0)^{-1/2}| < C|z-z_0|$, where $C$ is the maximum of the derivative of $h(z)^{-1/2}$ on $[z_0/2, z_0]$. On $[0, z_0/2]$, the same bound is obtained by choosing $C$ the supremum of $|h(z)^{-1/2} - h(z_0)^{-1/2}|/|z_0/2|$ on $[0, z_0/2]$. Together, $(*) \le C|z-z_0|$, and the above integral is bounded by

$$C\int_0^{|z_0|} r^{n-1}\sqrt{|z_0|-r}\,\mathrm{d}r = C\frac{\sqrt{\pi}|z_0|^{n+1/2}\Gamma(n)}{2\Gamma(n+\frac{3}{2})} = O\left(\frac{|z_0|^n}{n^{3/2}}\right)\ ,$$

i.e., it vanishes by $1/n$ faster than the asymptotics derived in Eq. (4.25), which justifies fixing $h(z)$ at $h(z_0)$.

From Eq. (4.25), it follows that the scaling is $e^{-n/\xi}/\sqrt{n}$, where the correlation length is again $\xi = -1/\log|z_0|$. The same scaling behavior can be shown to hold for appropriately chosen curved branch cuts from $z_0$ to 0 by relating the curved to a straight integral.

The situation gets more complicated if zeros of odd order $> 1$ appear. In order to get an estimate which holds in all scenarios, we apply Cauchy's theorem to contract the unit circle in the integration (4.23) to a circle of radius $r > |z_0|$, where $z_0$ is the largest zero inside the unit circle. Then, the integrand can be bounded by $C_r r^{n-1}$ (where $C_r < \infty$ is the supremum of $1/\sqrt{g}$ on the circle), and this gives a bound $2\pi C_r r^{n-1}$ for the integral. This holds for all $r > |z_0|$, i.e., the correlation decay faster than $e^{n\log r}$ for all $r > |z_0|$. This does not imply that the correlations decay as $e^{n\log|z_0|}$, but it is still reasonable to define $-1/\log|z_0|$ as the correlation length. □

**Theorem 4.10.** *Consider a 1D chain together with a family of Hamiltonians $H(\Delta)$ with gap $\Delta > 0$, where $H(\Delta)$ is continuous for $\Delta \to 0$ in the sense that all*

*entries of $H$ converge. Then, the ground state correlations scale exponentially, and for sufficiently small $\Delta$ the correlation length is*

$$\xi \simeq \frac{1}{\sqrt{\Delta m^*}} \ .$$

*Here, $m^* = \left( \frac{\mathrm{d}^2 \mathcal{E}(\phi)}{\mathrm{d}\phi^2} \Big|_{\phi=\phi_\Delta} \right)^{-1}$ is the effective mass at the band gap.*

For the discretized Klein-Gordon field (4.1), e.g., we have $\Delta = \sqrt{1-|\kappa|}$ and $m^* = 2\sqrt{1-|\kappa|}/|\kappa|$, and for small $\Delta$ (corresponding to $|\kappa|$ close to 1), one obtains

$$\xi \simeq \sqrt{\frac{|\kappa|}{2(1-|\kappa|)}} \simeq \frac{1}{\sqrt{2}\Delta} \ .$$

Hence, the $\xi \propto 1/\Delta$ law holds if the coupling is increased relative to the on-site energy (in which case $m^* \propto \Delta$).

More generally, if we expand the spectral function [Eq. (4.16)] around the band gap we are generically[3] led to the dispersion relation $\mathcal{E}(k) \simeq \sqrt{\Delta^2 + v^2 k^2}$ ($k \equiv \phi$). By the definition of the effective mass and Theorem 4.10 this leads exactly to the folk theorem

$$\xi \simeq \frac{v}{\Delta} \ . \tag{4.26}$$

**Proof.** According to Theorem 4.9, what remains to be done is to determine the position of the largest zero $\tilde{z}$ of $g$ in the unit circle. Due to the restriction on $H(\Delta)$, the coefficients of the polynomial $g(z)z^L$ and thus also the zeros of $g$ continuously depend on $\Delta$, i.e., for sufficiently small $\Delta$, the zero closest to the unit circle is the one closest to the gap. In order to determine the position of this zero, we will expand $g$ around the gap. We only discuss the generic case where the gap appears only for one angle $\phi_0$, $g(\phi_0) = \Delta$. In the case of multiple occurrences of the gap in the spectrum, one will pick the gap which gives the zero closest to the unit circle, i.e., the largest correlation length. Furthermore, we assume $\phi_0 = 0$ without loss of generality. Otherwise, one considers $g(ze^{-i\phi_0})$ instead of $g(z)$, which on the unit circle coincides with the (rotated) spectrum.

The knowledge on $g =: u + iv$ (with $u, v : \mathbb{C} \to \mathbb{R}$) which will be used in the proof is

$$\begin{aligned} u(1) &= \Delta^2 & v(1) &= 0 \\ u_\phi(1) &= 0 & v_\phi(1) &= 0 \\ u_{\phi\phi}(1) &= 2\Delta/m^* > 0 & v_{\phi\phi}(1) &= 0 \end{aligned} \tag{4.27}$$

[3]This makes the natural assumption that the minimum under the square root is quadratic. In fact, if it is of higher order, then $m^* = \infty$ and thus $\xi = 0$, which is consistent with the findings of the following section. An example of such a behavior is given by so called 'quadratic interactions' [79] for which $H = V \oplus \mathbb{1}$, where $V$ is the square of a banded matrix.

where the subscripts denote the partial derivative with respect to the respective subscript (in Euclidean coordinates $z \equiv x + iy$, in polar coordinates $z \equiv re^{i\phi}$). Note that $z = 1$ is the point where the gap appears, and that $g(e^{i\phi}) = \mathcal{E}(\phi)^2$ is real. Therefore, the derivatives of the imaginary part $v$ along the circle vanish, while the derivatives of the real part $u$ are found to be $u(1) = \mathcal{E}(0)^2 = \Delta^2$, $u_\phi(1) = 2\mathcal{E}(0)\mathcal{E}'(0) = 0$, and $u_{\phi\phi}(0) = 2\mathcal{E}'(0)^2 + 2\mathcal{E}(0)\mathcal{E}''(0) = 2\Delta/m^*$, where $m^* = 1/\mathcal{E}''(\phi)$ is the effective mass at the band gap.

We need to exploit the relation between Euclidean and polar coordinates,

$$\begin{aligned} g_x(1) &= g_r(1) \quad ; \quad g_y(1) = g_\phi(1) \\ g_{xx}(1) &= g_{rr}(1) \; ; \quad g_{yy}(1) = g_{\phi\phi}(1) + g_r(1) \end{aligned}$$

and the Cauchy-Riemann equations $u_x = v_y$, $u_y = -v_x$, and $g_{xx} + g_{yy} = 0$, which together with the information (4.27) lead to

$$\begin{gathered} u(1) = \Delta^2 \; ; \; v(1) = 0 \; ; \\ u_x(1) = u_y(1) = v_x(1) = v_y(1) = 0 \; ; \\ u_{xx}(1) = -2\Delta/m^* \; ; \; u_{yy}(1) = 2\Delta/m^* \; ; \\ v_{xx}(1) = 0 \; ; \; v_{yy}(1) = 0 \; . \end{gathered}$$

Note that it is not possible to derive information about the mixed second derivates using only the information (4.27). However, as long as $v_{xy}$ does not vanish at 1, $v$ will only stay zero in direction of $x$ or $y$, but not diagonally. Since $\Delta^2 > 0$ and $2\Delta/m^* > 0$, the closest zero is—to second order—approximately located along the $x$ axis. By intersecting with the parabola $\Delta^2 - \frac{\Delta}{m^*}(x-1)^2$, one finds that the zero is located at $x_0 \approx 1 - \sqrt{\Delta m^*}$. For small $\Delta$, the correlations thus decay with correlation length $\xi \approx -1/\log(1 - \sqrt{\Delta m^*}) \approx 1/\sqrt{\Delta m^*}$. □

## 4.6 Critical systems

We now turn towards critical systems, i.e., systems without an energy gap, $\Delta = 0$.[4] In that case, the Hamiltonian will get singular and some entries of the ground state covariance matrix will diverge, which leads to difficulties and ambiguities in the description of the asymptotic behavior of correlations. We will therefore restrict to Hamiltonians of the type

$$H = V \oplus \mathbb{1} \; ,$$

for which the ground state CM is $\gamma = V^{-1/2} \oplus V^{1/2}$. While the $Q$ part diverges, the entries of the $P$-block stay finite. Following Theorem 4.1b the extension to interactions of the form $H = H_Q \oplus H_P$ is straight forward.

[4]Note that there are different meanings of the notion criticality referring either to a vanishing energy gap or to an algebraic decay of correlations. In this section we discuss in which cases these two properties are equivalent.

In order to compute the correlations we have to determine the asymptotics of $V^{1/2}$, i.e.,

$$(V^{1/2})_n = \frac{1}{(2\pi)^d} \int_{\mathcal{T}^d} \sqrt{\hat{V}(\phi)} e^{in\phi} \mathrm{d}\phi \ .$$

We will restrict to the cases in which the excitation spectrum $\mathcal{E} = \sqrt{\hat{V}}$ has only a finite number of zeros, i.e., finitely many points of criticality. In addition, we will also consider the special case in which the Hamiltonian exhibits a tensor product structure.

We proceed as follows. First, we consider one-dimensional critical chains and show that the correlations decay typically as $O(n^{-2})$ and characterize those special cases where the correlations decay more rapidly. The practically important case of exactly cubic decaying interactions will be investigated in greater detail. Depending on the sign of the interaction this case will lead to a logarithmic deviation from the $n^{-2}$ behavior. Then, we turn to higher dimensional systems and show that generically the correlations decay as $n^{-(d+1)} \log n$, where $d$ is the spatial dimension of the lattice.

### 4.6.1 One dimension

First, we prove a lemma which shows that although taking the square root of a smooth function destroys its differentiability, the derivatives will stay bounded.

**Lemma 4.11.** *Let $f \in \mathscr{C}^m([-1;1])$, $f(x) \geq 0$ with the only zero at $x = 0$, and let $2\nu \leq m$ be the order of the minimum at $x = 0$, i.e., $f^{(k)}(0) = 0 \ \forall k < 2\nu$, $f^{(2\nu)}(0) > 0$.*

*Define $g(x) := \sqrt{f(x)}$. Then, the following holds:*

- *For odd $\nu$, $g \in \mathscr{C}^{\nu-1}([-1;1])$, and $g \in \mathscr{C}^{m-\nu}([-1;0])$, $g \in \mathscr{C}^{m-\nu}([0;1])$, i.e., the first $m - \nu$ derivatives (for $x \neq 0$) are bounded.*
- *For even $\nu$, $g \in \mathscr{C}^{m-\nu}([-1;1])$.*

**Proof.** Using the Taylor expansion $f(x) = \sum_{k=2\nu}^{m} c_k x^k + \rho(x)$, $\rho^{(k)}(x) = o(x^{m-k})$ for $k \leq m$, we express $g$ as $g(x) = (\operatorname{sgn} x)^\nu x^\nu r(x)$ with

$$r(x) = \sqrt{\sum_{k=2\nu}^{m} c_k x^{k-2\nu} + \frac{\rho(x)}{x^{2\nu}}} \ ,$$

where we used that $(\operatorname{sgn} x)^\nu x^\nu = \sqrt{x^{2\nu}}$. Let us now consider the derivatives of $r(x)$. While the sum leads to a $O(1)$ contribution, the $k$'th derivative of the remainder behaves as $o(1)/x^{2\nu-m+k}$. Together, this leads to

$$\begin{aligned} r^{(k)}(x) &= O(1) && 2\nu - m + k \leq 0 \ , \\ r^{(k)}(x) &= o(1)/x^{2\nu-m+k} && 2\nu - m + k \geq 1 \ . \end{aligned}$$

Now consider the $k$'th derivative of $g(x)$ for $x \neq 0$,

$$g^{(k)}(x) = (\operatorname{sgn} x)^\nu \sum_{l=0}^{k} \underbrace{\binom{k}{l} \left[\frac{\mathrm{d}^l}{\mathrm{d}x^l} x^\nu\right] r^{(k-l)}(x)}_{s_l} \ .$$

Assume first that $k \leq \nu$. Then, $s_l \propto O(1)x^{\nu-l}$ for $2\nu - m + k - l \leq 0$, and $s_l \propto o(1)x^{m-\nu-k}$ for $2\nu - m + k - l \geq 1$, and as $m \geq 2\nu$, it follows that $g^{(k)} = O(x)$ for $k < \nu$, which cancels the discontinuity originating from $\operatorname{sgn} x$. For $k = \nu$, on the contrary, $s_k = O(1)$, and $\operatorname{sgn} x$ introduces a discontinuity on $g^{(k)}$, yet, it remains bounded and piecewise differentiable on $[-1; 0]$ and $[0; 1]$. The first non-bounded $s_l$ is found as soon as $m - \nu - k = -1$, and $g \in \mathscr{C}^{m-\nu}([0; 1])$ directly follows.

This also implies that for $m - \nu - k \geq 0$, $g(x)/(\operatorname{sgn} x)^\nu \in \mathscr{C}^{m-\nu}([-1; 1])$, i.e., the discontinuity is only due to $(\operatorname{sgn} x)^\nu$. Since, however, this is only discontinuous for odd $\nu$, it follows that $g \in \mathscr{C}^{m-\nu}([-1; 1])$ if $\nu$ even. □

**Theorem 4.12.** *Consider a one-dimensional critical chain with Hamiltonian $H = V \oplus \mathbb{1}$, where $V_n = O(n^{-\alpha})$, $\alpha > 4$ and where $\hat{V}$ has a finite number of critical points which are all quadratic minima of $\hat{V}$. Then, $(\gamma_P)_n = O^*(n^{-2})$. For $V_n \propto n^{-\alpha}$, $\alpha > 3$ it even follows that $(\gamma_P) = \Theta(n^{-2})$.*

Note that for $V_n \propto n^{-\alpha}$, the extrema of $\hat{V}$ are always quadratic.

**Proof.** We want to estimate

$$(V^{1/2})_n = \frac{1}{2\pi} \int_{\mathcal{S}^1} g(\phi) e^{in\phi} \mathrm{d}\phi \ , \tag{4.28}$$

where $g = \hat{V}^{1/2}$. Under both assumptions, $\hat{V} \in \mathscr{C}^2(\mathcal{S}^1)$, and all critical points are minima of order 2. It follows from Lemma 4.11 that $g$ is continuous with bounded derivative. Therefore, we can integrate by parts, the bracket vanishes, and we obtain

$$(V^{1/2})_n = -\frac{1}{2\pi i n} \int_0^{2\pi} g'(\phi) e^{in\phi} \mathrm{d}\phi \ .$$

Now, split $\mathcal{S}^1$ at the zeros of $g$ into closed intervals $\mathcal{I}_j$, $\bigcup_j \mathcal{I}_j = \mathcal{S}^1$, and rewrite the above integral as a sum of integrals over $\mathcal{I}_j$. As $g' \in \mathscr{C}(\mathcal{I}_j)$ (and differentiable on the inner of $\mathcal{I}_j$), one can once more integrate by parts which yields

$$(V^{1/2})_n = -\frac{1}{2\pi(in)^2} \sum_j \left( \left[g'(\phi) e^{in\phi}\right]_{\mathcal{I}_j} - \int_{\mathcal{I}_j} g''(\phi) e^{in\phi} \mathrm{d}\phi \right) . \tag{4.29}$$

Neither of the terms will vanish, but since $g' \in \mathscr{C}(\mathcal{I}_j)$, the bracket is bounded. In case $V_n \in O(n^{-\alpha})$, $\alpha > 4$, we have $\hat{V} \in \mathscr{C}^3(\mathcal{S}^1)$, therefore $g''$ is bounded

(Lemma 4.11), and the integrals vanish as $o(1)$. Unless the contributions of the brackets for the different $\mathcal{I}_j$ cancel out, the $n^{-2}$ bound is tight, $(V^{1/2})_n = O^*(n^{-2})$. The tightness of the bound is also illustrated by the example which follows the proof.

For the case of an exactly polynomial decay, we additionally have to show that $g''$ is absolutely integrable for $3 < \alpha \leq 4$. Then, the exactness of the bound holds because the bracket in Eq. (4.29) does not oscillate (the critical point is either at $\phi = 0$ or at $\phi = \pi$), and because the integral is $o(1)$ for $g'' \in \mathcal{L}^1(\mathcal{S}^1)$. In case the critical point is at $\phi = \pi$, the latter holds since $\hat{V} \in \mathscr{C}^\infty((0; 2\pi))$ implies that $g''$ is bounded at $\pi$, and $\hat{V} \in \mathscr{C}^2(\mathcal{S}^1)$ that $g \in \mathscr{C}^2((-\pi, \pi))$, which together proves that $g''$ is bounded on $\mathcal{S}^1$.

In case the critical point is at $\phi = 0$, the situation is more involved (and for $\alpha = 3$, a logarithmic correction appears, cf. Theorem 4.14). Since

$$\hat{V}^{(3)}(\phi) = -\mathrm{Im}\big[\mathrm{Li}_{\alpha-3}(e^{i\phi})\big] = O(\phi^{\alpha-4}) \ ,$$

we have

$$\begin{aligned} \hat{V}''(\phi) &= \hat{V}''(0) + O(\phi^{\alpha-3}) \ , \\ \hat{V}'(\phi) &= \hat{V}''(0)\phi + O(\phi^{\alpha-2}) \ , \\ \hat{V}(\phi) &= \tfrac{1}{2}\hat{V}''(0)\phi^2 + O(\phi^{\alpha-1}) \ . \end{aligned}$$

With this information,

$$g''(\phi) = \frac{2\hat{V}(\phi)\hat{V}''(\phi) - \hat{V}'(\phi)^2}{4V(\phi)^{3/2}} = O(\phi^{\alpha-4}) \ ,$$

which indeed proves that $g'' \in \mathcal{L}^1(\mathcal{S}^1)$, and thus $(V^{1/2})_n = \Theta(n^{-2})$. □

As an example, consider again the discretized Klein-Gordon field of Eq. (4.1) which is critical for $\kappa = \pm 1$, corresponding to $\hat{V}(\phi) = 1 \mp \cos\phi$. The Fourier integral is solvable and yields

$$(\gamma_P)_n = -\frac{2\sqrt{2}}{\pi}\frac{(\mathrm{sgn}\,\kappa)^n}{4n^2 - 1} = \Theta(n^{-2}) \ .$$

**Generalizations of Theorem 4.12**

Using Lemma 4.11, several generalizations for the 1D critical case can be found. In the following, we mention some of them. In all cases $H = V \oplus \mathbb{1}$ is critical.
*Critical points of even order.*—If $V_n = o(n^{-\infty})$ and the critical points are minima of order $2\nu$, $\nu$ even, the correlations decay as $(\gamma_P)_n = o(n^{-\infty})$. This is the case, e.g., if $V = X^2$ with $X$ itself rapidly decaying.
*Critical points of higher order.*—If $\hat{V}$ has critical points of order at least $2\nu$, $\nu$

odd, and $V_n = O(n^{-\alpha})$, $\alpha > 2\nu + 2$, then $(\gamma_P)_n = O(n^{-(\nu+1)})$.
*Minima of different orders.*—If $\hat{V}$ has minima of different orders $2\nu_i$, in general the minimum with the lowest odd $\nu_i \equiv \nu_1$ will determine the asymptotics, $(\gamma_P)_n = O(n^{-(\nu_1+1)})$. As $\hat{V} \in \mathscr{C}^{(2\max\{\nu_i\})}(\mathcal{S}^1)$ is required anyway, the piecewise differentiability of $\hat{V}^{1/2}$ is guaranteed.
*Weaker requirements on $V$.*—It is possible to ease the requirements imposed on $V$ in Theorem 4.12 to $V_n = O(n^{-\alpha})$, $\alpha > 3$ or $\hat{V} \in \mathscr{C}^2(\mathcal{S}^1)$, respectively. The price one has to pay is that one gets an additional log correction as in the multidimensional critical case, Theorem 4.15. The method to bound $g''$ is the same which is used there to derive (4.39).

The above proof does not cover the case of the relevant $1/n^3$ interaction, which for instance appears for the motional degrees of freedom of trapped ions. In the following, we separately discuss this case. It will turn out that the scaling will depend on the sign of the coupling: while a positive sign (corresponding to the radial degrees of freedom) again gives a $\Theta(\frac{1}{n^2})$ scaling as before, for the negative sign (corresponding to the axial degree of freedom) one gets $\Theta\big(\frac{\sqrt{\log n}}{n^2}\big)$.

**Theorem 4.13.** *Consider a critical 1D chain with a $1/n^3$ coupling with positive sign, i.e., $H = V \oplus \mathbb{1}$, $V_n = c/n^3$, $V_0 = 3c\zeta(3)/2$, $c > 0$, with $\zeta$ the Riemann zeta function. Then, the ground state correlations scale as $(\gamma_P)_n = \Theta(\frac{1}{n^2})$.*

**Proof.** We take w.l.o.g. $c = \frac{1}{2}$. For this sign of the coupling, the critical point is at $\pi$, $\hat{V}(\pi) = 0$. From the proof of Lemma 4.7, we know that $\hat{V} \in \mathscr{C}^1(\mathcal{S}^1)$, $\hat{V} \in \mathscr{C}^\infty((0;2\pi))$, and that $\hat{V}''(\phi) = \log(2\sin(\phi/2))$ on $(0;2\pi)$. With $g := \hat{V}^{1/2}$, it follows from Lemma 4.11 that $g \in \mathscr{C}(\mathcal{S}^1)$, $g \in \mathscr{C}^1([-\pi;\pi])$, and $g \in \mathscr{C}^\infty((0;\pi])$, $g \in \mathscr{C}^\infty([-\pi;0))$. This means that all derivatives $g^{(k)}$, $k \geq 1$ can exhibit jumps at the critical point $\pi$ but they all remain bounded. In contrast, around $\phi = 0$, $g'$ is continuous but $g''$ has a log divergence.

Thus, the Fourier integral

$$(V^{1/2})_n = \frac{1}{2\pi} \int_{\mathcal{S}^1} g(\phi) e^{in\phi} \mathrm{d}\phi$$

can be split at 0 and $\pi$, and then integrated by parts twice. The brackets of the first integration cancel out due to continuity of $g$, and one remains with

$$(V^{1/2})_n = \frac{1}{\pi(in)^2} \left( [g'(\phi)\cos(n\phi)]_0^\pi + \int_0^\pi g''(\phi)\cos(n\phi)\mathrm{d}\phi \right) ,$$

where we used the symmetry of $g$. One finds $[g'(\phi)\cos(n\phi)]_0^\pi = -\sqrt{\frac{\log 2}{2}}(-1)^n$, and since $g''$ is integrable, the integral is $o(1)$ due to the Riemann-Lebesgue lemma. Together, this proves $(\gamma_P)_n = \Theta(\frac{1}{n^2})$. □

**Theorem 4.14.** *Consider a critical 1D chain with a $1/n^3$ coupling with negative sign, i.e., $H = V \oplus \mathbb{1}$, $V_n = -c/n^3$, $V_0 = 2c\zeta(3)$, $c > 0$, with $\zeta$ the Riemann zeta function. Then, the ground state correlations scale as $(\gamma_P)_n = \Theta\big(\frac{\sqrt{\log n}}{n^2}\big)$.*

**Proof.** Again, take w.l.o.g. $c = \frac{1}{2}$. For the negative sign of the interaction, the critical point is at $\phi = 0$. Since at this point $\hat{V}''$ diverges, Lemma 4.11 cannot be applied, and the situation gets more involved.

As in the previous proof, we use that $\hat{V} \in \mathscr{C}^1(\mathcal{S}^1)$, $\hat{V} \in \mathscr{C}^\infty((0;2\pi))$, and thus $\hat{V}^{1/2} \in \mathscr{C}(\mathcal{S}^1)$, $\hat{V}^{1/2} \in \mathscr{C}^\infty((0;2\pi))$. Further, $\hat{V}''(\phi) = -\log(2\sin(\phi/2))$ on $(0;2\pi)$, cf. the proof of Lemma 4.7, and with $\sin x = x(1 + O(x^2))$ we have

$$\hat{V}''(\phi) = -\log(\phi) + O(\phi^2)$$

for $\phi \to 0$ (and similarly for $\phi \to 2\pi$), and therefore

$$\begin{aligned} \hat{V}'(\phi) &= \phi(1 - \log\phi) + O(\phi^3) \ , \\ \hat{V}(\phi) &= \tfrac{1}{4}\phi^2(3 - 2\log\phi) + O(\phi^4) \ . \end{aligned} \tag{4.30}$$

As $\hat{V}^{1/2} \in \mathscr{C}(\mathcal{S}^1)$, we can integrate by parts one time,

$$(V^{1/2})_n = \frac{1}{2\pi}\int_{\mathcal{S}^1} \hat{V}^{1/2}(\phi)e^{in\phi}\mathrm{d}\phi = \frac{1}{\pi n}\int_0^\pi g'(\phi)\sin(n\phi)\mathrm{d}\phi \tag{4.31}$$

where we exploited the symmetry of $\hat{V}$, and with $g := \hat{V}^{1/2}$. Then, from (4.30),

$$\begin{aligned} g'(\phi) &= \frac{1 - \log\phi}{\sqrt{3 - 2\log\phi}} + O\bigg(\frac{\phi^2}{\sqrt{|\log\phi|}}\bigg) \ , \\ g''(\phi) &= \frac{-2 + \log\phi}{\phi(3 - 2\log\phi)^{3/2}} + O\bigg(\frac{\phi}{\sqrt{|\log\phi|}}\bigg) \ , \end{aligned}$$

and after another round of approximation,

$$\begin{aligned} g'(\phi) &= \frac{\sqrt{|\log\phi|}}{\sqrt{2}} + O\bigg(\frac{1}{\sqrt{|\log\phi|}}\bigg) \ , \\ g''(\phi) &= -\frac{1}{2^{3/2}}\frac{1}{\phi\sqrt{|\log\phi|}} + O\bigg(\frac{1}{\phi|\log\phi|^{3/2}}\bigg) \ . \end{aligned}$$

This shows that the remainder $g'(\phi) - \sqrt{|\log\phi|/2}$ is continuous with an absolutely integrable derivative, and by integration by parts it follows that it only leads to a contribution $O(1/n)$ in the integral (4.31). Thus, it remains to investigate the asymptotics of the sine Fourier coefficients of $h(\phi) = \sqrt{|\log\phi|}$. For convenience, we split the integral (4.31) at 1, and $[1;\pi]$ only contributes with $O(1/n)$, as $h$ is

continuous with absolutely integrable derivative on $[1;\pi]$. On $[0;1]$, we have to compute the asymptotics of

$$\mathcal{I} = \int_0^1 \sqrt{-\log\phi}\sin(n\phi)\mathrm{d}\phi \ . \tag{4.32}$$

Therefore, split the integral at $1/n$. The left integral can be bounded directly, and the right after integration by parts [cf. the treatment of Eq. (4.44)]. One gets

$$\mathcal{I} \leq \int_0^{1/n} \sqrt{-\log\phi}\,\mathrm{d}\phi + \frac{\sqrt{\log n}}{n} + \frac{1}{n}\int_{1/n}^1 \frac{1}{2\phi\sqrt{-\log\phi}}\mathrm{d}\phi = O\left(\frac{\sqrt{\log n}}{n}\right) \ .$$

In order to prove that this is also a lower bound for the asymptotics, it suffices to show this for the integral (4.32) as all other contributions vanish more quickly. To this end, split the integral (4.32) into single oscillations of the sine, $J_k = [\frac{2\pi k}{n}, \frac{2\pi(k+1)}{n}]$, $k \geq 0$. As $\sqrt{-\log\phi}$ has negative slope on $(0;1)$, each of the $J_k$ gives a positive contribution to $\mathcal{I}$, and thus we can truncate the integral at $\frac{1}{2}$,

$$\mathcal{I} \geq \sum_{\frac{2\pi(k+1)}{n} \leq \frac{1}{2}} \int_{J_k} \sqrt{-\log\phi}\,\sin(n\phi)\,\mathrm{d}\phi \ . \tag{4.33}$$

On $[0;\frac{1}{2}]$, $\sqrt{-\log\phi}$ has a positive curvature, and thus, each of the integrals can be estimated by linearly approximating $\sqrt{-\log\phi}$ at the middle of each $J_k$ but with the slope at $\frac{2\pi(k+1)}{n}$, which gives

$$\int_{J_k} \sqrt{-\log\phi}\,\sin(n\phi)\,\mathrm{d}\phi \geq \frac{\pi}{n^2}\frac{1}{\frac{2\pi(k+1)}{n}\sqrt{-\log\left[\frac{2\pi(k+1)}{n}\right]}} \ .$$

Now, we plug this into the sum (4.33) and bound the sum by the integral from $\frac{2\pi}{n}$ to $\frac{1}{2}$ (the integrand in monotonically decreasing), which indeed gives a lower bound $\frac{1}{n}(\sqrt{\log\frac{n}{2\pi}} - \sqrt{\log 2})$ on $\mathcal{I}$ and thus proves the $\Theta(\sqrt{\log n}/n^2)$ scaling. □

### 4.6.2 Higher dimensions

For more than one dimension, the situation is more involved. First of all, it is clear by taking many uncoupled copies of the one-dimensional chain that there exist cases where the correlations will decay as $n^{-2}$. However, these are very special examples corresponding to Hamiltonians with a tensor product structure $H_{i_1 i_2, j_1 j_2} = H_{i_1,j_1} H'_{i_2,j_2}$. In contrast, we show that for generic systems the correlations in the critical case decay as $O(n^{-(d+1)}\log n)$, where $d$ is the dimension of the lattice. The requirement is again that the energy spectrum $\mathcal{E}(\phi)$ has only a finite number of zeroes, i.e., finitely many critical points.

Note that the case of a Hamiltonian with a tensor product structure can also be solved, as in that case $\hat{V}$ becomes a product of terms depending on one $\phi_i$ each and thus the integral factorizes. Interestingly, although the correlations along the axes decay as $n^{-2}$, the correlations in a fixed diagonal direction will decay as $n_1^{-2} \cdots n_d^{-2} \propto \|n\|^{-2d}$ and thus even faster than in the following theorem. The $O\big(\|n\|^{-(d+1)} \log \|n\|\big)$ decay of the theorem holds isotropically, i.e., independent of the direction of $n$.

**Theorem 4.15.** *Consider a $d$-dimensional bosonic lattice with a critical Hamiltonian $H = V \oplus \mathbb{1}$. Then the $P$-correlations of the ground state decay as*

$$O\left(\|n\|^{-(d+1)} \log \|n\|\right)$$

*if the following holds: $\hat{V} \in \mathscr{C}^{d+1}$ [e.g., the correlations decay as $O(\|n\|^{-(2d+1+\varepsilon)})$, $\varepsilon > 0$], and $\hat{V}$ has only a finite number of zeros which are quadratic minima, i.e., the Hessian $\left(\frac{\partial^2 \hat{V}(\phi)}{\partial \phi_i \partial \phi_j}\right)_{ij}$ is positive definite at all zeros.*

**Proof.** We have to evaluate the asymptotic behavior of the integral

$$(\hat{V}^{1/2})_n = \frac{1}{(2\pi)^d} \int_{\mathcal{T}^d} \mathrm{d}^d \phi \sqrt{\hat{V}(\phi)} \cos[n\phi] \ .$$

Let us first briefly sketch the proof. We start by showing that it suffices to analyze each critical point separately. To this end, we show that is is possible to smoothly cut out some environment of each critical point which reproduces the asymptotic behavior. Then, we rotate the coordinate system such that we always look at the correlations in a fixed direction, and integrate by parts—which surprisingly can be carried out as often as $\hat{V}$ is differentiable, as all the brackets vanish. Therefore, the information about the asymptotics is contained in the remaining integral, and after a properly chosen number of partial integrations, we will attempt to estimate this term.

Let now $\zeta_i$, $i = 1, \ldots, I$ be the zeros of $\hat{V}$. Clearly, these will be the only points which contribute to the asymptotics as everywhere else $\sqrt{\hat{V}}$ is $\mathscr{C}^{d+1}$. In order to separate the contributions coming from the different $\zeta_i$, we will make use of so-called *neutralizers* [103]. For our purposes, these are functions $\mathcal{N}_{\xi_0,r} \in \mathscr{C}^\infty(\mathbb{R}^d \to [0;1])$ which satifsy

$$\mathcal{N}_{\xi_0,r}(\xi) = \begin{cases} 1 : \|\xi - \xi_0\| \leq r/2 \\ 0 : \|\xi - \xi_0\| \geq r \end{cases}$$

and are rotationally symmetric (cf. [103] for an explicit construction). For each $\zeta_i$, there exists an $r_i$ such that the balls $B_{r_i}(\zeta_i)$ do not intersect. We now define the functions

$$f_i(\phi) := \sqrt{\hat{V}(\phi)}\, \mathcal{N}_{\zeta_i, r_i}(\phi) \ , \quad \rho(\phi) := \sqrt{\hat{V}(\phi)} - \sum_{i=1}^{I} f_i(\phi) \ .$$

Clearly, $\rho$ is $\mathscr{C}^{d+1}$, and so is each $f_i$ except at $\zeta_i$. Furthermore, each $f_i$ is still the square root of a $\mathscr{C}^{d+1}$ function. By definition,

$$(\hat{V}^{-1/2})_n = \frac{1}{(2\pi)^d} \sum_{i=1}^{I} \int_{\mathcal{T}^d} \mathrm{d}^d\phi f_i(\phi) \cos[n\phi] + \frac{1}{(2\pi)^d} \int_{\mathcal{T}^d} \mathrm{d}^d\phi \rho(\phi) \cos[n\phi] \ , \quad (4.34)$$

i.e., it suffices to look at the asymptotics of each $f_i$ separately. The contribution of $\rho$ is $O(\|n\|^{-(d+1)})$ as can be shown by successive integrations by parts just as for the non-critial lattice (cf. the proof of Lemma 4.5).

Let us now analyze the integrals

$$I_i = \int_{B_{r_i}(\zeta_i)} \mathrm{d}^d\phi f_i(\phi) \cos[n\phi] \ .$$

The integration range can be restriced to $B_{r_i}(\zeta_i)$ as $f_i$ vanishes outside the ball. By a rotation, this can be mapped to an integral where $n = (\|n\|, 0, \dots, 0)$, whereas $f_i$ is rotated to another function $\tilde{f}_i$ with the same properties,

$$I_i = \int_{B_{r_i}(\zeta_i)} \mathrm{d}^d\phi \tilde{f}_i(\phi) \cos[\|n\|\phi_1] \ .$$

Since the integrand is continuous and thus bounded, it is absolutely integrable, and from Fubini's theorem, one finds

$$I_i = \int_{B_{r_i}(\tilde{\zeta}_i)} \mathrm{d}^{d-1}\tilde{\phi} \underbrace{\int_{\zeta_{i,1}-r_i}^{\zeta_{i,1}+r_i} \mathrm{d}\phi_1 \tilde{f}_i(\phi_1, \tilde{\phi}) \cos[\|n\|\phi_1]}_{J_i(\tilde{\phi})} \ ,$$

where we separated out the integration over the first component. The vector $\tilde{\phi}$ denotes the components $2 \dots d$ of $\phi$. The extension of the integration range to a cylinder is possible as $\tilde{f}_i$ vanishes outside $B_{r_i}(\zeta_i)$.

Let us now require $\tilde{\phi} \neq \tilde{\zeta}_i$. This does not change the integral since the excluded set is of measure zero, but it ensures that $\tilde{f}_i$ is in $\mathscr{C}^{d+1}$. This allows us to integrate the inner integral $J_i(\tilde{\phi})$ by parts up to $d+1$ times, and each of the brackets

$$\left[ \tilde{f}_i^{(k)}(\phi_1, \tilde{\phi}) \frac{1}{\|n\|^k} \cos(\|n\|\phi_1 - k\pi/2) \right]_{\phi_1=\zeta_{i,1}-r_i}^{\zeta_{i,1}+r_i}$$

appearing in the $k$'th integration step vanishes. Here, $\tilde{f}_i^{(d)}(\phi_1, \tilde{\phi}) = \partial^d \tilde{f}_i(\phi_1, \tilde{\phi}) / \partial \phi_1^d$ is the $d$'th partial derivative with respect to the first argument. After integrating by parts $d$ times, we obtain

$$I_i = \frac{1}{\|n\|^d} \int_{B_{r_i}(\tilde{\zeta}_i)} \mathrm{d}^{d-1}\tilde{\phi} \int_{\zeta_{i,1}-r_i}^{\zeta_{i,1}+r_i} \mathrm{d}\phi_1 \tilde{f}_i^{(d)}(\phi_1, \tilde{\phi}) \cos[\|n\|\phi_1 - d\pi/2] \ . \quad (4.35)$$

Now we proceed as follows: first, we show that the order of integration can be interchanged, and second, we show that for the function obtained after integrating $\tilde{f}_i^{(d)}$ over $\tilde{\phi}$, the Fourier coefficients vanish as $\log(\|n\|)/\|n\|$.

The central issue for what follows is to find suitable bounds on $|\tilde{f}_i^{(k)}|$. Therefore, define $\tilde{f}_i^2 =: h_i \in \mathscr{C}^{d+1}$. By virtue of Taylor's theorem, and as $h_i(\zeta_i) = 0$ is a minimum,

$$h_i(\phi) = \tfrac{1}{2}(\phi - \zeta_i) \cdot (\mathbf{D}^2 h_i(\zeta_i))(\phi - \zeta_i) + o(\|\phi - \zeta_i\|^2)$$

with $\mathbf{D}^2$ the second derivative, i.e., the Hessian. As the first term is bounded by $\frac{1}{2}\|\mathbf{D}^2 h_i(\zeta_i)\|_\infty \|\phi - \zeta_i\|^2$ and the second vanished faster than $\|\phi - \zeta_i\|^2$, we can find $\varepsilon_i > 0$ and $C_1 > 0$ such that

$$|h_i(\phi)| \leq C_1 \|\phi - \zeta_i\|^2 \qquad \forall \|\phi - \zeta_i\| < \varepsilon_i \ . \tag{4.36}$$

By looking at the Tayor series of $h_i' \equiv \partial h_i / \partial \phi_1$ up to the first order we also find that there are $\varepsilon_i > 0$ and $C_2 > 0$ such that

$$|h_i'(\phi)| \leq C_2 \|\phi - \zeta_i\| \qquad \forall \|\phi - \zeta_i\| < \varepsilon_i \ . \tag{4.37}$$

In addition to these upper bounds, we will also need a lower bound on $|h_i|$. Again, by the Taylor expansion of $h_i$ around $\zeta_i$, we find

$$|h_i(\phi)| \geq \tfrac{1}{2}\, \lambda_{\min} \left[\mathbf{D}^2 h_i(\zeta_i)\right]\, \|\phi - \zeta_i\|^2 - o(\|\phi - \zeta_i\|^2) \ ,$$

and as all the zeros are quadratic minima, i.e., $\lambda_{\min}\left[\mathbf{D^2} h_i(\zeta_i)\right] > 0$, there exist $\varepsilon_i > 0$, $C_3 > 0$ such that

$$|h_i(\phi)| \geq C_3 \|\phi - \zeta_i\|^2 \qquad \forall \|\phi - \zeta_i\| < \varepsilon_i \ . \tag{4.38}$$

Clearly, $\varepsilon_i$ can be chosen equal in Eqs. (4.36–4.38). Note that the bounds can be chosen to be invariant under rotation of $h_i$ and thus of $\tilde{f}_i$. This holds in particular for the $\varepsilon_i$ as the remainders of Taylor series vanish uniformly. Thus, the bound we will obtain for the correlation function indeed only depends on $\|n\|$ and not on the direction of $n$.

Now, we use the conditions (4.36–4.38) to derive bounds on $|\tilde{f}_i^{(k)}|$. Therefore, note that from $\tilde{f}_i \equiv \sqrt{h_i}$ it follows that

$$\tilde{f}_i^{(k)} = \frac{\sum\limits_{\substack{j_1+\cdots+j_k=k \\ j_\nu = 0,1,2,\ldots}} c_{j_1 \ldots j_k} h_i^{(j_1)} \cdots h_i^{(j_k)}}{h_i^{(2k-1)/2}} \ .$$

One can easily check that for each term in the numerator, the number $K_0$ of zeroth derivatives and the number $K_1$ of first derivatives of $h_i$ satisfy $2K_0 + K_1 \geq k$. By bounding all higher derivatives of $h_i$ from above by constants, we find that

the modulus of each summand in the numerator, and thus the modulus of the numerator itself, can be bounded above by $C'\|\phi-\zeta_i\|^k$ in the ball $B_{\varepsilon_i}(\zeta_i)$ with some $C'>0$. On the other hand, it follows directly from (4.38) that the modulus of the denominator is bounded below by $C''\|\phi-\zeta_i\|^{2k-1}$, $C''>0$, such that in total

$$|\tilde{f}_i^{(k)}(\phi)| \leq C\frac{1}{\|\phi-\zeta_i\|^{k-1}}\ ; \quad 1\leq k\leq d+1\ . \tag{4.39}$$

Note that this holds not only inside $B_{\varepsilon_i}(\zeta_i)$ but in the whole domain of $f_i$, as outside $B_{\varepsilon_i}(\zeta_i)$, $f_i$ is $\mathscr{C}^{d+1}$ and thus all the derivatives are bounded.

Eq. (4.39) is the key result for the remaining part of the proof. First, it can be used to bound the integrand in (4.35) by an integrable singularity (this is most easily seen in spherical coordinates, where $1/r^{d-1}$ is integrable in a $d$-dimensional space). Hence, the order of integration in (4.35) can be interchanged, and it remains to investigate the asymptotics of the integral

$$I_i = \frac{1}{\|n\|^d}\int_{\zeta_{i,1}-r_i}^{\zeta_{i,1}+r_i} \mathrm{d}\phi_1 g_i(\phi_1)\cos[\|n\|\phi_1 - d\pi/2]\ , \quad \text{with} \tag{4.40}$$

$$g_i(\phi_1) \equiv \int_{B_{r_i}(\tilde{\zeta}_i)} \mathrm{d}^{d-1}\tilde{\phi}\ \tilde{f}_i^{(d)}(\phi_1,\tilde{\phi})\ . \tag{4.41}$$

From (4.39), we now derive bounds on $g_i(\phi_1)$ and its first derivative. Again, we may safely fix $\phi_1\neq\zeta_{i,1}$ as this has measure zero. Then, using (4.39) we find that

$$|g_i(\phi_1)| \leq \int_0^{r_i} \frac{C}{((\phi_1-\zeta_{i,1})^2+r^2))^{(d-1)/2}} S_{d-1} r^{d-2}\mathrm{d}r$$

where we have transformed into spherical coordinates [$S_{d-1}$ is the surface of the $(d-1)$-dimensional unit sphere] and assumed the $l_2$-norm. Since $(\phi_1-\zeta_1)^2+r^2\geq r^2$, the integrand can be bounded once again, and we find

$$\begin{aligned} |g_i(\phi_1)| &\leq \int_0^{r_i}\frac{CS_{d-1}}{((\phi_1-\zeta_{i,1})^2+r^2)^{1/2}}\mathrm{d}r \\ &= C\left(-\log|\phi_1-\zeta_{i,1}| + \log\left[r_i+\sqrt{r_i^2+(\phi_1-\zeta_{i,1})^2}\right]\right) \\ &\leq -C\log|\phi_1-\zeta_{i,1}| \end{aligned} \tag{4.42}$$

where in the last step we used that in (4.40) $|\phi_1-\zeta_{i,1}|<r_i$ and that $r_i$ can be chosen sufficiently small.

Next, we derive a bound on $g_i'(\phi_1)$. As we fix $\phi_1\neq\zeta_1$, the integrand in (4.41) is $\mathscr{C}^1$ and we can take the differentiation into the integral,

$$g_i'(\phi_1) = \int_{B_{r_i}(\tilde{\zeta}_i)} \mathrm{d}^{d-1}\tilde{\phi}\,\tilde{f}_i^{(d+1)}(\phi_1,\tilde{\phi})\ .$$

Again, we bound the integrand by virtue of Eq. (4.39) and obtain

$$\begin{aligned} |g_i'(\phi_1)| &\leq \int_0^{r_i} \frac{CS_{d-1}}{((\phi_1-\zeta_{i,1})^2+r^2)}\mathrm{d}r \\ &= C\frac{\arctan\left[\frac{r_i}{|\phi_1-\zeta_{i,1}|}\right]}{|\phi_1-\zeta_{i,1}|} \leq \frac{C'}{|\phi_1-\zeta_{i,1}|} \ . \end{aligned} \tag{4.43}$$

Finally, these two bounds will allow us to estimate (4.40) and thus the asymptotics of the correlations in the lattice. We consider one half of the integral (4.40),

$$\int_{\zeta_{i,1}}^{\zeta_{i,1}+r_i} \mathrm{d}\phi_1 g_i(\phi_1)\cos[\|n\|\phi_1 - d\pi/2] \ , \tag{4.44}$$

as both halves contribute equally to the asymptotics. We then split the integral at $\zeta_{i,1}+r_i/\|n\|$. The left part gives

$$\begin{aligned} \left|\int_{\zeta_{i,1}}^{\zeta_{i,1}+r_i/\|n\|} \mathrm{d}\phi_1 g_i(\phi_1)\cos[\|n\|\phi_1 - d\pi/2]\right| &\overset{(4.42)}{\leq} C\int_{\zeta_{i,1}}^{\zeta_{i,1}+r_i/\|n\|} \mathrm{d}\phi_1(-\log|\phi_1-\zeta_{i,1}|) \\ &= C\,\frac{r_i - r_i\log r_i + r_i\log\|n\|}{\|n\|} \ . \end{aligned} \tag{4.45}$$

The right part of the split integral (4.44) can be estimated by integration by parts,

$$\begin{aligned} &\left|\int_{\zeta_{i,1}+r_i/\|n\|}^{\zeta_{i,1}+r_i} \mathrm{d}\phi_1 g_i(\phi_1)\cos[\|n\|\phi_1 - d\pi/2]\right| \leq \\ &\quad\leq \left|\left[g_i(\phi_1)\frac{1}{\|n\|}\cos[\|n\|\phi_1-(d+1)\pi/2]\right]_{\zeta_{i,1}+r_i/\|n\|}^{\zeta_{i,1}+r_i}\right| + \frac{1}{\|n\|}\int_{\zeta_{i,1}+r_i/\|n\|}^{\zeta_{i,1}+r_i}\mathrm{d}\phi_1|g_i'(\phi_1)| \\ &\quad\overset{(4.42,4.43)}{\leq} C\frac{\log\|n\|}{\|n\|} + C'\frac{|\log r_i|}{\|n\|} \ . \end{aligned} \tag{4.46}$$

Thus, both halves [Eqs. (4.45),(4.46)] give a $\log\|n\|/\|n\|$ bound for the integral (4.44), and therefore the integral $I_i$ is asymptotically bounded by $\log\|n\|/\|n\|^{d+1}$ following Eq. (4.40). As the number of such integrals in (4.34) is finite, this proves that the correlations of the ground state decay at least as $\log\|n\|/\|n\|^{d+1}$. □

# Chapter 5

# Entropy and approximability by Matrix Product States

## 5.1 Introduction

Understanding the behaviour of quantum many-body systems is a central problem in physics. Recently, Matrix Product States (MPS) have received much interest as a variational ansatz for the simulation of correlated one-dimensional systems. They have proven particularly powerful in approximating the ground states of local Hamiltonians, as used in the DMRG method [17, 18], but have also been applied, e.g., to simulate the time evolution of slightly entangled quantum systems [22]. Despite considerable progress [104], it is still not fully understood which property exactly a state has to fulfil to be well approximated by MPS. This knowledge is not only of practical interest, but could also tell us how to extend the MPS ansatz to, e.g., higher dimensional systems.

It is generally believed that the relevant criterion for efficient approximability by MPS is that the states under consideration obey an area law, i.e., the von Neumann entropy of a block is bounded. Although indeed both ground states of local Hamiltonians and MPS obey an area law, there are reasons to doubt this immediate connection: Firstly, the von Neumann entropy is an asymtotic concept, quantifying what happens when dealing with a large number of copies of a state. Conversely, it has been shown recently that a rigorous connection can be established by looking at Rényi entropies instead [26]. Unfortunately, the argument used breaks down as the von Neumann entropy is approached. Finally, the continuity inequality for the von Neumann entropy carries a size-dependent constant, and thus states which are close to each other need not be close in entropy [105].

In the following, we fully explore the connection between entropy scaling and approximability by MPS. The results are summarized in Table 5.1: An at most logarithmic scaling of Rényi entropies $S_\alpha$, $\alpha < 1$, implies approximability by

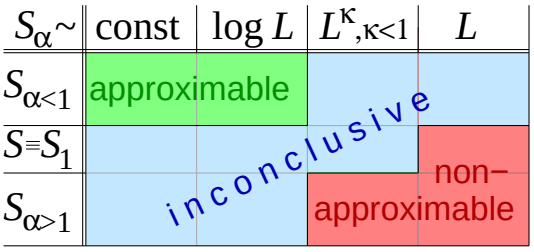


**Table 5.1:** Relation between scaling of block Rényi entropies and approximability by MPS. In the "inconclusive" region, the scaling does not imply anything about approximability.

MPS. On the other side, a faster than logarithmic increase of $S_\alpha$, $\alpha > 1$, rules out efficient approximability by MPS, as does linear growth of the von Neumann entropy. For all other cases, the scaling of the block entropy does not allow for conclusions about approximability. In particular, this holds for the case of constant von Neumann entropy, which demonstates that the reason why MPS describe ground states well is not simply that those states obey an area law, but rather some additional property.

Finally, we apply our results to illustrate that quantum computers might outperform classical computers in simulating time evolutions. It is long-known that quantum computers can simulate the evolution of quantum systems [106]. However, this does not necessarily imply that they will perform better than classical computers in this task: e.g., ground states of gapped quantum systems appear to be classically efficiently approximable [104]. On the other hand, it is known that time evolution under a translational invariant Hamiltonian can implement quantum computations if either translational invariance is broken by the initial [107] or boundary conditions [108], or the Hamiltonian is time dependent [109], and is thus hard to simulate. We extend these results by showing that even the evolution of a translational invariant spin $\frac{1}{2}$ system with translational invariant initial conditions under a time independent Hamiltonian cannot be simulated efficiently using MPS; this provides strong evidence that quantum computers will outperform classical computers in simulating these systems.

## 5.2 Definitions

Let us first introduce the relevant quantities and notations. We want to obtain approximations which reproduce accurately not only the local properties such as energy, but also the non-local ones such as correlations. This is ensured by bounding the error made for an arbitrary observable $O$,

$$|\mathrm{tr}[\psi O] - \mathrm{tr}[\phi O]| \leq \|O\|_{\mathrm{op}} \|\psi - \phi\|_{\mathrm{tr}} \ ,$$

where $\psi \equiv |\psi\rangle\langle\psi|$. We focus on non-extensive observables,[1] therefore w.l.o.g. $\|O\|_{\rm op} \leq 1$. It follows that by imposing

$$\|\psi - \phi\|_{\rm tr} \leq \delta \ , \tag{5.1}$$

we bound the error made in any observable by $\delta$.

For some of the proofs it will be more convenient to consider the two-norm distance $\big||\psi\rangle - |\phi\rangle\big|_2$ or the fidelity

$$\frac{\big|\langle\phi|\psi\rangle\big|}{\big||\phi\rangle\big|_2\big||\psi\rangle\big|_2} =: \cos(\theta) \ .$$

Fortunately, these measures turn all out to be equvalent: Since the best approximating MPS will generally not be normalized, it is appropriate to consider the optimized quantities, and one finds that

$$T(\phi,\psi) := \inf_{\alpha>0} \frac{\|\psi - \alpha\phi\|_{\rm tr}}{\|\psi\|_{\rm tr}} \equiv \sin(2\theta)$$

and

$$V(\phi,\psi) := \inf_{\alpha\in\mathbb{C}} \frac{\big||\psi\rangle - \alpha|\phi\rangle\big|_2}{||\psi\rangle|_2} \equiv \sin(\theta)$$

for $0 \leq \theta \leq \frac{\pi}{4}$.

We now introduce Matrix Product States (MPS) [19]. Consider a chain of $N$ $d$-level systems with the corresponding Hilbert space

$$\mathcal{H}_N := (\mathbb{C}^d)^{\otimes N} \ .$$

We call $|\phi_D\rangle \in \mathcal{H}_N$ a Matrix Product State (MPS) with *bond dimension* $D$ (or, briefly, a $D$-MPS) if it can be written as

$$|\phi_D\rangle = \sum_{i_1,\ldots,i_N=1}^{d} A^{[1]}_{i_1} A^{[2]}_{i_2} \cdots A^{[N]}_{i_N} |i_1, i_2, \ldots, i_N\rangle \tag{5.2}$$

with $A^{[k]}_i$ $D \times D$ matrices for $2 \leq k \leq N-1$, and vectors for $k = 1, N$.

Given a family $(|\psi_N\rangle)$ of states, $|\psi_N\rangle \in \mathcal{H}_N$, we say that it *can be approximated efficiently by MPS* if for every $\delta > 0$, there exists a sequence $|\phi_{N,D}\rangle$ of MPS with $D \equiv D(N) = O(\mathrm{poly}_\delta(N))$ such that $\|\psi_N - \phi_{N,D}\|_{\rm tr} \leq \delta$. On the contrary, if there is some $\delta > 0$ such that no sequence of MPS with polynomial bond dimension can

[1] For extensive observables, where $\|\psi - \psi_D\|_{\rm tr} \leq \delta/N$, the results are the same except that $S \sim N^\kappa$ now implies inapproximability [replace $\delta$ by $\delta/N$ in (5.5)]. In the approximability example for linearly growing $\alpha < 1$ Rényi entropy, one has to set $p_N = 1/N^3$. As an example of a non-approximable state with bounded von Neumann entropy, a single copy of (5.6) is now sufficient, yielding a translation invariant example.

approximate $|\psi\rangle$ up to $\delta$, we say that $(|\psi_N\rangle)$ *cannot be approximated efficiently by MPS*. For brevity, we will sometimes drop the word "efficiently".

We will measure entropies using the *Rényi entropies*

$$S_\alpha(\rho) = \frac{\log \mathrm{tr}\rho^\alpha}{1-\alpha} \ , \ 0 \leq \alpha \leq \infty \ ,$$

which are a generalization of the von Neumann entropy $S(\rho) = -\mathrm{tr}[\rho \log \rho]$. In particular, $\lim_{\alpha\to 1} S_\alpha(\rho) = S(\rho)$. Note that all logs are to the basis 2.

## 5.3 Approximability and truncation error

We aim to relate approximability by MPS to the scaling of block entropies. To this end, we first show that the error made in approximating some state by a $D$-MPS is determined by the error made when truncating the Schmidt spectrum of its bipartitions after $D$ values. Therefore, let $|\psi\rangle \in \mathcal{H}_N$, $\rho_k = \mathrm{tr}_{k+1,\ldots,N}|\psi\rangle\langle\psi|$, and let $\lambda_1^{[k]} \geq \lambda_2^{[k]} \geq \cdots \geq \lambda_{d^k}^{[k]}$ be the ordered spectrum of $\rho_k$. Then, define the *truncation error*

$$\epsilon_k(D) := \sum_{i=D+1}^{k^d} \lambda_i^{[k]} \ .$$

Let us now relate the truncation error to approximability by MPS. The intuition is that the best an MPS with bond dimension $D$ (i.e., Schmidt rank $D$ in any bipartition) can do is preserve the $D$ largest eigenvalues, resulting in an error of $\epsilon_k(D)$ for the cut at $k$ (which can, but need not, accumulate). On the one side, it has been shown in [26] that for a state $|\psi\rangle \in \mathcal{H}_N$, there always exists an MPS $|\phi_D\rangle$ with bond dimension $D$ such that

$$\big||\psi\rangle - |\phi_D\rangle\big|_2 \leq 2\sum_{k=1}^{N-1} \epsilon_k(D) \ . \tag{5.3}$$

On the other hand, any $D$-MPS $|\phi_D\rangle$ satisfies

$$\|\psi - \phi_D\|_{\mathrm{tr}} \geq \epsilon_k(D) \quad \forall k \ , \tag{5.4}$$

since with $\rho_k = \mathrm{tr}_{1,\ldots,k}\psi$, $\sigma_{D,k} = \mathrm{tr}_{1,\ldots,k}\phi_D$, we have

$$\|\psi - \psi_D\|_{\mathrm{tr}} \geq \|\rho_k - \rho_{D,k}\|_{\mathrm{tr}} \geq \epsilon_k(D) \ ,$$

where we have used the contractivity of the partial trace, that for fixed spectra, the trace norm distance is extremal for commuting operators [110], and that $\mathrm{rank}\,\rho_{D,k} \leq D$.

## 5.4 Conclusive cases

We start the discussion of Table 5.1 by proving the cases for which conclusive statements can be made. In the following, $\rho_N^L$ will denote any $L$-particle reduced block of a state $|\psi_N\rangle \in \mathcal{H}_N$. The case of at most logarithmically growing Rényi entropy with $\alpha < 1$ was discussed in [26], where is was shown that it implies approximability. More formally, if for a family of states $(|\psi_N\rangle)$ there exist $c, c' > 0$ and $0 \leq \alpha < 1$ such that $S_\alpha(\rho_N^L) \leq c \log(L) + c'$ for all reduced blocks $\rho_N^L$, then it can be approximated efficiently by MPS.

Let us now show that a linearly growing von Neumann entropy implies inapproximability. Formally, if for a family $(|\psi_N\rangle)$, $S(\rho_N^L) \geq cL$ holds for some $c > 0$, $L \equiv L(N) \geq \eta N$, $\eta > 0$, and some reduced blocks $\rho_N^L$, then it cannot be approximated efficiently by MPS.

To prove this, we use Fannes' inequality in its improved version by Audenaert [105]: For density operators $\rho$, $\sigma$ on a $K$-dimensional Hilbert space,

$$|S(\rho) - S(\sigma)| \leq T \log(K-1) + H(T, 1-T) \ ,$$

where $2T = \|\rho - \sigma\|_{\mathrm{tr}} \leq \delta$, and $H(T, 1-T) \leq 1$ is the binary entropy. Let $(|\phi_{N,D}\rangle)$ be a sequence of MPS approximating $(|\psi_N\rangle)$ up to $\delta$, and $\rho_N^L$, $\sigma_{N,D}^L$ the corresponding reduces states for which $S(\rho_N^L) \geq cL$. Then,

$$|S(\rho_N^L) - S(\sigma_{N,D}^L)| \leq \tfrac{1}{2}\delta L \log d + 1 \ ,$$

and thus, for $L \geq \eta N$,

$$\begin{aligned} \log D(N) \geq S(\sigma_{N,D}^L) &\geq S(\rho_N^L) - \tfrac{1}{2}\delta L \log d - 1 \\ &\geq \eta(c - \tfrac{1}{2}\delta \log d)N - 1 \end{aligned} \tag{5.5}$$

as soon as the error $\delta < 2c/\log d$, i.e., the bond dimension grows exponentially in $N$, which completes the proof.

In the following, we show that a faster than logarithmic increase of any Rényi entropy with $\alpha > 1$ also implies inapproximability, i.e., if for a family $(|\psi_N\rangle)$, there exist $\alpha > 1$ and $\kappa > 0$ s.th. $S_\alpha(\rho_N^L) \geq cL^\kappa$ for some $c > 0$, $L \equiv L(N) \geq \eta N$, and some reduced blocks $\rho_N^L$, then it cannot be approximated efficiently by MPS.

This is proven by lower bounding the truncation error $\epsilon \equiv \epsilon(D)$ of a block $\rho_N^L$ for given $S_\alpha(\rho)$ $(\alpha > 1)$ and then applying (5.4). This, however, is the same as maximizing the entropy while keeping $\epsilon$ fixed. Since both the entropy and $\epsilon$ only depend on the spectrum, the problem reduces to a classical one. It is easy to see that the probability distribution

$$p_1, \ldots, p_D = \frac{1-\epsilon}{D}; \ p_{D+1}, \ldots, p_{2^L} = \frac{\epsilon}{2^L - D}$$

is majorized by all decreasingly ordered probability distributions $(q_i)$ which satisfy $q_{D+1} + \cdots + q_{2^N} = \epsilon$, and since Renyi entropies are Schur concave functions, it

has maximal entropy [110]. Therefore, we obtain the inequality

$$S_\alpha(\rho_N^L) \leq \frac{-1}{\alpha-1}\log\left[\frac{(1-\epsilon)^\alpha}{D^{\alpha-1}} + \frac{\epsilon^\alpha}{(2^L-D)^{\alpha-1}}\right]$$
$$\leq \frac{-1}{\alpha-1}\log\left[\frac{(1-\epsilon)^\alpha}{D^{\alpha-1}}\right] = \log D - \frac{\alpha}{\alpha-1}\log(1-\epsilon)\ .$$

Since from (5.4) the total error is $\delta \geq \epsilon$, we find

$$\log D \geq S_\alpha(\rho_N^L) + \frac{\alpha}{\alpha-1}\Big|\log(1-\delta)\Big|\ ,$$

and from $S_\alpha(\rho_N^L) \geq cL^\kappa \geq c\eta^\kappa N^\kappa$, we infer that $D$ has to grow exponentially for any $\delta$.

## 5.5 Inconclusive cases

We now turn towards the inconclusive region in Table 5.1, where we provide examples for both approximability and inapproximability. This task is greatly simplified by the fact that approximability examples extend to the top and left in Table 5.1, while non-approximability extends to the right and bottom. This holds as approximability for a given scaling implies the same for more moderate scalings (and conversely for non-approximability), and since $S_\alpha(\rho)$ decreases monotonically in $\alpha$.

We aim to clarify the relation between entropy scaling laws and the approximability by MPS: Therefore, our examples are not constructed to be ground states. Yet, all of them form uniform families of states, i.e., they can be generated by a uniform family of time dependent Hamiltonians. The existence of time-independent realizations is plausible, as the central ingredient of the examples are properly distributed entangled pairs. These could be represented by pairs of localized excitations which are prepared locally and then propagated by a time-independent Hamiltonian.

All of the examples can be chosen to be translational invariant, with the only possible exception of the inapproximability example for constant von Neumann entropy. The question whether any translational invariant state with bounded von Neumann entropy can be approximated efficiently by MPS thus remains open.

The examples can be grouped into two classes; the first is based on states of the type

$$|\psi_{2N}\rangle = \sqrt{1-p_N}|2\rangle^{\otimes 2N} + \sqrt{\frac{p_N}{2^N}}\sum_{x\in\{0,1\}^N}|x\rangle|x\rangle\ . \tag{5.6}$$

By choosing $p_N = 1/N$, we obtain an example of a state with linearly growing Rényi entropies for all $\alpha < 1$ which can be approximated by MPS, as

$$\big||\psi_{2N}\rangle - \sqrt{1-p_N}|2\rangle^{\otimes 2N}\big|_2 = \sqrt{p_N} \to 0\ .$$

On the other hand, for $L \leq N$,

$$\rho_{2N}^L = (1-p_N)|2\rangle\langle 2|^{\otimes L} + \frac{p_N}{2^L} \sum_{y\in\{0,1\}^L} |y\rangle\langle y| \ ,$$

and therefore

$$\begin{aligned} S_\alpha(\rho_{2N}^L) &= \frac{1}{1-\alpha} \log \left[(1-p_N)^\alpha + 2^{(1-\alpha)L} p_N^\alpha\right] \\ &\geq L - \frac{\alpha}{1-\alpha} \log N \ . \end{aligned}$$

Note that the infavourable scaling of $c_\alpha := \frac{\alpha}{1-\alpha}$ for $\alpha \to 1$ can be compensated by e.g. choosing $p_N = N^{-1/c_\alpha}$.

The next example provides states with algebraically (but sublinearly) growing von Neumann entropy which can be approximated efficiently by MPS. Therefore, fix $0 < \kappa < 1$ and $\epsilon > 0$, and set $p_N = N^{-\epsilon(1-\kappa)}$ in (5.6). As in the previous example, $p_N \to 0$ implies approximability, and

$$S(\rho_{2N}^L) = H(p_N, 1-p_N) + p_N \log[2^L] \geq L/N^{\epsilon(1-\kappa)} \ ,$$

which implies $S(\rho_{2N}^L) \geq L^\kappa$ for $L \geq N^\epsilon$.

We now construct a state which obeys a strict area law for the von Neumann entropy but yet cannot be approximated by MPS. Therefore, set $M = 2N^3$ and define $|\chi_M\rangle = |\psi_{2N}\rangle^{\otimes N^2}$ with $|\psi_{2N}\rangle$ from (5.6), where $p_N = 1/N$. Then, $S(\rho_M^L)$ is at most twice the maximum entropy of a cut through $|\psi_{2N}\rangle$, and thus

$$S(\rho_M^L) \leq 2\left(H(p_N, 1-p_N) + p_N N\right) \leq 4 \ .$$

To prove hardness of approximation, observe that for a given $D$, the best MPS approximation to $|\psi_{2N}\rangle^{\otimes N^2}$ is of the form $|\phi_D\rangle^{\otimes N^2}$, i.e., it carries the same product structure.[2] From the multiplicativity of the fidelity and the relations following Eq. (5.1) one infers

$$T(\phi^{\otimes K}, \psi^{\otimes K}) \geq \sqrt{K/8}\, T(\phi, \psi)$$

[2] It holds in full generality that the optimal $D$-MPS approximation to a state of the form $|\psi_A\rangle|\psi_B\rangle \in \mathcal{H}_N \otimes \mathcal{H}_M$ carries the same product structure: Given any normalized $D$-MPS $|\phi_D\rangle$, insert $\sum_k |k\rangle\langle k|$ in (5.2) between $A^{[N]}$ and $A^{[N+1]}$. This gives a decomposition $|\phi_D\rangle = \sum_k |\alpha_k\rangle|\beta_k\rangle$ with $|\alpha_k\rangle$ ($|\beta_k\rangle$) spanning a subspace of $D$-MPS, as they only differ by one boundary condition. Thus, the same holds for the Schmidt decomposition $|\phi_D\rangle = \sum_k \lambda_k |\tilde{\alpha}_k\rangle|\tilde{\beta}_k\rangle$ (with normalized vectors). Define $a_k := \langle\psi_A|\tilde{\alpha}_k\rangle$, $b_k := \langle\psi_B|\tilde{\beta}_k\rangle$, and the product approximation

$$|\phi_D'\rangle = \left(\sum_k \lambda_k \tfrac{|a_k|}{a_k} |\tilde{\alpha}_k\rangle\right) \left(\tfrac{1}{|\mathcal{L}|^{1/2}} \sum_{l\in\mathcal{L}} \tfrac{|b_l|}{b_l} |\tilde{\beta}_l\rangle\right) \ ,$$

where $\mathcal{L} = \{l : |b_l| \geq |b_j| \, \forall j\}$. Then, $|\phi_D'\rangle$ is a normalized $D$-MPS, and $|\langle\psi_A, \psi_B|\phi_D\rangle| < |\langle\psi_A, \psi_B|\phi_D'\rangle|$ unless the Schmidt rank of $|\phi_D\rangle$ is one.

for $T(\phi,\psi)^2 \leq 2/K$. Second, from the truncation error $\epsilon_N(D)$ for $|\psi_{2N}\rangle$,

$$T(\phi_D,\psi_{2N}) \geq (2^N-(D-1))p_N/2^N$$

for any $D$-MPS $|\phi_D\rangle$. Together, this shows that

$$D \geq 2^N(1-8T(\Phi_D,\chi_M))+1$$

which is exponential in the system size $M=2N^3$.

It is unclear how to make this example translational invariant. However, for the adjacent cases in Table 5.1, those examples exist: For $S\sim\log L$, take the preceding example and make it translational invariant by adding a tagging system $|10\ldots0\rangle^{\otimes N^2}$ and superposing all translations. The resulting state is hard to approximate as the translational invariance can be broken by local projections on the tags, and since the reduced state $\rho_N^L$ is the translational invariant mixture of the original, tagged reduced states, the entropy is increased by at most $\log L$. For the case $S_\alpha\sim$ const., $\alpha>1$, the state (5.6) with constant $p_N$ does the job.

The last two examples are of a different type: We consider $N$ spins on an ring and take maximally entangled pairs between opposite spins with a certain density, while the remaining qubits are in $|0\rangle$. The first of them provides a state with $S_\infty\sim\log L$ which is approximable. Therefore, distribute the maximally entangled pairs with density $\log N/N$. This state is an MPS with $D=2^{\log N}=N$, and for any $c>0$,

$$S_\infty(\rho_N^L) \geq \left\lfloor \tfrac{L}{N}\log N\right\rfloor \geq c\log L-1 \quad\text{for}\quad L\geq cN\ .$$

The example can be made translational invariant by taking the superposition of all translations: this increases the bond dimension by at most a factor of $N$ [111], while the largest eigenvalue of a block of length $N/\log N$ is $\frac{1}{2}$ (for the $|0\cdots0\rangle$ state), leaving the $\log L$ lower bound on the $S_\infty$ entropy unchanged.

The second example illustrates that for any $\kappa>0$, there is a state with $S_0\sim N^\kappa$ which cannot be approximated by MPS. Again, given $N$ qubits on a ring, take maximally entangled pairs between opposite spins with density $N^\kappa/N$ ($\kappa>0$), and set all other spins to $|0\rangle$. Then,

$$S_0(\rho_N^L) \leq N^\kappa L/N+1 \leq 2L^\kappa\ ,$$

while inapproximability follows from the superlogarithmic number of maximally entangled pairs. Translational invariance is achieved by taking the superposition of all translations for $\kappa'<\kappa$. The spectrum of a block of length $N/N^{\kappa'}$ is broadened to $(\frac{1}{2},\frac{N^{\kappa'}}{2N},\ldots,\frac{N^{\kappa'}}{2N})$: this clearly increases the truncation error, and the entropy scaling gets a log correction

$$S_0(\rho_N^L) \leq 2L^{\kappa'}(1+(1-\kappa')\log L)$$

which is bounded by $4L^\kappa$ for properly chosen $\kappa'$ and $L$.

## 5.6 Hardness of simulating time evolution

Let us now prove the hardness of simulating time evolutions with MPS-based approaches, using the results obtained. To this end, take a spin $\frac{1}{2}$ chain of $N$ spins which is initially in the ground state for infinite magnetic field,

$$|\psi(t=0)\rangle = |0\dots 0\rangle \ ,$$

and apply a critical Ising Hamiltonian

$$\mathcal{H} = -\frac{1}{2}\sum_j \left[\sigma_j^x \sigma_{j+1}^x + \sigma_j^z\right]$$

with periodic boundary conditions. There is good evidence [112] that in this case the block entropy of any block grows linearly in time, and indeed, a lower bound

$$S(\rho_N^L(t)) \geq 8t/3\pi + O(\log t)$$

for $t \leq L/3$, $L \leq N/2$, can be rigorously proven [113]. By plugging this into (5.5) with $L = 3t$, one finds that

$$\log D \geq \left(\frac{8}{3\pi} - \frac{3\delta}{2}\right) t + O(\log t) \ .$$

Thus, for an error $\delta < 16/9\pi \approx 0.57$, the required bond dimension, and therefore the effort to simulate the time evolution using MPS, grows exponentially in time.

## 5.7 Smooth Rényi entopies

At the end of this chapter, let us note that more refined criteria for approximability can be found: e.g., one can consider *smooth Rényi entropies* [114]

$$S_\alpha^\epsilon(\rho) = \min\{S(\sigma) : \|\rho - \sigma\|_{\mathrm{tr}} \leq \epsilon\}$$

and adapt the approximability proof of [26] to show that if

$$S_\alpha^{1/N^{1+\epsilon}}(\rho_N^L) \leq c \log N$$

for some $\alpha < 1$, $\epsilon > 0$, and $c > 0$, then approximability follows. This criterion is indeed more powerful: The state (5.6) with $p_N = 1/N^2$ has linearly growing Rényi entropies and is thus in the "inconclusive" region, while looking at smooth entropies shows its approximability. As smooth entropies are a lower bound on their non-smooth counterpart, this is in fact the only region where their scaling will imply approximability.

## 5.8 Conclusions

We have fully expored the connection between the scaling of block entropies and the approximability of families of states by MPS. We found that approximability is implied by a logarithmic scaling of $S_\alpha$ for $\alpha < 1$, while non-approximability follows whenever $S_\alpha \sim L^\kappa$ for $\alpha > 1$ and $\kappa > 0$, or from a linear growth of the von Neumann entropy. In all other cases, no conclusive statement can be made by simply looking at the scaling of block entropies. Except for the case of constant von Neumann entropy, this is true even under translational invariance. We then applied these results to prove that even the evolution of completely translational invariant and time independent systems cannot by simulated efficiently using MPS, and finally showed that a more refined criterion is obtained from looking at smooth Rényi entropies.

Open questions remain, in the first place whether translational invariant states with bounded von Neumann entropy can be approximated efficiently by MPS. (Note that this is not true for approximability w.r.t. extensive observables, see Footnote 1 on pg. 85). Furthermore, it would be interesting to find other and more refined criteria for approximability. As we have shown, one such criterion is obtained by looking at smooth entropies, but there are more possibilities: one can, e.g., consider the joint scaling behaviour of several entropies or other spectral properties of the reduced states.

# Chapter 6

# The computational complexity of PEPS

## 6.1 Introduction

As we have seen in the preceding chapter, Matrix Product States (MPS) prove very useful to describe the properties of correlated quantum many-body systems. Indeed, the Density Matrix Renormalization Group (DMRG) method [17], which has been extremely successful in the description of one-dimensional phenomena, can be interpreted as a variational method over the class of MPS [18]. MPS structure the exponentially large state space into a hierarchy of states with polynomial description complexity [111], and it turns out that already the lowest levels of this hierarchy approximate many physical states of interest extremely well. MPS have a natural extension to two and higher dimensional lattices, called Projected Entangled Pair States (PEPS), which also have an efficient description and are promising candidates for variational methods in higher dimensions [24]. It has been shown that MPS can be created efficiently by a quantum computer [27], and that they also can be simulated efficiently classically [21]. In contrast, in two or more dimensions it seems to be hard to create arbitrary PEPS, as well as to classically compute expectation values. In fact, it has been demonstrated that there exist 2D PEPS which encode solutions to NP-complete problems [28], thus posing lower bounds on their complexity and computational power.

In this chapter, we determine both the power of creating PEPS and the complexity of classically simulating them. We investigate which kind of problems we could solve if we had a way to efficiently create PEPS, and find that these are exactly the problems in the complexity class PP (deciding whether a boolean formula has more satisfying than non-satisfying assingments). Second, we show that classically computing local expectation values on PEPS is a #P-complete problem (counting the satisfying assignments of a boolean formula). This result can be extended to the contraction of arbitrary tensor networks, which turns out

to be #P-comlete as well.

The main tool in our proofs is a duality between PEPS and postselection, which permits to use existing results from quantum complexity [115]: any PEPS can be created by a postselected quantum circuit, and any output of such a circuit can be written as a PEPS. We also apply this duality to show that ground states of gapped local Hamiltonians in $D$ dimensions can be efficiently approximated by the boundary of a $D+1$-dimensional PEPS. Finally, we compare the power of creating PEPS to the power of creating ground states of local Hamiltonians. While in general they are equally hard, we find that when restricting to gapped Hamiltonians, creating ground states becomes easier: it is in the weaker class QMA, the quantum analogue of NP.

## 6.2 PEPS and postselection

We start by recalling the definition of PEPS [116]. Consider an arbitrary undirected graph where each of the vertices corresponds to a quantum system (a spin) of Hilbert space dimension $d$. A PEPS on these $N$ spins is constructed by placing as many virtual spins of dimension $D$ on each vertex as there are adjacent edges. Along each edge, these virtual spins form maximally entanged states $\sum_{i=1}^{D} |i\rangle|i\rangle$. The physical spins are now obtained from the virtual ones by applying a linear map $P^{[v]} : \mathbb{C}^D \otimes \ldots \otimes \mathbb{C}^D \to \mathbb{C}^d$ at each vertex $v$. For the sake of readability, we will mostly supress the dependence of $P$ on $v$. The graph underlying the PEPS will usually be chosen according to the physical setup, typically a two or higher dimensional lattice.

Let us now turn to *postselected quantum circuits* [115]. Roughly speaking, postselection means that we can measure a qubit with the promise of obtaining a certain outcome. More precisely, the postselected circuits we consider start from the $|0\cdots 0\rangle$ state, perform a sequence of unitary one- and two-qubit gates, and postselect on the first qubit being $|0\rangle$. Thereby, the state $\alpha|0\rangle|\phi_0\rangle + \beta|1\rangle|\phi_1\rangle$ is projected onto the state $|\phi_0\rangle$, which is the state created by the postselected quantum circuit.[1] Note that a state with $\alpha = 0$ will not be considered a valid input.

In the following, we show that the output of a postselected quantum circuit can be expressed efficiently as a PEPS on a 2D square lattice with both $D = d = 2$. We start by briefly recalling the concept of measurement based quantum computation [118, 119]: One starts from the 2D cluster state (which is a PEPS with $D = d = 2$ [116]) and implements the quantum circuit by a sequence of projective measurements on the individual spins. Finally, the output is found

[1] We do not impose polynomial-size and uniformity conditions on the circuit, which would yield a natural extension PostΨP of the class ΨP defined in [117]. The reason is that we will show that there exists a uniform and efficient *transformation* between PEPS and postselected quantum states, although none of the two has to satisfy any efficiency or uniformity condition.

in the unmeasured qubits, up to Pauli corrections which depend on the previous measurement outcomes. In order to express the output of a postselected circuit as a PEPS, we therefore start by implementing its unitary part in the measurement based model. We do this by projecting each qubit on the outcome $|a\rangle$ which does *not* give a Pauli correction, by replacing the original cluster projector $P_C$ with $|a\rangle\langle a|P_C$. This leaves us with a set of qubits holding the output of the circuit, and by projecting the first qubit on $|0\rangle$, we obtain the output of the postselected quantum circuit. The transformation between the representations can be carried out efficiently, and the resulting PEPS has a size polynomial in the length of the circuit.

Conversely, any PEPS can be efficiently created by a postselected quantum computer. This holds for PEPS on an arbitary graph with degree (the maximum number of edges adjacent to a vertex) at most logarithmic in the system size, which ensures that the $P$'s are polynomial-size matrices. The key point is that any linear map $P$ can be implemented deterministically using postselection. To this end, append rows or colums of zeros to make $P$ a square matrix $\tilde{P}$. By appropriate normalization, we can assume w.l.o.g. that $P^\dagger P \leq \mathbb{1}$. Hence, there exists a unitary $U$ on the original system and one ancilla such that $\langle 0|_{\text{anc}} U |0\rangle_{\text{anc}} = \tilde{P}$. This is, by adding an ancilla $|0\rangle_{\text{anc}}$, performing $U$ and postselecting the ancilla we can implement $\tilde{P}$. In order to generate a PEPS using postselection, we thus have to encode each of the virtual spins in $\lceil \log D \rceil$ qubits, create the maximally entangled pairs, and implement the $U$'s corresponding to the maps $P$, which can be all done efficiently. We are thus left with $N$ ancillas, all of which we have to postselect on $|0\rangle$. This, however, can be done with a single postselection by computing the OR of all ancillas into a new ancilla and postselecting it on $|0\rangle$.

In summary, on the one side we have that any postselected quantum circuit can be translated efficiently into a 2D PEPS with $D = d = 2$, while conversely there is also an efficient transform from any PEPS to a postselected quantum circuit. In turn, this shows that all the features and the full complexity of PEPS can already be found in the simplest case of two-dimensional PEPS, making them an even more interesting subject for investigations.

## 6.3 The power of creating PEPS

Let us first briefly introduce the complexity classes #P and PP [120]. Consider an efficiently computable boolean function $f : \{0,1\}^N \to \{0,1\}$, and let $s \equiv s(f) := |\{x : f(x) = 1\}|$ be the number of satisfying assignments. Then, finding $s$ defines the counting class #P, while determining whether $s \geq 2^{n-1}$ (i.e., finding the first bit of $s$) defines the decision class PP. This class contains NP and BQP as well as QMA, the quantum version of NP.

First, we investigate the computational power of creating PEPS. More precisely, we consider the scenario of Fig. 6.1: We want to know which decision

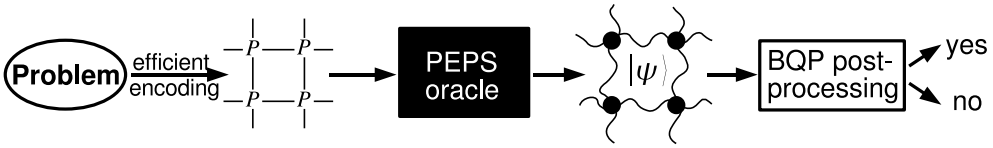


**Figure 6.1:** The power of creating PEPS: The original decision problem is transformed into a PEPS description by a polynomial-time algorithm. The black box creates the corresponding quantum state, and an efficient quantum postprocessing returns the solution. Which kind of problems can we solve this way?

problems we can solve with one use of a PEPS oracle, i.e., a black box which creates the quantum state from its classical PEPS description, together with efficient classical pre-processing and quantum post-processing.

We now use the PEPS–postselection duality to show that the power of creating PEPS equals PP. It has been shown that PostBQP—the class of decision problems which can be solved by a postselected quantum computer—equals PP, PostBQP = PP [115]. This readily implies that a PEPS oracle allows us to solve PP problems instantaneously by preparing the output of the postselected circuit as a PEPS and just measuring one output qubit in the computational basis. On the other hand, this is the best we can do with a single use of the PEPS oracle, since every PEPS can be generated efficiently by a postselected quantum computer. BQP postprocessing instead of a simple one-qubit measurement does not increase the computational power, since it commutes with the postselection and can thus be incorporated in the PEPS.

The fact that creating PEPS allows to solve PP-complete problems strongly suggests the existence of PEPS which cannot be created efficiently by a quantum computer. Note however that the states which appear in the PP-hardness proof above are not of this type: once the corresponding counting problem is solved, they can be easily constructed. While it appears very unlikely that all PEPS can be constructed efficiently from some normal form (it would imply QMA = QCMA and BQP/qpoly = BQP/poly [121]), an example of such a state is still missing.

## 6.4 The classical complexity of PEPS

Let us now investigate the complexity of classically simulating PEPS, and its generalization, the contraction of tensor networks. For the case of PEPS, there are at least three possible definitions of the problem: compute the normalization of the PEPS (NORM), compute the unnormalized expectation value of some observable (UEV), and compute the normalized expectation value (NEV). Since they can be transformed easily into each other,[2] we will use whichever is most

[2] UEV of $\mathbb{1}$ gives NORM, while UEV of an operator $A$ is obtained by applying NORM twice, $\langle\psi|A|\psi\rangle = \langle\tilde{\psi}|\tilde{\psi}\rangle - \|A\|\langle\psi|\psi\rangle$. Here, $|\tilde{\psi}\rangle$ is derived from $|\psi\rangle$ by replacing the relevant $P$ by

appropriate.

We first show that contracting PEPS is #P-hard, i.e., that for any (polynomial) boolean function $f$, $s(f)$ can be found by simulating a PEPS. Therefore, we take a quantum circuit which creates $\sum_x |x\rangle_A |f(x)\rangle_B$ and encode it in a PEPS. Then, the normalized expectation value of $\sigma_z$ of $B$ allows to compute $s(f)$.

To show that the simulation of PEPS is inside #P, we have to show that the normalization of the PEPS, or equivalently the success probability for the postselection, can be computed by counting the satisfying assigments of some boolean function. This can be done by adapting well-established quantum complexity techniques (see [115] and references therein): First, approximate the postselected circuit using only Toffoli and Hadamard gates [122, 123]. The probability $p_x$ for a state $|x\rangle$ before postselection is obtained as a kind of path integral [124], by summing the amplitudes for all possible "computational paths" $\zeta = (\zeta_1, \ldots, \zeta_{T-1})$, where $|\zeta_t\rangle$ is the state step $t$ and $T$ the length of the circuit:

$$p_x = \Big| \sum_{\zeta} \alpha_{x,\zeta} \Big|^2 = \sum_{\zeta,\zeta'} \alpha_{x,\zeta} \alpha^*_{x,\zeta'} \ ,$$

with $\alpha_{x,\zeta}$ a product over transition amplitudes $A_{\zeta_t \to \zeta_{t+1}}$ along the path $\zeta$. The normalization of the PEPS is obtained as the sum over all states where the postselection succeeds, $\sum_{\bar{x}} p_{(0,\bar{x})}$. This can be rewritten as the sum over an efficiently computable function

$$f(\bar{x}, \zeta, \zeta') = \alpha_{(0,\bar{x}),\zeta} \alpha^*_{(0,\bar{x}),\zeta'}$$

which takes values in $\{0, \pm 1\}$, as the circuit consisted only of Toffoli and Hadamard gates. Now this sum can be computed by counting the satisfying assigments of the function

$$f_{\mathrm{bool}}(\xi, z) := (f(\xi) \geq z) \ , \quad z \in \{0, 1\} \ ,$$

which shows that the simulation of PEPS is in #P. Together, we find that the classical simulation of PEPS is #P-complete under weakly parsimonious reductions (see Footnote 2).

It is natural to ask whether this also shows that contracting general tensor networks is in #P. For a tensor network $T$, let us denote its contraction by $\mathcal{C}(T) \in \mathbb{C}$. Since the contraction of PEPS is a special case, it is clear that the problem is #P-hard. To place it within #P, observe first that $|\mathcal{C}(T)|^2 = \mathcal{C}(T \otimes T^*)$ can be found by attaching a physical system of dimension one to each site and computing the normalization of the resulting PEPS. To determine the phase of

$(A + \|A\| \mathbb{1})^{1/2} P$. Clearly, UEV and NORM allow to compute NEV. Conversely, to compute the norm of a PEPS write it as a quantum circuit, but stop before the postselection. Then, its norm equals the NEV of $\mathrm{diag}(1, 0)$ on the qubit to be postselected, which equals the NEV on a PEPS. All reductions are weakly parsimonious: problem A can be solved by *one* call to problem B, with efficient pre-processing of the input *and* post-processing of the output. Note that two (or more) parallel #P-queries can be encoded in a single one, by considering $h(x, y, b)$, defined as $f(x)$ for $b = 0$ and $g(y)$ for $b = 1$, $x = 0$.

$\mathcal{C}(T)$, observe that $\mathcal{C}(T \oplus T') = \mathcal{C}(T) + \mathcal{C}(T')$. Thus, by setting $T' = T^*$, we get $|\mathrm{Re}(\mathcal{C}(T))|$, while the sign can be determined by adding another $T'' \equiv c > 0$. This proves that contracting tensor networks is #P-complete.

The obtained hardness results are stable under approximations. To see why, note that any counting problem can be reduced to any of our three primitives with only linear postprocessing, and thus approximating these primitives is as hard as approximating counting problems can be. For NEV, this again works by preparing $\sum |x\rangle_A |f(x)\rangle_B$ and computing the expectation value of $B$. For NORM and thus UEV, note that the output of any normal quantum circuit and thus $\sum |x\rangle_A |f(x)\rangle_B$ has a known norm when written as a PEPS, since the success probability of each cluster projector is known, and the probability of the two measurement outcomes in the cluster is unbiased [119]. Thus, the probability for $|1\rangle_B$ can be readily determined from the norm of the PEPS where we postselected on $|1\rangle_B$.

## 6.5 PEPS and ground states

The interest in MPS and PEPS stems mainly from the fact that those states perform extremely well in approximating ground states. In the following, we use the PEPS–postselection duality, and a relation between postselection and cooling, to shed new light on the connection between PEPS and ground states. In particular, we show that the unique ground state of a gapped Hamiltonian on a $D$-dimensional lattice can be approximated efficiently by the border of a PEPS with $D + 1$ dimensions.

Consider a Hamiltonian on $N$ spins, $H = \sum_i H_i$, where each $H_i$ acts on a finite number of spins, with a unique ground state and a polynomial energy gap $\Delta \geq 1/\mathrm{poly}(N)$. Starting from a random state $|\chi\rangle$, the ground state can be efficiently approximated via $|\psi_0\rangle \approx \exp[-\beta H]|\chi\rangle$. The imaginary time evolution can in turn be approximated using the Trotter decomposition, which only requires operations $\exp[-\beta/N H_i]$ acting on finitely many spins. Since those operations are linear, they can be implemented using postselection, and we see that postselection can be used to cool into the ground state. By embedding the postselected cooling procedure in a PEPS, the ground state of any gapped $N$-particle Hamiltonian can be approximated up to $\epsilon$ by the boundary of a PEPS, where the extra dimension has depth $M \sim \mathrm{poly}(N, 1/\epsilon)$.[3] In case the $H_i$ are local, the PEPS can be simplified considerably since any local linear operation can be implemented directly on the level of the PEPS without the need for ancilla qubits.

[3] One might object that the performance of 1D variational methods is much better. However, there are several differences: Our method works for any dimension and for non-local Hamiltonians, it is constructive, it does not break translational symmetry, and it implements the complete evolution $\exp[-\beta H]$.

## 6.6 The power of creating ground states

As we have seen, PEPS can encapsulate problems as hard as PP. However, these PEPS are quite artificial, while in practice one is often interested in PEPS in connection with ground states. Therefore, let us have a look at the computational power of a ground state oracle, i.e., a black box which creates the ground state from the Hamiltonian.

First, let us introduce the complexity class QMA [125]. Colloquially, QMA is the quantum version of NP, i.e., it contains all decision problems where for the "yes" instance, there exists an efficiently checkable *quantum* proof, while there is no proof for any "no" instance. In a seminal work, Kitaev [125, 126] has shown that the problem of determining ground state energies of local Hamiltonians up to polynomial accuracy is QMA-complete. More precisely, in LOCAL HAMILTONIAN one is given an $N$-qubit local Hamiltonian $H = \sum H_i$ with the promise that the ground state energy $E_0 < a$ or $E_0 > b$, $b - a > 1/\mathrm{poly}(N)$, and the task is to decide whether $E_0 < a$. Clearly, the ground state of $H$ serves as a proof for a "yes" instance. In successive works, the class of Hamiltonians has been restricted down to two-particle neareast neighbor Hamiltonians on a 2D lattice of qubits [127].

Let us briefly reconsider our cooling protocol in the light of QMA. It is easy to see that the QMA proof need not necessarily be the ground state, as long as it is close enough in energy (depending on the verifier). Since our cooling protocol suppresses higher energy levels exponentially, the correspondence between post-selection and cooling shows that a postselected quantum computer can be used to create proofs for QMA problems, or differently speaking, that any QMA proof can be efficiently expressed as a PEPS.

In the following, we give some observations which indicate that creating ground states of gapped Hamiltonians is easier than creating PEPS. First, note that a ground state oracle for arbitrary Hamiltonians is still as powerful as PP. To see why, take a PP problem and encapsulate it in a PEPS. By perturbing the $P$'s randomly by a small amount, one obtains a PEPS which is the unique ground state of a local Hamiltonian, which can be derived efficiently from the $P$'s [111, 128]. This shows that an unrestricted ground state oracle enables us to solve PP problems. However, the gap $\Delta$ of the above Hamiltonian will be exponentially small: if not, one could add a small penalty, say $\Delta/100$, on the "answer" qubit, and use that the original Hamiltonian has ground state energy $E_0 = 0$: Then, determining the value of that qubit could be solved in QMA, thus proving QMA = PP which is considered unlikely [129].

Since ground states of general Hamiltonians are not easier to create than PEPS, let us now assume an oracle which only works for local Hamiltonians with a unique ground state, known ground state energy, and a polynomial spectral gap to the first excited state. (Alternatively, one could consider "proof oracles" for the LOCAL HAMILTONIAN problem.) It is easy to see that this restricted oracle,

even with BQP postprocessing, is at most as powerful as QMA. The proof is the ground state, and the verifier is constructed as follows. Let $V_1$ be the verifier for the ground state, it accepts the ground state with $p_{\mathrm{GS}}$, and any excited state with probability at most $p_{\mathrm{ES}} = p_{\mathrm{GS}} - \Delta$, $\Delta = 1/\mathrm{poly}(N)$. Further, let $V_2$ be the postprocessing circuit which has a polynomial separation between the "yes" and the "no" answer if applied to the ground state, $p_{\mathrm{yes}} = 1/2 + \delta$, $\delta = 1/\mathrm{poly}(N)$. Take $Q = \frac{\Delta/2+1}{\Delta+1}$, and construct the complete verifier as follows: with probability $Q$, run $V_1$, and with $(1-Q)$, run $V_2$. One can readily check that this gives a polynomial separation between the cases where the proof is the ground state *and* the postprocessing return "yes", and the cases where either the proof is not the ground state or the postprocessing returns "no". The same strategy can be used to show that a PEPS oracle cannot be tested on all inputs unless QMA = PP: Otherwise, one could take a PP-hard PEPS and construct a verifier which either runs the testing routine or reads out the PP solution.

These observations show that imposing a constraint on the spectral gap of a Hamiltonian has direct implications on its computational complexity, and we think that the complexity properties of gapped Hamiltonians are worth being considered. On the one side, in the above scenario it is not clear whether all QMA problems can be solved using this oracle, on the other side, it is not clear how important the knowledge of the ground state energy is—note that we however also had this knowledge in the PP-hard case. It is also an interesting question whether the problem LOCAL HAMILTONIAN remains QMA-complete when restricting to polynomially gapped Hamiltonians. If not, GAPPED LOCAL HAMILTONIAN should be a natural candidate for a physically motivated class of problems weaker than QMA.

# Chapter 7

# Gaussian Matrix Product States

## 7.1 Introduction

As we have seen in the previous chapters, Matrix Product States (MPS) and their higher dimensional generalization, Projected Entangled Pair States (PEPS), provide an efficient and useful way to describe quantum many-body systems. Also, we have seen in Chapter 4 that lattices of harmonic oscillators interacting via a quadratic Hamiltonian can both be used to describe physical scenarios as e.g. vibrational modes in solid state systems or ion traps, and can serve as an accessible model for more complicated system due to their efficient description in terms of Gaussian states. In this chapter, we introduce Gaussian Matrix Product States (GMPS), a generalization of MPS and PEPS to the case of harmonic lattices.

We define GMPS as projected entangled pairs, and prove that every (translation invariant) Gaussian state can be represented as a (translation invariant) GMPS. We then investigate how much entanglement is really needed in the bonds, since different from the finite dimensional case, a maximally entangled bond would carry an infinite amount of entanglement. We continue by discussing correlation functions and show that they can be computed efficiently from the GMPS representation; for the case of pure one-dimensional GMPS with one mode per site, we prove that the correlations decay exponentially (as it is the case in finite dimensions) and explicitly derive the correlation length. Finally, again in analogy to the finite dimensional case, we show that every GMPS is the ground state of a local Hamiltonian.

Throughout this chapter, we will use the basic concepts and notations for Gaussian states and lattices as introduced at the beginning of Chapter 4.

## 7.2 Definition of Gaussian MPS

We start by defining Gaussian matrix product states (GMPS). The definition resembles the physical interpretation of finite-dimensional matrix product states as projected entangled pairs. In finite dimensions, MPS can be described by taking maximally entangled pairs of dimension $D$ between adjacent sites, and applying arbitrary local operations on each site, mapping the $D \times D$ input to a $d$-dimensional output state. Similarly, GMPS are obtained by taking a number of entangled bonds and applying local (not necessarily trace-preserving) operations $\mathcal{T}^{[i]}$, where the boundary conditions can be taken either open or closed. Any GMPS is completely described by the type of the bonds and by the operations $\mathcal{T}^{[i]}$. Note that this construction holds independent of the spatial dimension. For one dimension, it is illustrated in Fig. 7.1. As matrix product states are frequently used to describe translationally invariant systems, an inportant case is given if all maps are identical, $\mathcal{T}^{[i]} = \mathcal{T}\ \forall i$.

In order to define MPS in the Gaussian world, we have to decide on the type of the bonds as well as on the type of operations. We choose both the bonds to be Gaussian states and the operations to be Gaussian operations, i.e., operations mapping Gaussian inputs to Gaussian outputs. For now, we will take the bonds to be maximally entangled (i.e., EPR) states, such that the only parameter originating from the bonds is the number $M$ of EPRs. We show later on how the case of finitely entangled bonds can be easily embedded.

As to the operations, we will allow for arbitrary Gaussian operations. Operations of this type are most easily described by the Jamiolkowski isomorphism [130]. There, any Gaussian operation $\mathcal{T}$ which maps $N$ input modes to $M$ output modes can be described by an $N+M$ mode covariance matrix $\Gamma$ with block $B$ (input) and $C$ (output). The corresponding map on some input state $\gamma_{\text{in}}$ in mode $A$ is implemented by projecting the modes $A$ and $B$ onto an EPR state as shown in Fig. 7.2, such that the output state $\mathcal{T}(\gamma_{\text{in}})$ is obtained in mode $C$. Conversely, the matrix $\Gamma$ which represents the channel $\mathcal{T}$ is obtained by applying

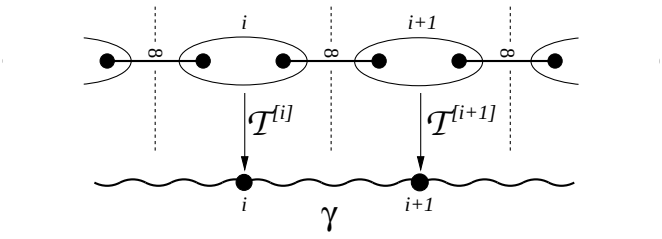


**Figure 7.1:** Construction of Gaussian Matrix Product States (GMPS). GMPS are obtained by taking a fixed number $M$ of maximally entangled (i.e., EPR) pairs shared by adjacent sites, and applying an arbitrary $2M$ to 1 mode Gaussian operation $\mathcal{T}^{[i]}$ on site $i$.

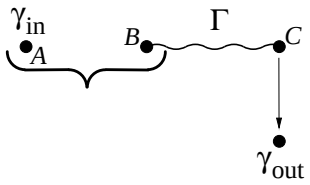


**Figure 7.2:** The Jamiolkowski isomorphism. The Gaussian channel described by the state $\Gamma$ can be implemented by projecting the input state $\gamma_{\rm in}$ (mode $A$) and the input port of $\Gamma$ (mode $B$) onto the EPR state (symbolized by curly brackets). In case of success, the output is obtained in mode $C$. The operation can be made trace-preserving by measuring in a basis of displaced EPR states, and displacing $C$ according to the measurement outcome.

the channel to one half of a maximally entangled state. The duality between $\mathcal{T}$ and $\Gamma$ is most easily understood in terms of teleportation, and shows that this characterization encompasses all Gaussian operations. Note that the protocol of Fig. 7.2 can be always made trace-preserving by projecting onto the set of phase-space displaced EPR states and correcting the displacement of mode $C$ according to the measurement outcome [69].

In the following, we will denote all maps $\mathcal{T}$ by their corresponding CM $\Gamma$. Sometimes, we will speak of the modes $B$ and $C$ as input and output ports of $\Gamma$, respectively.

We now discuss how the covariance matrix of the output will depend on the CM of the input and on the channel $\Gamma$ [68, 69]. This is most easily computed in the framework of characteristic functions [66]. The characteristic function of the output is given by

$$\chi_C(\xi_C) \propto \int e^{-\xi_A^T \gamma_{\rm in} \xi_A} e^{-\xi_{BC}^T \Gamma \xi_{BC}} \delta(x_A - x_B) \delta(p_A + p_B) \mathrm{d}\xi_{AB} \ ,$$

and by integrating out subsystem $A$,

$$\chi_C(\xi_C) \propto \int e^{-\xi_{BC}^T M \xi_{BC}} \mathrm{d}\xi_B \ ,$$

with

$$M = \left( \begin{array}{cc} \theta\gamma\theta + \Gamma_B & \Gamma_{BC} \\ \Gamma_{CB} & \Gamma_C \end{array} \right) \ .$$

Basically, the integration $\int \mathrm{d}\xi_A \delta(x_A - x_B)\delta(p_A + p_B)$ does the following: first, it applies the partial transposition $\theta \equiv \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ to one of the subsystems, and second, it collapses the two systems $A$ and $B$ in the covariance matrix by adding the corresponding entries. The integration over $\xi_B$, one the other hand, leads to a state whose CM is the Schur complement of $M_{11}$, $M_{22} - M_{21} M_{11}^{-1} M_{12}$, such that the output state is described by the CM

$$\gamma_{\rm out} = \Gamma_C - \Gamma_{CB} \frac{1}{\Gamma_B + \theta \gamma_{\rm in} \theta} \Gamma_{BC} \ .$$

Let us briefly summarize how to perform projective measurements onto the EPR state in the framework of CMs, where we denote the measured modes by $A$ and $B$, while $C$ is the remaining part of the system. First, apply the partial transposition to $B$, second, collapse $A$ and $B$, and third, take the Schur complement of the collapsed mode $AB$, which gives the output CM of $C$.

As we discuss Gaussian matrix product states in connection with ground states of Hamiltonians, we are mainly interested in pure GMPS. Particularly, a GMPS is pure if the $\Gamma^{[i]}$ which describe the operations $\mathcal{T}^{[i]}$ are taken to be pure, which we assume from now on.

Let us finally emphasize that the given defintion of MPS holds independent of the spatial dimension of the system, as do most of the following results, and in fact applies to an arbitrary graph.

## 7.3 Completeness of Gaussian MPS

In the following, we show that any pure and translational invariant state can be approximated arbitrarily well by translational invariant Gaussian matrix product states, i.e., GMPS with identical local operations $\mathcal{T}$. (Without translational invariance, this is clear anyway: the complete state is prepared locally and teleported to its destination using the bonds.) The proof is presented for one dimension, but can be extended to higher spatial dimensions.

Given a translational invariant state $\gamma$, there is a translational invariant Hamiltonian $H$ which transforms the separable state $\mathbb{1}$ into $\gamma$, $\gamma = SS^T$, $S = e^{\sigma H}$. It has been shown [131] that this time evolution can be approximated arbitrarily well by a sequence of translational invariant local (one-mode) and nearest neighbor (two-mode) Hamiltonians $H_j$,

$$e^{\sigma H} \approx \prod_{j=1}^{J} e^{\bigoplus_n \sigma H_j} \ , \tag{7.1}$$

where the $H_j$ act on one or two modes, respectively, and approach the identity for growing $J$.

Clearly, translational invariant local Hamiltionians can be implemented by local maps without using any EPR bonds. In the following, we show how translational invariant nearest-neighbor interactions can be implemented by exploiting the entanglement of the bonds. The whole procedure is illustrated in Fig. 7.3 and requires two EPR pairs per site. We start with some initial state $\gamma_{\text{in}}$ onto which we want to apply $S_{\oplus} = e^{\bigoplus \sigma H_j} \approx \mathbb{1} + \bigoplus_n \sigma H_j$.

First, we perform local EPR measurements between the modes of $\gamma_{\text{in}}$ and one of the bonds in order to teleport the modes of $\gamma_{\text{in}}$ to the left, cf. Fig. 7.3a. Then, the infinitesimal symplectic operation $S = e^{\sigma H_j}$ is applied to the left-teleported mode and the second bond, Fig. 7.3b. In the last step, another EPR measurement

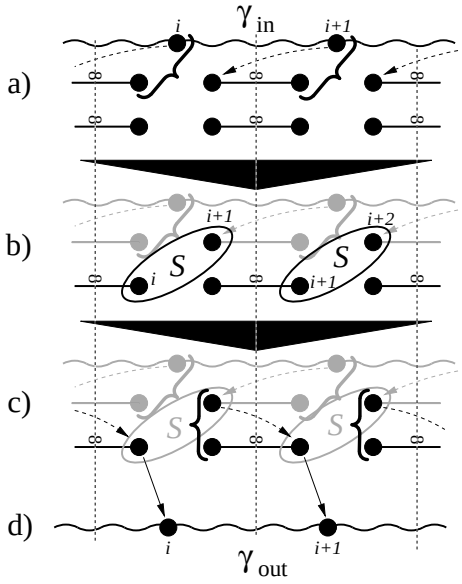


**Figure 7.3:** Implementation of a translational invariant nearest neighbor Hamiltonian in a translational invariant fashion. Starting from $\gamma_{\text{in}}$, the input is first teleported to the left, then, the infinitesimal time evolution $S = e^{\sigma H}$, $H \ll 1$, is performed, and finally, the state is teleported back.

is performed which teleports the left-teleported mode back to the right, and "into" the mode on which the adjacent $S$ was applied. As the operations $e^{\sigma H_j} \approx \mathbb{1} + \sigma H_j$ all commute, the "nested" application of the nearest neighbor symplectic operations $S$ indeed give $S_\oplus$. Thus, the remaining mode indeed contains the output $\gamma_{\text{out}} = S_\oplus \gamma_{\text{in}} S_\oplus^T$. The whole decomposition (7.1) can be implemented by iterated application of the whole protocol of Fig. 7.3.

## 7.4 GMPS with finitely entangled bonds

In this section we show that in general, infinitely entangled bonds can be replaced by finitely entangled ones. Intuitively, this should be possible whenever the channel $\mathcal{T}^{[i]}$ destroys some of the entanglement of the bond anyway, i.e., $\Gamma^{[i]}$ is non-maximally entangled. In that case, it should be possible to use a less entangled bond while choosing a channel which does not destroy entanglement any more.

The method is illustrated in Fig. 7.4. Again, for reasons of clarity we restrict to one dimension and one bond. The argument however appies independent of the spatial dimension and the number of bonds. The only restriction we have to make is the restriction to pure GMPS, i.e., those with pure $\Gamma^{[i]}$.

Consider a GMPS with local channels given by $\Gamma^{[i]}$ and infinitely entangled bonds, Fig. 7.4a. First, apply a Schmidt decomposition [132] to $\Gamma^{[i]}$ in the partition $A|BC$, which can be always done as long as $\Gamma^{[i]}$ is pure. The Schmidt decomposition allows us to rewrite the state as shown in Fig. 7.4b—an entangled state between modes $A$ and $C$ with two-mode squeezing $r^{[i]}$, $B$ in the coherent state $\mathbb{1}$, and sympectic operations $S_A^{[i]}$ and $S_{BC}^{[i]}$ which are applied to modes $A$ and $BC$, respectively. As the bond itself is infinitely entangled, we can tele-

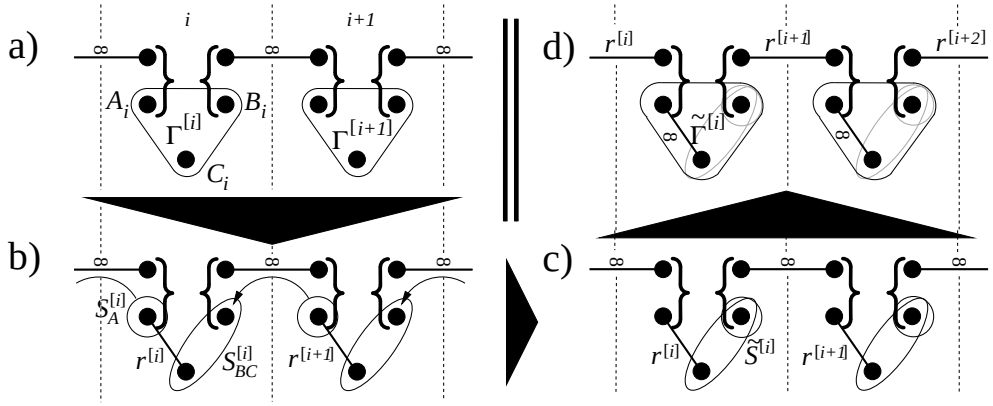


**Figure 7.4:** How to make the bonds of GMPS finitely entangled. **a)** The initial MPS. **b)** Do a Schmidt decomposition of the original map $\Gamma$. **c)** Move the $S_A^{[i]}$ through the infinitely entangled bond to the next site. **d)** Swap the finitely and the infinitely entangled pair.

port the sympectic operation through the bond to the next site as indicated in Fig. 7.4b. Then, $S_A^{[i+1]}$ can be merged with $S_{BC}^{[i]}$ to a new operation $\tilde{S}^{[i]}$ acting on modes $B$ and $C$ of site $i$ (Fig. 7.4c). Finally, in the triples consisting of one maximally entangled state, one non-maximally entangled state, and the projection onto the EPR state, the maximally and the non-maximally entangled state can be swapped, resulting in Fig. 7.4d. There, we have finitely entangled bonds, while the infinite entanglement has been moved into the new maps $\tilde{\Gamma}^{[i]}$.

It is tempting to apply this construction to the completeness proof of the preceding section in order to obtain a construction which is less wasting with respect to resources. However, for any iterative protocol this is most likely difficult to achieve. The reason for this is found in the no-distillation theorem which states that with Gaussian operations, it is not possible to increase the amount of entanglement [69] between two parties. Particularly, this implies that in each step of an iterative protocol, the bonds need to have at least as much entanglement as can be obtained at the output of this step, maximized over all inputs where the entanglement is increased. This is indeed a severe restriction, although it does not imply the impossibility of such a protocol. One could, e.g., create a highly entangled state in the first step and then approach the desired state by decreasing the entanglement in each step. Still, it seems most likely that a sequence of MPS which approach a given state efficiently will have to involve more and more bonds simultaneously and thus cannot be constructed in an iterative manner.

## 7.5 Correlation functions of Gaussian MPS

In this section, we show how to compute correlations functions from the maps $\Gamma^{[i]}$ which describe the GMPS. We show that this can be done efficiently, i.e.,

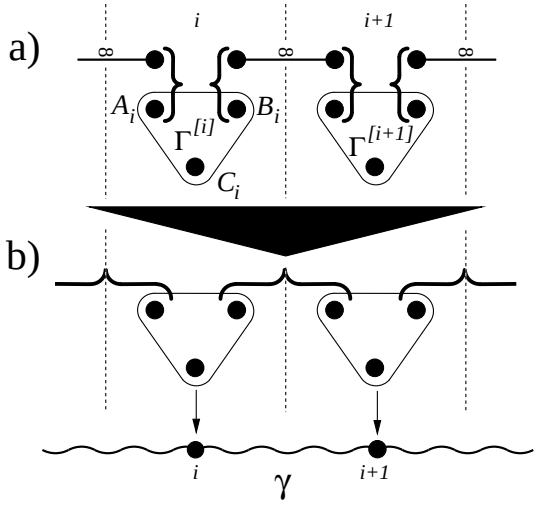


**Figure 7.5:** If the local operations are described by states $\Gamma^{[i]}$ via the Jamiolkowski isomorphism, the construction of GMPS can be simplified by replacing the measurement-bond-measurement triples by a simple projection onto the EPR state.

in polynomial time independent of the dimension of the graph which is different to the finite dimensional case. Of course, this is not too surprising as Gaussian states can be fully characterized by a number of paramaters quadratic in the number of modes.

Let us start with the general case of different $\Gamma^{[i]}$, as in Fig. 7.5a. The calculation can be facilitated by the simple observation that the triples consisting of two projective measurements and one EPR pair can be replaced by a single projection onto the EPR state, Fig. 7.5b. It follows that we can apply the formalism for projective measurements onto the EPR state which we presented in Sec. 7.2. We start from $\bigoplus_i \Gamma^{[i]}$. First we partially transpose all $B$ modes, then we collapse $A_{i+1}$ and $B_i$ for all $i$, and finally we take the Schur complement of the merged mode. In case of periodic boundary conditions, this can be expressed by the transformation matrix

$$\Pi = \begin{pmatrix} \mathbb{1}_A & \mathcal{R}\theta_B & 0 \\ 0 & 0 & \mathbb{1}_C \end{pmatrix} \tag{7.2}$$

which maps $ABC$ onto $A'C$, where $\theta_B \equiv \theta \otimes \mathbb{1}$ is the partial transposition on system $B$, and $\mathcal{R}$ is the circulant right shift operator, $(\mathcal{R})_{ij} = \delta_{i,j+1 \mod N} \otimes \mathbb{1}$. Then, the output state, i.e., the GMPS characterized by $\Gamma^{[i]}$, is

$$\gamma = \mathrm{SC}_{A'}\left[\Pi\left(\bigoplus_i \Gamma^{[i]}\right)\Pi^T\right] ,$$

where $\mathrm{SC}_X(U)$ is the Schur complement of the $X$ part of $U$, $\mathrm{SC}_X(U) = U_{YY} - U_{YX}U_{XX}^{-1}U_{XY}$. For fixed boundary conditions, the matrix $\Pi$ has to be modified accordingly at the boundaries. All the involved operations scale polynomially in the product $NM$ of the number of sites $N$ and the number of modes $M$.

In case all the local maps are chosen equal, $\Gamma^{[i]} \equiv \Gamma \ \forall i$, the above formula can be simplified considerably. Therefore, note that the Fourier transform can

be taken into the Schur complement, and that $\Pi$ as well as $\bigoplus_{i=1}^{N} \Gamma^{[i]} = \Gamma \otimes \mathbb{1}_N$ are blockwise circulant so that both are diagonalized by the Fourier transform. We adapt the notation used in Section 4.3 of writing the diagonal of the Fourier transformed matrices as functions of an angle $\phi$. In that case, $\Gamma \otimes \mathbb{1}$ is mapped onto the constant function $\Gamma$, and the same holds for $\mathbb{1}$ and $\theta$ in (7.2). The right shift operator $\mathcal{R}$, on the other hand, is transformed to $e^{i\phi}\mathbb{1}$: the EPR measurement performed between adjacent sites leads to a complex phase of $\phi$. Altogether, we have

$$\hat{\Pi} = \begin{pmatrix} \mathbb{1}_A & e^{i\phi}\theta_B & 0 \\ 0 & 0 & \mathbb{1}_C \end{pmatrix} \; ; \quad \hat{\gamma} = \mathrm{SC}_{A'} \left[ \hat{\Pi}\, \Gamma\, \hat{\Pi}^\dagger \right] \; .$$

Directly expressed in terms of the map $\Gamma$, this reads

$$\hat{\gamma}(\phi) = \Gamma_C - \Gamma_{C|AB}\, \hat{\Lambda} \frac{1}{\hat{\Lambda}\, \Gamma_{AB|AB}\, \hat{\Lambda}^\dagger}\, \hat{\Lambda}^\dagger\, \Gamma_{AB|C} \tag{7.3}$$

where $\hat{\Lambda} = (\mathbb{1}_A\,;\, e^{i\phi}\theta_B)$ is the upper left subblock of $\hat{\Pi}$.

## 7.6 States with rational trigonometric functions as Fourier transforms

If one restricts to pure MPS (i.e., those for which $\Gamma$ is pure) with one mode per site, then it follows from Theorem 4.2 that these states have reflection symmetry, and therefore $\hat{\gamma}(\phi) = \gamma_0 + 2\sum_{n\geq 0} \gamma_n \cos(n\phi)$ is real. This implies that the sines in (7.3) can only appear in even powers $\sin^{2n}\phi = (1-\cos^2\phi)^n$. Therefore, the Fourier transform $\hat{\gamma}$ of any pure Gaussian MPS, which is a $2\times 2$ matrix valued function of $\phi$, has elements which are rational functons of $\cos(\phi)$, $(\hat{\gamma}(\phi))_{xy} = p_{xy}(\cos(\phi))/q_{xy}(\cos(\phi))$ with $p$, $q$ polynomials. The degree of the polynomials is limited by the size of $\hat{\Lambda}\Gamma_{AB}\hat{\Lambda}^\dagger$, and thus by the number $M$ of the bonds. One can easily check that $\dim p \leq 2M+1$ and $\dim q \leq 2M$.

For the following discussion, let us write those rational functions with a common denominator $d$,

$$\hat{\gamma}(\phi) = \frac{1}{d(\cos(\phi))} \begin{pmatrix} q(\cos(\phi)) & r(\cos(\phi)) \\ r(\cos(\phi)) & p(\cos(\phi)) \end{pmatrix} , \tag{7.4}$$

where $q$, $p$, $r$, and $d$ are polynomials of degree $L$. Then, the set of all such $\hat{\gamma}$ with $L \geq 2M+1$ encompasses the set of translational invariant GMPS with $M$ bonds. Computing correlation functions in a lattice of size $N$ can be done straightforwardly in this representation by taking the discrete Fourier transform of $\hat{\gamma}(\phi)$ which scales polynomially with $N$, and in the following section we show that for one dimension, the correlations can be even computed exactly in the limit of an infinite chain.

It is interesting to note that $\gamma(\phi)$ is already determined up to a finite number of possibilities by fixing $r$ and $d$. Since $\gamma$ is pure, $1 = \det\gamma = \det\hat{\gamma}$, and therefore, $pq = d^2 + r^2$. Therefore, the zeros of $pq$ are the zeros of $d^2 + r^2$, such that the only freedom is to choose how to distribute the zeros on $p$ and $q$. On the contrary, fixing only $q$ and $d$ does not give sufficient information, while choosing $p$, $q$ and $d$ (i.e., the diagonal of $\hat{\gamma}$) does not ensure that there exists a polynomial $r$ such that $pq - r^2 = d^2$.

From the above, it follows that $2L + 1$ parameters are sufficient to describe $\hat{\gamma}(\phi)$, where $L$ is still the degree of the polynomials. This encloses all translational invariant Gaussian MPS with bond number $M \leq (L-1)/2$, which need $(2M + 1)(2M + 2) = L(L + 1)$ parameters. Therefore, the class of states where $\hat{\gamma}(\phi)$ is a rational function of $\cos(\phi)$ is a more efficient description of translationally invariant states than Gaussian MPS are.

Let us stress once more that the results of this section hold for arbitrary spatial dimension.

## 7.7 Correlation length

In the following, we show that the correlations of one-dimensional GMPS decay exponentially and explicitly derive the correlation length. The derivation only makes use of the representation (7.4) of Gaussian MPS and thus holds for the whole class of states where the Fourier transform is a rational function of the cosine. We will restrict to the case where the state $\Gamma$ associated to the GMPS map has only finite entries, which corresponds to the case where the denominator $d(\cos(\phi))$ in (7.4) has no zero on the unit circle.[1]

The correlations are directly obtained by back-transforming the elements of $\hat{\gamma}(\phi)$, which are rational functions $[\hat{\gamma}(\phi)]_s = s(\cos(\phi))/d(\cos(\phi))$, $s \in \{p, q, r\}$; in the limit of an infinite chain,

$$(\gamma_s)_n = \frac{1}{2\pi}\int_0^{2\pi} \frac{s(\cos(\phi))}{d(\cos(\phi))} e^{in\phi} \mathrm{d}\phi \ .$$

Now transform $s$, $d$ to complex polynomials via $\cos(\phi) \to (z+1/z)/2$, and expand with $z^K$, $\tilde{s}(z) := z^K s(z)$, $\tilde{d}(z) := z^K d(z)$, where $K$ is chosen large enough to make

[1] The case where $d$ has zeros on the unit circle corresponds to critical systems, which is why the correlations diverge. In the case of a Hamiltonian $H = V \oplus \mathbb{1}$, however, the ground state correlations of $P$ do not diverge. As in that case one has $pq = d^2$, $p/d = d/q$ need not have a singularity just because $q/d$ has one.

$\tilde{s}$, $\tilde{d}$ polynomials in $z$. Then,

$$\begin{aligned}(\gamma_s)_n &= \frac{1}{2\pi i}\int_{\mathcal{S}^1}\frac{\tilde{s}(z)z^{n-1}}{\tilde{d}(z)}\mathrm{d}z \\ &= \sum_{z_i:\tilde{d}(z_i)=0}\frac{1}{(\nu_i-1)!}\underbrace{\frac{\mathrm{d}^{\nu_i-1}}{\mathrm{d}z^{\nu_i-1}}\left[\frac{\tilde{s}(z)z^{n-1}}{\tilde{d}_i(z)}\right]\Big|_{z=z_i}}_{D_i}\end{aligned}$$

by the calculus of residues, where $\nu_i$ is the order of the zero $z_i$ in $\tilde{d}$ and $\tilde{d}_i(z)(z-z_i)^{\nu_i}=\tilde{d}(z)$. For $n>\nu_i$, $D_i\propto z_i^{(n-\nu_i)}$, and it follows that the correlations decay exponentially, where the correlation length is given by the largest zero of $q(z)$ inside the unit circle.

This proof holds only for one-dimensional GMPS. However, it can be proven for arbitrary spatial dimensions that the correlations decay as $o(\|n\|^{-\infty})$ by iterated integration by parts as in Lemma 4.5.

## 7.8 GMPS as ground states of local Hamiltonians

Finally, we prove that every Gaussian MPS is the ground state of a local Hamiltonian, and show that only a special class of local Hamiltonians has a GMPS as an exact ground state. Again, the proof only relies on the representation (7.4).

Given a state $\gamma$ with Fourier transform (7.4), define $H$ to be the Hamiltonian matrix with Fourier transform

$$\hat{H}(\phi)=\begin{pmatrix} p(\cos(\phi)) & -r(\cos(\phi)) \\ -r(\cos(\phi)) & q(\cos(\phi))\end{pmatrix}. \tag{7.5}$$

Then, $H$ corresponds to a local Hamiltonian—the interaction range is the degree of $p,q,r$—and according to (4.9), $\mathcal{E}(\phi)=\left[\sqrt{pq-r^2}\ \right](\cos\phi)=d(\cos\phi)$, which together with Eq. (4.10) proves that $\gamma$ is the ground state of $H$.

It is interesting to have a brief look at the converse as well. Given a local Hamiltonian, when will it have a GMPS as its ground state? Any local Hamiltonian has a Fourier transform which consists of polynomials in $\cos(\phi)$, and thus we adapt the notation of Eq. (7.5). Then, the ground state is represented by a rational function of $\cos(\phi)$ in Fourier space exactly if $pq-r^2=d^2$ is the square of another polynomial, as can be seen from Eq. (4.10). In terms of the original Hamiltonian, this implies that $H_QH_P-H_{QP}^2$ has to be the square of another banded matrix. For example, for the usual case $H=V\oplus\mathbb{1}$ one would need $V=X^2$ with $X$ again a banded matrix. The Klein-Gordon Hamiltonian (4.1), e.g., does not have a GMPS as its ground state.

# Chapter 8

# Conclusions and outlook

Entanglement is one of the most intriguing phenomena in quantum physics, and it is at the core of quantum information theory. In this thesis, we investigated several topics relating to both the theory and the application of entanglement.

We started in Chapter 2 by investigating how the notion of entanglement changes under the constraint given by superselection rules (SSR). We introduced a new quantity, the superselection induced variance SiV, and showed that together with the entropy of entanglement, it allows to fully quantify the amount of nonlocality of pure states. We extended this to mixed states by generalizing the concept of entanglement of formation and distillable entanglement to SiV and developing protocols how to distill both SiV and entanglement. We demonstrated that SiV is a resource as it allows to overcome the SSR restrictions, and that the new restrictions also give rise to new protocols such as perfect quantum data hiding. Further work might point in various directions: first, finding more cryptographic tasks which become possible in the presence of SSR, second, investigating the behaviour of other mixed state entanglement measures when applied to SiV, and third, searching for better protocols to transform and distill mixed states. Finally, it is an interesting direction to investigate how the concept of entanglement is modified by restrictions beyond particle number conservation, such as permutational symmetry, antisymmetry in fermionic systems, or the lack of a joint reference frame [133, 134].

Next, we investigated the optimal generation of quantum resources in Chapter 3. We showed how to optimally generate squeezing from a given noisy device, and found that the optimal squeezing per iteration only depends on the squeezing of the input. In turn, this led to the result that in order to optimally generate squeezing in an interative scenario, it is sufficient to optimize the squeezing created in each step. We then showed that this has direct implications for the optimal generation of entanglement, given noiseless passive operations. The main tool for the proof was a decomposition of the input into a sub-quantum "coherent kernel" and an additional noise term, which should also prove useful in other contexts where states can be optimized over some group of operations. Indeed, we showed

that the same idea can be used to derive a very similar result for the optimal generation of entanglement in a scenario of two separated parties which have a noisy joint symmetric operation available. Directions of future research include investigating the conjecture on optimal inputs for symmetric channels on which the latter result is based, extending the result to optimal entanglement generation using non-symmetric channels, and considering entanglement generation in the multipartite scenario.

We then turned towards quantum many-body systems. In Chapter 4, we considered lattices of harmonic oscillators using tools developed for Gaussian states in the context of quantum information theory. We investigated both the case of non-critical Hamiltonians, where we found that in general the correlations decay as the interaction, and the case of critical Hamiltonians, where we showed that a vanishing energy gap indeed implies an algebraic decay of correlations. In particular, we found that the correlations in one-dimensional systems decay as $1/n^2$, while in higher dimension, they decay as $\log n/n^{d+1}$. A special case is given for an interaction decaying as $1/n^3$, which corresponds to the relevant case of ions in a trap: while for a positive sign of the interaction, the $1/n^2$ law still holds, a $\sqrt{\log n}$ correction is encountered for the negative sign. Future research directions include another look at the higher-dimensional critical case, where it remains to be seen whether the log correction is tight, as well as the case of non-critical Hamiltonians with exponentially decaying interaction which might imply an exponential decay of correlations. Finally, one can try to infer more than the asymptotic decay of correlations: especially for practical cases as ions in a trap, the short-range scaling of correlations is of particular interest.

The remaining part of the thesis was devoted to the description of quantum many-body systems in terms of Matrix Product States (MPS) and Projected Entangled Pair States (PEPS). We started in Chapter 5 by investigating how approximability by MPS is related to the scaling of block entropies. We showed that a logarithmic scaling of $\alpha < 1$ Rényi entropies is sufficient for approximability, while conversely a linear scaling of the von Neumann entropy and an faster than logarithmic scaling of $\alpha > 1$ Rényi entropies both imply non-approximability. We also demonstrated that all other entropy scalings do not allow for conclusions about approximability, which in particular holds for the case of bounded von Neumann entropy. We then applied the tools developed to show that time evolution even of translational invariant and time independent systems cannot be simulated efficiently using MPS, demonstrating that a quantum computer will likely outperform classical computers in this task. Future research will investigate whether an area law does imply approximability for translational invariant systems [135], and aim at finding better criteria for approximability, e.g., by looking at the joint scaling of entropies, at smooth entropies (cf. the example at the end of Chapter 5), or at other more refined spectral properties. More generally, it is of big interest to get a good understanding of the mechanisms underlying approximability by MPS and related states in order to apply these techniques in

the best possible way.

In Chapter 6, we turned towards PEPS, the higher-dimensional generalization of MPS, and investigated the computational complexity of creating and simulating them. We found that these tasks are PP- and #P-complete, respectively, where the central tool for our proofs was a duality between PEPS and postselection. Future research will consider the relation between the complexity of ground state problems and the spectral properties of the Hamiltonian, in particular the energy gap, as hinted at the end of Chapter 6, as well as the relation of PEPS and ground states (cf. [104]). Moreover, it remains to be investigated whether there exist interesting subclasses of PEPS which are computationally easier than the general case, as well as to find an example of a PEPS which cannot be created efficiently even with free classical side-processing. Finally, the PEPS–postselection duality can likely be applied in many more contexts: for instance, the proof that quantum messages can be simulated by classical messages of essentially the same length [121] relies on postselected circuits on the receiver's side, which implies that every quantum message can be replaced by a PEPS without the need for extra computational power. The implications of this observation, and in general the scenarios in which PEPS can (or cannot) replace general quantum states, remain to be investigated.

Finally, in Chapter 7 we introduced MPS for Gaussian states on a lattice, so-called Gaussian Matrix Product States (GMPS). We showed that GMPS can approximate every state even under translational invariance, derived a decomposition bounding the entanglement contained in the bonds, clarified their relation to local Hamiltonians, and demonstrated how to compute the scaling of correlations and the correlation length. As a tool, we introduced a description of translational invariant Gaussian states using covariance matrices with rational functions as their Fourier transforms, which encompasses the set of GMPS and shares a lot of their properties. As to future directions, GMPS can serve as a toy model for spin systems due to the typically less involved structure of Gaussian states, and one might investigate several questions which are of interest for PEPS, e.g. whether GMPS can be disentangled, how well they can approximate ground states, or how the correlations scale in higher dimensional systems. Other directions of research are to extend GMPS to more general or different scenarios, as non-Gaussian bosonic states or fermionic Gaussian states, and to combine infinite dimensional Gaussian bonds with finite dimensional physical systems in order to get a family of states which is more general than MPS [136].

# Acknowledgements

First and most of all, I want to thank Ignacio Cirac—for the opportunity to do my PhD in his group, for his confidence and his advice, and for generally being the best possible supervisor. His inventiveness, his creativity, and his enthusiasm in doing physics have taught me a lot and made a lasting impression on me.

This work has been carried out in collaboration with Frank Verstraete and Michael Wolf. It was a great experience to work with them, and I learned a lot from both, although in somewhat different ways: Frank's undamped enthusiasm has taught me to discuss without reservation and to defend my ideas; Michael always had an open ear, and his didactical abilities in answering all kind of questions are really remarkable. I am also grateful to both for advice regarding non-quantum problems in a physicist's life.

I want to thank Karl Gerd Vollbrecht for the numerous insightful discussions we had, and in particular for conveying the beauty of simplicity to me. Also, I enjoyed a lot collaborating with David Pérez García.

This work profited from discussions with a lot of people—thanks to Koenraad Audenaert, Toby Cubitt, Chris Dawson, Jens Eisert, Géza Giedke, Otfried Gühne, Klemens Hammerer, Diego Porras, and Tommaso Roscilde.

The time in this group would have been much less enjoyable without all the people around who made both work and leisure more pleasant: thanks to all of you for the good time. In particular, I want to thank Valentin Murg for being a calm and pleasant office mate. Thanks also go to our secretaries Renate, Marion, and Verena, who made life in the presence of administrative issues a lot easier.

Finally, I want to thank my family for all the support I have received throughout.

This work was supported in parts by the EU (projects RESQ, QUPRODIS, COVAQIAL), the DFG-Forschergruppe 635, the "Kompetenznetzwerk Quanteninformationsverarbeitung der Bayerischen Staatsregierung", and the Elite Network of Bavaria project QCCC.

# Bibliography


[1] C. Bennett and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, in *Proceedings of IEEE International Conference on Computers Systems and Signal Processing*, pp. 175–179, Bangalore, India, 1984.

[2] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[3] C. H. Bennett *et al.*, Phys. Rev. Lett. **70**, 1895 (1993).

[4] C. H. Bennett and S. J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).

[5] A. Ekert and R. Jozsa, Rev. Mod. Phys. **68**, 733 (1996).

[6] Y. Shi, J. Phys. A: Math. Gen. **37**, 6807 (2004).

[7] A. Osterloh, L. Amico, G. Falci, and R. Fazio, Nature **416**, 608 (2002).

[8] E. R. G. Vidal, J.I. Latorre and A. Kitaev, Phys. Rev. Lett. **90**, 227902 (2003).

[9] A. Kitaev and J. Preskill, Phys. Rev. Lett. **96**, 110404 (2006).

[10] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996), quant-ph/9511030.

[11] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996), quant-ph/9604024.

[12] C. H. Bennett *et al.*, Phys. Rev. Lett. **76**, 722 (1996), quant-ph/9511027.

[13] L.-M. Duan, M. Lukin, I. Cirac, and P. Zoller, Nature **414**, 413 (2001), quant-ph/0105105.

[14] A. Acín, J. I. Cirac, and M. Lewenstein, Nature Physics **3**, 256 (2007), quant-ph/0612167.

[15] G. Vidal, K. Hammerer, and J. I. Cirac, Phys. Rev. Lett. **88**, 237902 (2002), quant-ph/0112168.

[16] B. Kraus, K. Hammerer, G. Giedke, and J. I. Cirac, Phys. Rev. A **67**, 042314 (2003), arXiv:quant-ph/0210136.

[17] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[18] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005), cond-mat/0409292.

[19] A. Klümper, A. Schadschneider, and J. Zittartz, Z. Phys. B **87**, 281 (1992).

[20] F. Verstraete, D. Porras, and J. I. Cirac, Phys. Rev. Lett. **93**, 227205 (2004), cond-mat/0404706.

[21] G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003), quant-ph/0301063.

[22] G. Vidal, Phys. Rev. Lett. **93**, 040502 (2004), quant-ph/0310089.

[23] F. Verstraete, J. J. Garcia-Ripoll, and J. I. Cirac, Phys. Rev. Lett. **93**, 207204 (2004), cond-mat/0406426.

[24] F. Verstraete and J. I. Cirac, (2004), cond-mat/0407066.

[25] M. Wolf, J. Eisert, and M. Plenio, Phys. Rev. Lett. **90**, 047904 (2003).

[26] F. Verstraete and J. I. Cirac, Phys. Rev. B **73**, 094423 (2006), cond-mat/0505140.

[27] C. Schön, E. Solano, F. Verstraete, J. I. Cirac, and M. M. Wolf, Phys. Rev. Lett. **95**, 110503 (2005), quant-ph/0501096.

[28] F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, Phys. Rev. Lett. **96**, 220601 (2006), quant-ph/0601075.

[29] N. Schuch, F. Verstraete, and J. I. Cirac, Phys. Rev. A **70**, 042310 (2004), quant-ph/0404079.

[30] N. Schuch, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **92**, 087904 (2004), quant-ph/0310124.

[31] F. Verstraete, N. Schuch, and J. I. Cirac, Proc. SPIE **5468**, 93 (2004).

[32] N. Schuch, M. M. Wolf, and J. I. Cirac, Proc. SPIE **5842**, 37 (2005).

[33] N. Schuch, M. M. Wolf, and J. I. Cirac, Phys. Rev. Lett. **96**, 023004 (2006), quant-ph/0505145.

[34] N. Schuch, J. I. Cirac, and M. M. Wolf, Commun. Math. Phys. **267**, 65 (2006), quant-ph/0509166.

[35] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **98**, 140506 (2007), quant-ph/0611050.

[36] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, (2007), arXiv:0705.0292.

[37] P. Zanardi, Phys. Rev. Lett. **87**, 077901 (2001), quant-ph/0103030; P. Zanardi, D. Lidar, and S. Lloyd, Phys. Rev. Lett. **92**, 060402 (2004), quant-ph/0308043.

[38] H. Barnum, E. Knill, G. Ortiz, and L. Viola, Phys. Rev. A **68**, 032308 (2003), quant-ph/0207149; H. Barnum, E. Knill, G. Ortiz, R. Somma, and L. Viola, Phys. Rev. Lett. **92**, 107902 (2004), quant-ph/0305023.

[39] S. Popescu, personal communication.

[40] F. Verstraete and J. I. Cirac, Phys. Rev. Lett. **91**, 010404 (2003), quant-ph/0302039.

[41] H. M. Wiseman and J. A. Vaccaro, Phys. Rev. Lett. **91**, 097902 (2003), quant-ph/0210002; S. D. Bartlett and H. M. Wiseman, Phys. Rev. Lett. **91**, 097903 (2003), quant-ph/0303140; H. M. Wiseman, S. D. Bartlett, and J. A. Vaccaro, quant-ph/0309046; J. A. Vaccaro, F. Anselmi, and H. M. Wiseman, Int. J. Quant. Inf., **1**, 427 (2003), quant-ph/0311028.

[42] B. M. Terhal, D. P. DiVincenzo, and D. W. Leung, Phys. Rev. Lett. **86**, 5807 (2001).

[43] A. Kitaev, D. Mayers, and J. Preskill, (2003), quant-ph/0310088.

[44] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, (2004), quant-ph/0403161.

[45] M. Nielsen, Phys. Rev. Lett. **83**, 436 (1999), quant-ph/9811053.

[46] D. Jonathan and M. B. Plenio, Phys. Rev. Lett. **83**, 1455 (1999), quant-ph/9903054.

[47] H.-K. Lo and S. Popescu, Phys. Rev. A **63**, 022301 (2001), quant-ph/9707038.

[48] G. Vidal, J. Mod. Opt. **47**, 355 (2000), quant-ph/9807077.

[49] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley-Interscience, 1991).

[50] M. A. Nielsen and I. A. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

[51] P. M. Hayden, M. Horodecki, and B. M. Terhal, J. Phys. A **34**, 6891 (2001), quant-ph/0008134.

[52] A. Pomeransky, Phys. Rev. A **68**, 032317 (2003), quant-ph/0305056.

[53] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998), quant-ph/9709029.

[54] F. Verstraete, J. Dehaene, and B. DeMoor, Phys. Rev. A **64**, 010101(R) (2001), quant-ph/0011111.

[55] A. Chefles, Contemporary Physics **41**, 401 (2000), quant-ph/0010114.

[56] K. Banaszek, Phys. Rev. A **62**, 024301 (2000), quant-ph/0002088.

[57] T. Rudolph and B. C. Sanders, Phys. Rev. Lett. **87**, 077903 (2001), quant-ph/0103147.

[58] C. Caves, Phys. Rev. D **23**, 1693 (1981).

[59] E. Polzik, J. Carri, and H. Kimble, Phys. Rev. Lett. **68**, 3020 (1992).

[60] N. P. Georgiades, E. S. Polzik, K. Edamatsu, H. J. Kimble, and A. S. Parkins, Phys. Rev. Lett. **75**, 3426 (1995).

[61] M. Xiao, L.-A. Wu, and H. Kimble, Phys. Rev. Lett. **59**, 278 (1987).

[62] P. Grangier, R. E. Slusher, B. Yurke, and A. LaPorta, Phys. Rev. Lett. **59**, 2153 (1987).

[63] Z. Y. Ou, S. F. Pereira, H. J. Kimble, and K. C. Peng, Phys. Rev. Lett. **68**, 3663 (1992).

[64] C. Silberhorn *et al.*, Phys. Rev. Lett. **86**, 4267 (2001).

[65] W. P. Bowen, R. Schnabel, P. K. Lam, and T. C. Ralph, Phys. Rev. Lett. **90**, 043601 (2003).

[66] A. S. Holevo, *Probabilistic and statistical aspects of quantum theory* (North-Holland Publishing Company, 1982).

[67] B. Demoen, P. Vanheuverzwijn, and A. Verbeure, Lett. Math. Phys. **2**, 161 (1977).

[68] J. Fiurášek, Phys. Rev. Lett. **89**, 137904 (2002), quant-ph/0204069.

[69] G. Giedke and J. I. Cirac, Phys. Rev. A **66**, 032316 (2002), quant-ph/0204085.

[70] J. Eisert and M. Plenio, Int. J. Quant. Inf. **1**, 479 (2003).

[71] R. Simon, N. Mukunda, and B. Dutta, Phys. Rev. A **49**, 1567 (1994).

[72] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Phys. Rev. Lett. **73**, 58 (1994).

[73] J. Geremia, J. Stockton, and H. Mabuchi, Science **304**, 270 (2004).

[74] G. Vidal and R. Werner, Phys. Rev. A **65**, 032314 (2002).

[75] K.␣G. Vollbrecht, private communication.

[76] G. Giedke, M. M. Wolf, O. Krueger, R. F. Werner, and J. I. Cirac, (2003), quant-ph/0304042.

[77] M. M. Wolf, F. Verstraete, and J. Cirac, Phys. Rev. Lett. **92**, 087903 (2004), quant-ph/0307060.

[78] M. B. Plenio, J. Eisert, J. Dreissig, and M. Cramer, Phys. Rev. Lett. **94**, 060503 (2005), quant-ph/0405142.

[79] M. Cramer, J. Eisert, M. B. Plenio, and J. Dreissig, Phys.Rev. A **73**, 012309 (2006), quant-ph/0505092.

[80] M. M. Wolf, Phys. Rev. Lett. **96**, 010404 (2006), quant-ph/0503219.

[81] K. Audenaert, J. Eisert, M. B. Plenio, and R. F. Werner, Phys. Rev. A **66**, 042327 (2002).

[82] A. Botero and B. Reznik, Phys. Rev. A **70**, 052329 (2004), quant-ph/0403233.

[83] M. Asoudeh and V. Karimipour, Phys. Rev. A **72**, 0332339 (2005), quant-ph/0506022.

[84] M. B. Plenio and F. L. Semiao, New J. Phys. **7**, 73 (2005), quant-ph/0407034.

[85] M. B. Plenio, J. Hartley, and J. Eisert, New J. Phys. **6**, 36 (2004), quant-ph/0402004.

[86] J. Eisert, M. B. Plenio, S. Bose, and J. Hartley, Phys. Rev. Lett. **93**, 190402 (2004), quant-ph/0311113.

[87] M. M. Wolf, F. Verstraete, and J. Cirac, Int. J. Quant. Inf. **1**, 465 (2003), quant-ph/0311051.

[88] B. Nachtergaele and R. Sims, Commun. Math. Phys. **265**, 119 (2006), math-ph/0506030.

[89] M. B. Hastings and T. Koma, Commun. Math. Phys. **265**, 781 (2006), math-ph/0507008.

[90] M. Cramer and J. Eisert, New J. Phys. **8**, 71 (2006), quant-ph/0509167.

[91] A. Auerbach, *Interacting electrons and quantum magnetism* (Springer Verlag, New York, 1994).

[92] D. James, Appl. Phys. B **66**, 181 (1998), quant-ph/9702053.

[93] J. Williamson, American Journal of Mathematics **58**, 141 (1936).

[94] G. Birkl, S. Kassner, and H. Walther, Nature **357**, 310 (1992).

[95] D. H. E. Dubin, Phys. Rev. Lett. **71**, 2753 (1993).

[96] D. G. Enzer *et al.*, Phys. Rev. Lett. **85**, 2466 (2000).

[97] T. B. Mitchell *et al.*, Science **282**, 1290 (1998).

[98] J. Manuceau and A. Verbeure, Comm. Math. Phys. **9**, 293 (1968).

[99] A. S. Holevo, Theor. Math. Phys. **6**, 1 (1971).

[100] Arvind, B. Dutta, N. Mukunda, and R. Simon, Pramana **45**, 471 (1995), quant-ph/9509002.

[101] M. M. Wolf, G. Giedke, O. Krüger, R. F. Werner, and J. I. Cirac, Phys. Rev. A **69**, 052320 (2004), quant-ph/0306177.

[102] M. Benzi and G. H. Golub, BIT Numerical Mathematics **39**, 417 (1999).

[103] N. Bleistein and R. A. Handelsman, *Asymptotic expansions of integrals* (Dover Publication, New York, 1986).

[104] M. B. Hastings, (2007), cond-mat/0701055.

[105] K. M. R. Audenaert, (2006), quant-ph/0610146.

[106] S. Lloyd, Science **273**, 1073 (1996).

[107] K. G. H. Vollbrecht and J. I. Cirac, (2005), quant-ph/0502143.

[108] R. Raussendorf, Phys. Rev. A **72**, 052301 (2005), quant-ph/0505122.

[109] K. G. Vollbrecht and J. I. Cirac, (2007), arXiv:0704.3432.

[110] R. Bhatia, *Matrix Analysis* (Springer, New York, 1996).

[111] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, (2006), quant-ph/0608197.

[112] P. Calabrese and J. Cardy, J. Stat. Mech. P04010 (2005), cond-mat/0503393.

[113] N. Schuch *et al.* , in preparation.

[114] R. Renner, (2005), quant-ph/0512258.

[115] S. Aaronson, Proc. R. Soc. Lond. A **461**, 3473 (2005), quant-ph/0412187.

[116] F. Verstraete and J. I. Cirac, Phys. Rev. A **70**, 060302 (2004), quant-ph/0311130.

[117] S. Aaronson, Proc. ACM STOC '04 , 118 (2004), quant-ph/0311039.

[118] R. Raussendorf and H. J. Briegel, Phys. Rev. Lett. **86**, 5188 (2001), quant-ph/0010033.

[119] R. Raussendorf, D. E. Browne, and H. J. Briegel, Phys. Rev. A **68**, 022312 (2003), quant-ph/0301052.

[120] C. M. Papadimitriou, *Computational complexity* (Addison-Wesley, Reading, MA, 1994).

[121] S. Aaronson, Theory of Computing **1**, 1 (2005), quant-ph/0402095.

[122] Y. Shi, Quant. Inf. Comput. **3**, 84 (2003), quant-ph/0205115.

[123] D. Aharonov, (2003), quant-ph/0301040.

[124] C. M. Dawson *et al.*, Quant. Inf. Comput. **5**, 102 (2005), quant-ph/0408129.

[125] D. Aharonov and T. Naveh, (2002), quant-ph/0210077.

[126] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and quantum computation* (American Mathematical Society, Providence, Rhode Island, 2002).

[127] R. Oliveira and B. M. Terhal, (2005), quant-ph/0504050.

[128] D. Perez-Garcia, M. M. Wolf, F. Verstraete, and J. I. Cirac, in prepraration.

[129] M. Vyalyi, Electronic Colloquium on Computational Complexity **10** (2003), `http://eccc.hpi-web.de/eccc-reports/2003/TR03-021/index.html`.

[130] A. Jamiołkowski, Rep. Math. Phys. **3**, 275 (1972).

[131] C. V. Kraus, M. M. Wolf, and J. I. Cirac, (2006), quant-ph/0607094.

[132] A. S. Holevo and R. F. Werner, Phys. Rev. A **63**, 032312 (2001), quant-ph/9912067.

[133] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Rev. Mod. Phys. **79**, 555 (2007), quant-ph/0610030.

[134] M.-C. Bañuls, J. I. Cirac, and M. M. Wolf, (2007), arXiv:0705.1103.

[135] M. B. Hastings, (2007), arXiv:0705.2024.

[136] S. Iblisdir, R. Orus, and J. I. Latorre, Physical Review B **75**, 104305 (2007), cond-mat/0610530.