Freie Universität Berlin
Dahlem Center for Complex Quantum Systems

Locality and complexity in simulations of complex quantum systems

Im Fachbereich Physik der
Freien Universität Berlin
eingereichte Dissertation

Martin Kliesch

Berlin, 2014

Erstgutachter:	Prof. Dr. Jens Eisert, Freie Universität Berlin
Zweitgutachter:	Prof. Dr. Piet Brouwer, Freie Universität Berlin
Tag der Disputation:	22. April 2015

Contents

List of publications

This cumulative dissertation is based on the following first-author publications.

[Kli+11a] on pp. 32ff — M. Kliesch, T. Barthel, C. Gogolin, M. Kastoryano, and J. Eisert, "*Dissipative Quantum Church-Turing Theorem*", Phys. Rev. Lett. **107**, 120501 (2011), DOI: 10.1103/PhysRevLett.107.120501

[KGE14b] on pp. 13ff — M. Kliesch, C. Gogolin, and J. Eisert, "*Lieb-Robinson Bounds and the Simulation of Time-Evolution of Local Observables in Lattice Systems*", in: Many-Electron Approaches in Physics, Chemistry and Mathematics, ed. by V. Bach and L. Delle Site, Mathematical Physics Studies, Springer International Publishing, 2014, pp. 301–318, DOI: 10.1007/978-3-319-06379-9_17, arXiv:1306.0716

[Kli+14] on pp. 47ff — M. Kliesch, C. Gogolin, M. J. Kastoryano, A. Riera, and J. Eisert, "*Locality of Temperature*", Phys. Rev. X **4**, 031019 (2014), DOI: 10.1103/PhysRevX.4.031019

[KGE14a] on pp. 69ff — M. Kliesch, D. Gross, and J. Eisert, "*Matrix-Product Operators and States: NP-Hardness and Undecidability*", Phys. Rev. Lett. **113**, 160503 (2014), DOI: 10.1103/PhysRevLett.113.160503

Other publications by the author that were written during this thesis are the following.

[BKE10] on pp. 101ff — T. Barthel, M. Kliesch, and J. Eisert, "*Real-Space Renormalization Yields Finite Correlations*", Phys. Rev. Lett. **105**, 010502 (2010), DOI: 10.1103/PhysRevLett.105.010502

[BK12] on pp. 108ff — T. Barthel and M. Kliesch, "*Quasilocality and Efficient Simulation of Markovian Quantum Dynamics*", Phys. Rev. Lett. **108**, 230504 (2012), DOI: 10.1103/PhysRevLett.108.230504

[GKAE13] on pp. 117ff — C. Gogolin, M. Kliesch, L. Aolita, and J. Eisert, "*Boson-Sampling in the light of sample complexity*", (2013), arXiv:quant-ph/1306.3995

[AGKE14] on pp. 139ff — L. Aolita, C. Gogolin, M. Kliesch, and J. Eisert, "*Quantum certification of photonic quantum simulations*", submitted to Nat. Comm. (2014), arXiv:quant-ph/1407.4817

The first-author publications are presented in the main text. In the discussion of these publications, references are repeatedly made also to the other publications, which are included in

Appendix B in order to keep the thesis self-contained.

The author's contributions

For each publication the author's contribution to conception, realization, and report writing is listed in the following (see also Appendix C.5 for a German version).

[Kli+11a] The author is the main contributor of this publication. In particular, he developed and wrote important parts of the proofs in main text and the Supplemental Material.

[KGE14b] The author is the main contributor of this publication. He provided important input to the conception of this book chapter. Moreover, substantial parts of the text and the figures were generated by the author.

[Kli+14] The author is the main contributor of this publication. He developed and wrote out important parts of the proofs of the main text and the appendix. Considerable parts of the whole text and the figures were generated by the author.

[KGE14a] The author is the main contributor of this publication. He adapted classical methods for hidden Markov models to the quantum case and wrote out the technical parts with its definitions, theorems, and proofs. Also the figures were generated by the author. Moreover, he wrote important parts of the introduction, of the review part in the main text, and of the appendix.

[BKE10] The author contributed a large portion to the development of the notation, the development of the technical framework, the proofs, and the *refinements* and examples in the Supplemental Material, and also took part in the writing of the manuscript. A central part of the publication are the figures that were in large parts generated by the author.

[BK12] The author contributed important parts to the conception of the work and to the construction of the proofs. He also took part in the writing of the main text and the Supplemental Material.

[GKAE13] Important contributions to the composition of the problem statement, which is an essential part of this publication, as well as the conception of the work were contributed by the author. Moreover, the author contributed important parts to the construction and writing of the proofs and helped to write the whole manuscript.

[AGKE14] The author made important contributions to the conception of this work, to the formulation of the theorems and development of their proofs, generating the figures, and to the writing of the manuscript.

1 Introduction

Understanding quantum many-body systems is an enduring challenge in modern physics. In the past decades, methods from computer science have played an increasingly important role in pursuing this challenge. The development of supercomputers and elaborate numerical techniques allows for more and more powerful simulations of quantum systems. At the same time, when it comes to harness certain quantum systems for computational tasks, their power as well as limitations can be grasped in the language of complexity theory.

The main task in computational complexity theory is to classify computational problems according to their difficulty and such classifications are made for different models of computation (see Appendix A.1 for details). For instance, the complexity classes P and NP contain roughly those *decision problems* (problems with yes-or-no answers) that can be solved or, respectively, whose solutions can be verified, by classical computers efficiently. Whether or not these two classes coincide is still an open problem; actually one of the seven Millennium Prize Problems with a prize of 10^6 US dollars. It is, however, strongly believed that there are problems in NP that are not in P, meaning that NP contains problems that are intractable on classical computers. It is the hope that one can use quantum resources for efficiently solving some of these and also other classically intractable problems.

More concretely, one would like to use quantum many-body systems to solve problems that are practically impossible to solve on classical computers. This hope is inspired by the study of *quantum algorithms* [NC00], which can be performed by sequential unitary operations, each only acting on two two-dimensional quantum systems called *qubits*. It has turned out that some problems believed to be intractable on classical computers, can be solved by such quantum algorithms efficiently. One of the most famous examples of this type is integer factoring [Sho97]. The ultimate goal certainly is, to actually build a *universal quantum computer*, i.e., a machine on which one can efficiently run any quantum algorithm. However, the much more modest goal of implementing a *quantum simulation*, i.e., of solving only some specific but still classically intractable problem with the help of quantum systems, would already constitute a great breakthrough [CZ12, Pre12].

Having potentially more computational power at hand than that of classical computers, one would like to assess what can actually be done with it. More explicitly, one would also like to classify problems in quantum systems according to various notions of complexity. This has de-

veloped into the research field of *Hamiltonian complexity theory*. One guiding question in this filed is "How hard is it to simulate a physical system?" [Osb12]. The situation is particularly well-investigated for one-dimensional lattice systems. Already with time evolutions generated by relatively simple local Hamiltonians [Nag10] or by using ground states of spin chains [AGIK09] one can perform universal quantum computations, meaning that such systems are expected to be classically intractable. As it turns out, ground states can be be very complex objects, as, e.g., only approximating the ground state energy of translation invariant spin chains [GI09] is QMA-hard, i.e., cannot be done efficiently even on a quantum computer. Once a system has an energy gap between the ground states and first excited states, the situation changes drastically: Any unique ground state features exponential clustering of correlations [HK06, NS06] and can be approximated classically efficiently [LVV13]. Most classical approximation algorithms are build on so-called *tensor networks* (cp. review part in Publication [KGE14a] on pages 69ff, and Publication [BKE10] in Appendix B for a discussion of famous examples), which essentially are computationally friendly parameterizations of the relevant degrees of freedom. This is worth explaining in more detail: The key for the success of tensor network methods is, to exploit the *locality structure* of those systems, i.e, to use that interactions involve only a constant number (usually two) of constituents and/or that interaction strengths decay rapidly with the distance. In this case, the speed of propagation of correlations is bounded by the so-called *Lieb-Robinson speed*. Rigorous formulations of such a statement are called *Lieb-Robinson bounds* [LR72] (see Publication [KGE14b] on pages 13ff for an introduction). While having a number of immediate physical consequences, Lieb-Robinson bounds also provide very powerful proof tools. As a consequence [Osb06], time evolution of local observables can be simulated efficiently in the system size [DKSV04, Hae+11, Vid04]. An important stepping stone for such simulations are *Trotter-Suzuki* approximations [Suz76, Tro59] that allow for a decomposition of the full evolution into small and tractable building blocks by approximately factoring certain matrix exponentials into local operators. Besides the consequences for the dynamics, also the above mentioned results on ground states follow: The Lieb-Robinson speed together with the Hamiltonian gap induces a bounded correlation length [NS06], which, in turn, allows for an efficient approximation by tensor network states for one-dimensional systems [BH13].

On a rigorous level, a lot less is known in cases where quantum states are mixed. Two important frameworks are the ones of Makovian time evolution of open quantum systems and Hamiltonian systems in thermal equilibrium, i.e., systems at non-zero temperature. Markovian and purely dissipative time evolution can be used to perform universal quantum computation [VWIC09], which is called *dissipative quantum computation*. However, for many standard tools used in the investigation of closed quantum systems, it is unclear if and how they carry over to the dissipative setting. For thermal states, the focus has been on so-called cluster properties in, and the uniqueness of, the thermodynamic limit, for classical systems [Rue64, Rue69], quantum gases [Gin65], i.e., translation invariant Hamiltonians in the continuum, and cubic lattices [BR97, Gre69]. For such systems the existence and uniqueness of thermal states in the

thermodynamic limit at high temperatures is proven and analyticity of correlations can be derived. Moreover, in the regime of high temperatures, n-point correlation functions are shown to cluster for spin gases [Gin65, Rue64] and bosonic lattices [PY95] in the translation invariant case. However, the questions of the existence of efficient classical simulation algorithms and stability results for thermal states are still open. This is also highly disturbing from a physical point of view: The desired mathematical properties of thermal states are intertwined with the definition of temperature as a *local* quantity [FGSA12, HM05, HMH04]. More precisely, in the case of interacting systems it is not clear how temperature can be defined locally, which we call the *locality of temperature problem*. Advances toward a better understanding of thermal states are hindered by a lack of methods to exploit the system's locality structure. This lack of methods also seems to be an obstacle to progress on some of the most interesting open questions related to foundations of statistical mechanics and the equilibration and thermalization behaviour of closed quantum systems [GLTZ06, LPSW09, RDO08, RGE12, RK12], such as equilibration time scales [CDEO08, CE10].

Going from classical to quantum simulations is expected to lead to new and more powerful ways to understand complex quantum systems. Indeed, it is an intriguing idea to simulate one (quantum) system using another quantum system. For the case, where the former is intractable on classical computers, this would lead to the celebrated "quantum supremacy" [Pre12] meaning that (for specific tasks) quantum systems are more powerful than classical ones. Indeed, this idea has opened the race to implement the first quantum simulation outperforming classical computers. To some extent, this goal has already been achieved with ultracold atoms in optical lattices [Bra+14, Tro+12], with Boson-Sampling in linear optical networks (see discussion and references in Publication [GKAE13]), and with superconducting flux qubits [Boi+13, Boi+14]. However, as quantum simulations are intractable classically, it is unclear how one can test whether or not a claimed quantum simulation does indeed work correctly. In the light of this problem, it is even unclear what a quantum simulation precisely is in the first place.

In this thesis, rigorous mathematical tools are used and developed further to generalize several pure state Hamiltonian complexity results to settings involving mixed quantum states. Also the structure of spaces of physically relevant quantum many-body states is investigated and the significance for simulations is pointed out. Finally, complementary results on verification of quantum simulations are discussed. Most of the results have in common that they rely on the spatial locality structure of the quantum many-body systems under investigation. Exploiting this locality structure leads to numerous implications concerning simulatability and intractability of certain properties of these systems as well as physical consequences.

First, tools from basic Hamiltonian complexity theory are generalized to the setting of Markovian open quantum spin lattice systems. In Publication [Kli+11a], we derive a dissipative Trotter-Suzuki-type approximation, which has a number of immediate consequences. Most importantly, Markovian dynamics can be simulated efficiently on a quantum computer,

or, more precisely, by a unitary quantum circuit of a size scaling polynomially in the simulation time and the size of the system. This means that dissipative quantum computing [VWIC09] is not more powerful than the unitary circuit model, which is the standard model of a quantum computer. As we will see, our results can be seen as a so-called Church-Turing type statement, as it guarantees under very natural assumptions, such as weak coupling to an environment, that the dynamics of open quantum systems can be simulated efficiently on a quantum computer. It then follows that Markovian dynamics is quasi-local and can be locally simulated on classical computers, where the simulation is efficient in the system size and the error but, of course, not efficient in time (Publication [BK12] in Appendix B). This results relies on a generalization of a dissipative Lieb-Robinson bound. An introduction to these works and to consequences of Lieb-Robinson bounds in general, is given in Publication [KGE14b], which appeared as a book chapter. Together, the dissipative Trotter-Suzuki approximation and the dissipative quasi-locality provide rigorous methods for classical simulations of open quantum systems. However, there is also a fundamental obstacle for classical reliable simulations of properties of stationary states of dissipative systems, which is identified in Publication [KGE14a]: Deciding whether or not a translation invariant operator in matrix product form is positive semi-definite, is NP-hard in the system size and undecidable if no bound on the system size is given. This means that this problem is as complex as the most complex problems in NP and cannot be decided by any algorithm in the thermodynamic limit. These findings imply that classical computers cannot check efficiently whether or not an operator parameterized in the computation friendly matrix product form is a density matrix. We also comment on possibilities to overcome this obstacle. From the earlier announced result that dissipative dynamics can be effectively simulated in the circuit model, we obtain an intriguing consequence for the structure of state space: Most states cannot be prepared efficiently by local dynamics [Kli+11a].

For thermal states on spin and fermionic lattices, I provide a perturbation formula and a strong version of exponential clustering of correlations at high temperatures in Publication [Kli+14]. Together, this constitutes a thermal Lieb-Robinson type bound having again a number of implications: It solves the locality of temperature problem, allows for efficient local classical simulation at high temperatures, and also provides a bound on critical temperatures of thermal phase transitions involving long-range order.

As a further discussion, of simulations and complex quantum systems, we will also discuss so-called *quantum simulations*. In particular, we will discuss the precise meaning of "outperforming classical computers". It turns out that the situation is more delicate as compared to the one where one is confronted with problems in NP, such as factoring of integers. This is due to the *verification problem*, i.e., that there is no obvious way to verify that a classically intractable quantum simulation actually does what it is intended to do. For Boson-Sampling simulations, min-entropy and sample complexity lower bounds are derived in Publication [GKAE13] in Appendix B. These results provide rigorous evidence that Boson-Sampling cannot be verified classically efficiently. Complementary to that, an explicit verification protocol was introduced

very recently (see Publication [AGKE14] in Appendix B) that uses very simple quantum resources. With this protocol, verification of quantum simulations of certain classes of multimode bosonic states, including many states occurring in quantum linear optics such as current Boson-Sampling experiments is already possible. The protocol is efficient in the number of modes, in the error one needs to allow for, and in the success probability of measurement post selection that can be allowed for.

This thesis is structured into two main chapters (Chapters 2 and 3). Chapter 2 contains the quantum lattice systems results on Markovian dynamics, the structure of states space, and on thermal states. Chapter 3 provides a further discussion of the verification problem, which is often neglected in the literature on quantum simulations. In particular, two complementary verification results on quantum simulations are discussed. All first-author publications are presented in the subsequent chapter. In order to keep the discussion self-contained, the author's other publications are included in Appendix B. The presentation of all publications is complemented by short comprehensive introductions, where the results are interpreted and evaluated, further connections between the results are discussed, and where the results are put into context with other known results.

In order to keep these introductions as self-contained as possible, some of the standard terminology used in the field, such as the precise meaning of *efficiency*, is is briefly explained in Appendix A.

2 Locality and complexity in lattice systems

Spin lattice systems are an important class of models in solid state physics [AM76], as they are used to study critical points and phase transitions [Voj03] of magnetic systems. They are also ideal candidates for testing locality related questions in Hamiltonian complexity theory, as the discrete locality structure allows for a (computational) complexity theoretic investigation.

This chapter starts with an introduction to Lieb-Robinson bounds and a review of some of their consequences in Section 2.1. Then, in Section 2.2 of this chapter, we generalize various results concerning time evolution from closed to open quantum systems. For thermal states, we derive a Lieb-Robinson type bound and discuss several physical implications in Section 2.3. We also provide surprising results concerning the structure of state space in Section 2.4. In particular, in Section 2.4.1, we show that in more than one spatial dimension, so-called real-space renormalization schemes yield states with bounded correlations and, in Section 2.4.2, we identify a roadblock for reliable simulation techniques for mixed states.

2.1 Introduction to Lieb-Robinson bounds and their implications

The locality structure of spin lattice systems has a very important physical consequence: It effectively limits the speed of propagation of any kind of correlation. This speed limit is the *Lieb-Robinson speed* and it is proportional to the interaction strength[1]. A precise mathematical statement of this type is referred to as *Lieb-Robinson bound*. Such statements are important mathematical proof tools and have various consequences for the dynamic as well as static properties of lattice systems. Lieb-Robinson bounds are one of the most fundamental and important mathematical tools for spin lattice systems and their consequences are perfect examples of how a spatial locality structure can lead to simulatability statements. Therefore, it is worth having a more detailed but still introductory discussion on that.

[1] In the literature on Lieb-Robinson bounds, *interactions* refer to interactions between spatial sites, i.e., interactions include hopping terms.

On the next pages, we first present an introduction to open and closed spin and fermionic lattice systems, Lieb-Robinson bounds, and their implications. This text appeared as the Book Chapter [KGE14b] and already anticipates some of the results presented in Section 2.2, where, e.g., the quasi-locality of Markovian dynamics is proven. For copyright reasons we present the preprint version [KGE13] of this publication.

Lieb-Robinson bounds and the simulation of time evolution of local observables in lattice systems

Martin Kliesch, Christian Gogolin, and Jens Eisert

Dahlem Center for Complex Quantum Systems,
Freie Universität Berlin, 14195 Berlin, Germany

August 19, 2014

Abstract

This is an introductory text reviewing Lieb-Robinson bounds for open and closed quantum many-body systems. We introduce the Heisenberg picture for time-dependent local Liouvillians and state a Lieb-Robinson bound that gives rise to a maximum speed of propagation of correlations in many body systems of locally interacting spins and fermions. Finally, we discuss a number of important consequences concerning the simulation of time evolution and properties of ground states and stationary states.

1 Introduction

In lattice systems one might expect that, due to the locality of the interaction, there is some limit to the speed with which correlations can propagate. Similar to the light cone in special relativity, there should be a space time cone, outside of which a local perturbation of such a system should not be able to influence any measurement in a significant way. That this intuition can indeed be made rigorous was first shown by Elliott H. Lieb and Derek W. Robinson in a seminal work [38] in 1972.

Today, the term Lieb-Robinson bound generally refers to upper bounds on the speed of propagation of some measure of correlation. Outside the space time cone defined by this speed, any signal is typically exponentially suppressed in the distance. The results of Lieb and Robinson, originally derived in the setting of translation invariant 1D spin systems with short range, or exponentially decaying interactions [38] have since been tightened [27, 43] and extended to more general graphs [31, 47] and to interactions decaying only polynomially with the distance, both, for spin systems [44] and fermionic systems [31] (see also Ref. [45] for a review). Lieb-Robinson bounds have been proven for Liouvillian dynamics first in Ref. [55], where Liouvillian dynamics is a generalization of Hamiltonian dynamics that can also capture the effect of a certain type of noise. The bounds have recently been strengthened for a specific subclass of Liouvillians in Ref. [14] and have been generalized to time-dependent Liouvillian dynamics in Refs. [6, 48]. Indeed, Lieb-Robinson bounds provide the basis for a wealth of statements in quantum many-body theory, mostly as a mathematical proof tool, but also as an argument justifying numerical techniques. We will touch upon these implications and discuss the simulation of time evolution in more detail.

To keep the presentation both self-contained and concise, we mainly focus on Liouvillian dynamics as presented in Ref. [6]. The chapter is structured as follows: In the beginning,

Appeared as a chapter in the Book [73].

we introduce the setting and the necessary notation in Sect. 2. This includes in particular an introduction to Liouvillian dynamics in both the Schrödinger and Heisenberg picture and a discussion of the relevant measures for approximation errors that are needed to state the Lieb-Robinson bound and their physical interpretation. In the last part of Sect. 2 we explain the setting of spin lattice systems. Next, we state a general Lieb-Robinson bound in Sect. 3 and mention various consequences. In particular, we explain the locality and simulability of time evolution in more detail in Sect. 4. Finally, in Sect. 5, we state the Lieb-Robinson bound for fermions and introduce the Jordan-Wigner transform, which is a mapping between spin systems and fermionic systems.

2 Setting and notation

In this section we introduce the necessary formalism to describe the dynamics of spin lattice systems evolving under local Liouvillian dynamics, including local Hamiltonian dynamics as a special case. While Hamiltonian time evolution describes the dynamics of closed systems, Liouvillian dynamics also captures the case of so-called open quantum systems [39], which are systems coupled to memoryless "baths". Such couplings can be used to model Markovian "noise" perturbing the evolution of the system. The formalism and results discussed here partially address the problem of developing a better understanding of "imperfect systems" and, in particular, their time evolution (see also the chapter of Claude Le Bris).

2.1 Schrödinger and Heisenberg picture for time-dependent Liouvillians

We start by introducing some notation and some basic mathematical facts. For some Hilbert space $\mathcal{H}$ of finite dimension $\dim(\mathcal{H})$ let us denote the space of linear operators on $\mathcal{H}$ by $\mathcal{B}(\mathcal{H})$. Together with the *Hilbert-Schmidt* inner product, defined by $\langle A, B\rangle := \mathrm{Tr}(A^\dagger B)$ for $A, B \in \mathcal{B}(\mathcal{H})$, the space of operators $\mathcal{B}(\mathcal{H})$ is also a Hilbert space. Importantly, this defines the *Hilbert-Schmidt adjoint* of a *superoperator*. A superoperator is a linear map $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, i.e., $T \in \mathcal{B}(\mathcal{B}(\mathcal{H}))$ and its (Hilbert-Schmidt) adjoint $T^\dagger \in \mathcal{B}(\mathcal{B}(\mathcal{H}))$ is defined via $\langle X, T^\dagger(Y)\rangle := \langle T(X), Y\rangle$ for all $X, Y \in \mathcal{B}(\mathcal{H})$. The subspace of *observables* $\mathcal{A}(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$ are the Hermitian, i.e. self-adjoint operators and the set of *states* $\mathcal{S}(\mathcal{H})$ (also called density operators) are positive semidefinite Hermitian operators with unit trace. Given an observable $A \in \mathcal{A}(\mathcal{H})$ and a state $\rho \in \mathcal{S}(\mathcal{H})$ the expectation value is

$$\langle A\rangle_\rho := \mathrm{Tr}(\rho A). \tag{1}$$

When considering time evolution one is confronted with the following scenario: At some time s the system is in some initial state ρ and at a later time

$$t \geq s \qquad \text{(throughout this chapter)} \tag{2}$$

one measures some observable A that gives rise to an expectation value $\langle A\rangle_\rho(s,t)$. The time evolution can be described either in the Schrödinger picture or the Heisenberg picture. In the Schrödinger picture, one evolves the initial state ρ, given at time s, forward in time until time t is reached at which the measurement is performed. In the Heisenberg picture, in turn, one evolves the observable A backwards in time from t to the time s at which the initial state is given.

In the Schrödinger picture one considers the states to be time-dependent. In the case of a closed quantum system evolving under a Hamiltonian H, the state of the system at time t is

the solution of the linear initial value problem

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_s(t) = -\mathrm{i}[H(t), \rho_s(t)], \quad \rho_s(s) = \rho, \tag{3}$$

where the solutions of the dynamical equations carry the initial time s as a label for reasons that become clear once we switch to the Heisenberg picture. If a system is coupled to further degrees of freedom giving rise to decoherence and dissipation, one can, e.g., for many physically relevant situations with weak coupling, describe the system as an *open quantum system* whose dynamic is given by the solution of the linear initial value problem

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_s(t) = \mathcal{L}_t^\dagger(\rho_s(t)) \,, \quad \rho_s(s) = \rho, \tag{4}$$

where $\mathcal{L}^\dagger : \mathbb{R} \to \mathcal{B}(\mathcal{B}(\mathcal{H}))$ is called the *Liouvillian*[1], and where the time dependence is given by the input $t \in \mathbb{R}$. The Liouvillian may explicitly depend on time, e.g. to be able to capture change of external control parameters. Throughout this chapter we restrict the time dependence to be *piecewise continuous*. For an equation of motion of this form, the only constraint is that the time evolution maps states to states, i.e., is completely positive and trace preserving. This is equivalent [70] to the Liouvillian $\mathcal{L}_t^\dagger$ having a Lindblad representation [39], i.e. it must be of the form

$$\mathcal{L}^\dagger(\rho) = -\mathrm{i}[H, \rho] + \sum_{\mu=1}^{\dim(\mathcal{H})^2} \left(2L_\mu \rho L_\mu^\dagger - L_\mu^\dagger L_\mu \rho - \rho L_\mu^\dagger L_\mu\right), \tag{5}$$

for some time-dependent operators $H : \mathbb{R} \to \mathcal{A}(\mathcal{H})$ and $L_\mu : \mathbb{R} \to \mathcal{B}(\mathcal{H})$.

Liouvillian dynamics is ubiquitous in many contexts in physics. It has recently been studied particularly intensely in the context of cold atoms in optical lattices [3, 15, 16, 52], trapped ions [5, 57], driven dissipative Rydberg gases [23], and macroscopic atomic ensembles [35]. Also dissipative state preparation [69], dissipative phase transitions [15], noise-driven criticality [17] and nonequilibrium topological phase transitions [4] have been considered.

The initial value problem (4) defines the *propagator* (also called dynamical map) $T_{\mathcal{L}^\dagger}(t, s) : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ via

$$T_{\mathcal{L}^\dagger}(t, s)(\rho) \coloneqq \rho_s(t), \tag{6}$$

which is also the unique solution of the initial value problem

$$\frac{\mathrm{d}}{\mathrm{d}t}T_{\mathcal{L}^\dagger}(t, s) = \mathcal{L}_t^\dagger T_{\mathcal{L}^\dagger}(t, s) \,, \quad T_{\mathcal{L}^\dagger}(s, s) = \mathrm{id}. \tag{7}$$

The expectation value at time t then is

$$\langle A \rangle_\rho (s, t) = \mathrm{Tr}\left[T_{\mathcal{L}^\dagger}(t, s)(\rho)\, A\right]. \tag{8}$$

If the Liouvillian $\mathcal{L}^\dagger$ is time-independent, a state satisfying $\mathcal{L}^\dagger(\rho) = 0$ is called stationary state. The role played by stationary states is reminiscent of the role of ground states of Hamiltonians. For the case of a unique stationary state the spectral gap of the Liouvillian is a measure of the speed of convergence [33] towards this stationary state.

Evolving some state ρ from s to $r \geq s$ and then from r to $t \geq r$ also yields $\rho_s(t)$ and hence

[1] As we will later mostly work in the Heisenberg picture it is convenient to denote the Liouvillian in the Schrödinger picture by $\mathcal{L}^\dagger$ rather than $\mathcal{L}$.

the propagator has the *composition property* $T_{\mathcal{L}^\dagger}(t,r)T_{\mathcal{L}^\dagger}(r,s) = T_{\mathcal{L}^\dagger}(t,s)$ for all $t \geq r \geq s$. For classical processes this property is stated by the *Chapman-Kolmogorov equation*. It is a good exercise to derive the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}s} T_{\mathcal{L}^\dagger}(t,s) = -T_{\mathcal{L}^\dagger}(t,s)\mathcal{L}_s^\dagger, \tag{9}$$

from this property.

We are now ready to introduce the Heisenberg picture, in which the states are constant and the observables are defined as solutions of a dynamical equation. Of course, both pictures must yield the same expectation values, i.e.,

$$\langle A \rangle_\rho (s,t) = \mathrm{Tr}\left(\rho\, \tau_{\mathcal{L}}(t,s)(A)\right), \tag{10}$$

where

$$\tau_{\mathcal{L}}(s,t) = T_{\mathcal{L}^\dagger}(t,s)^\dagger \tag{11}$$

is the adjoint of $T_{\mathcal{L}^\dagger}(t,s)$ in the Hilbert-Schmidt inner product. $\tau_{\mathcal{L}}$ is the *propagator in the Heisenberg picture*. Using Eq. (9), it is not hard to see that it is the unique solution of

$$\frac{\mathrm{d}}{\mathrm{d}s}\tau_{\mathcal{L}}(s,t) = -\mathcal{L}_s \tau_{\mathcal{L}}(s,t)\,, \quad \tau_{\mathcal{L}}(t,t) = \mathrm{id}, \tag{12}$$

where $\mathcal{L}$ and $\mathcal{L}^\dagger$ are Hilbert-Schmidt adjoints of each other and, in particular, $\mathcal{L}$ is given by

$$\mathcal{L}(A) = \mathrm{i}[H,A] + \sum_{\mu=1}^{\dim(\mathcal{H})^2} \left(2L_\mu^\dagger A L_\mu - L_\mu^\dagger L_\mu A - A L_\mu^\dagger L_\mu\right). \tag{13}$$

Now we define the *(backward) time evolved* observable $A_t(s)$ to be the solution of

$$\frac{\mathrm{d}}{\mathrm{d}s} A_t(s) = \mathcal{L}_s(A_t(s))\,, \quad A_t(t) = A, \tag{14}$$

which is equivalent to

$$A_t(s) = \tau_{\mathcal{L}}(t,s)(A). \tag{15}$$

In the case of time-independent Liouvillians, one can equivalently define the Heisenberg picture such that observables are evolved forward in time. More generally, this is always possible if $\tau_{\mathcal{L}}(s,t)\mathcal{L}_t = \mathcal{L}_t\tau_{\mathcal{L}}(s,t)$ for all $s \leq t$, i.e., when the propagator commutes with the Liouvillian. In this case, one can equivalently evolve observables forward in time with $T_{\mathcal{L}^\dagger}(t,s)^\dagger$ which is then $T_{\mathcal{L}^\dagger}(t,s)^\dagger = T_{\mathcal{L}}(t,s)$. If the propagator and the Liouvillian do not commute, there is no simple way to obtain a consistent forward time evolution for A.

2.2 The physically relevant norms

Norms are functions that quantify the "size" of a vector or operator and hence provide an important tool to measure errors when approximating observables. Let us explain this in more detail. The Hilbert space inner product induces a norm via $\| |\psi\rangle \| \coloneqq \sqrt{\langle\psi|\psi\rangle}$. This norm gives rise to a norm on operators: let $B \in \mathcal{B}(\mathcal{H})$, then its *operator norm* is defined to be the supremum

$$\|B\| \coloneqq \sup_{\| |\psi\rangle\|=1} \|B\,|\psi\rangle\|\,, \tag{16}$$

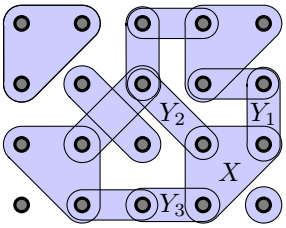

Fig. 1: An interaction hypergraph. The dots denote the vertices and the frames the hyperedges. The maximum number of nearest neighbors is $\mathcal{Z} = 4$: the edge X has the nearest neighbors Y_j and itself.

which coincides with the largest singular values of B. If B is an observable, then its norm is its largest eigenvalue in magnitude and thus a bound on the range of values one can obtain when B is measured, i.e.,

$$\|B\| = \sup_{\rho \in \mathcal{S}(\mathcal{H})} |\operatorname{Tr}(\rho B)|. \tag{17}$$

Considering the case where $B = A - A'$ is the difference of two observables $A, A' \in \mathcal{A}(\mathcal{H})$ this means that the operator norm is the physically relevant norm to measure closeness of the two observables: If $\|A - A'\|$ is small, then A and A' will have almost the same expectation value on all states, see Ref. [49] for a more detailed discussion.

2.3 Lattice systems and local Liouvillians

Quantum lattice systems are formally described by a set of (spatial) sites that are considered to be the vertices of a (hyper)graph. The interactions between the sites correspond to the edges of the (hyper)graph (see also Fig. 1). In this section we explain this setting for spin systems in detail and consider fermionic systems in Sect. 5.

Let us assume that the set of sites V is finite and that each site $x \in V$ is associated with a finite dimensional Hilbert space $\mathcal{H}_x$. The Hilbert space of some subsystem $X \subset V$ is denoted by $\mathcal{H}_X := \bigotimes_{x \in X} \mathcal{H}_x$ and $\mathcal{H} := \mathcal{H}_V$. For an operator $A \in \mathcal{B}(\mathcal{H})$ we define its support $\operatorname{supp}(A)$ to be the smallest subset $X \subset V$ such that it acts as the identity outside of X, i.e., $A_X = A \otimes \mathbb{1}_{V \setminus X}$. The set of operators supported on X is denoted by $\mathcal{B}_X(\mathcal{H}) := \{A \in \mathcal{B}(\mathcal{H}) : \operatorname{supp}(A) \subset X\}$ and the subspace of observables by $\mathcal{A}_X(\mathcal{H}) \subset \mathcal{B}_X(\mathcal{H})$. For a Liouvillian $\mathcal{L}$ on $\mathcal{B}(\mathcal{H})$ we define its support to be

$$\operatorname{supp}(\mathcal{L}) := \bigcup \{X \subset V : \mathcal{A}_{V \setminus X}(\mathcal{H}) \subset \ker(\mathcal{L})\}, \tag{18}$$

i.e., the part of the system where $\mathcal{L}$ corresponds to a non-trivial time evolution. The set of Liouvillians supported on X is denoted by $\mathbb{L}_X(\mathcal{H})$. Often we omit the Hilbert space and write, e.g., $\mathcal{A}_X$ instead of $\mathcal{A}_X(\mathcal{H})$.

We are interested in the time evolution under *local Liouvillians*. A Liouvillian $\mathcal{L}$ is called *local* if it is of the form

$$\mathcal{L} = \sum_{X \subset V} \mathcal{L}_X, \quad \mathcal{L}_X \in \mathbb{L}_X. \tag{19}$$

In many physically relevant situations many of the strictly local terms $\mathcal{L}_X$, in particular those belonging to large sets X, will be zero. This structure reflects interactions and dissipation processes that are finite-ranged. The *interaction graph* E of the Liouvillian is the set of all

subsets of V for which the Liouvillian contains a non zero term, i.e.,

$$E \coloneqq \{X \subset V : \mathcal{L}_X \neq 0\}. \tag{20}$$

As an example, consider the case of a 1D system with nearest neighbor interactions and open boundary conditions. If the sites are $V = \{1, \ldots, N\}$, the interaction graph is $E = \{\{1,2\}, \{2,3\}, \ldots, \{N-1, N\}\}$ in that case.

The interaction (hyper)graph E defines a distance $\mathrm{d}(X,Y)$ between any two sets $X, Y \subset V$ of vertices. The distance $\mathrm{d}(X,Y)$ is equal to 0 if and only if $X \cap Y \neq \emptyset$ and otherwise equal to the length of the shortest path connecting X and Y, and ∞ if there is no connecting path. A path between two sets $X, Y \subset V$ is a sequence of elements of E, such that the first element contains a vertex in X, each element of the path shares at least one vertex with the following element and the last element contains a vertex in Y. Note that d is a degenerate metric on subsets of V. In the above 1D example the graph distance of the two sets $\{j\}, \{k\} \subset V$ would simply be $\mathrm{d}(\{j\}, \{k\}) = |j - k|$, as one would expect.

3 A Lieb-Robinson bound

In this section we state and explain a very general Lieb-Robinson bound for the speed of propagation of correlations in spin systems under arbitrary time-dependent Liouvillian dynamics. Our goal is to make statements about *local* time evolution, i.e., time evolution of local observables arising from local interactions and local noise. In order to make this precise, let us impose some technical constraints on a possibly time-dependent Liouvillian $\mathcal{L}_s \in \mathbb{L}_V$, which we consider to be fixed from now on. Local time evolution is captured by a Liouvillian $\mathcal{L}$ that is a sum of strictly local terms $\mathcal{L}_X$, each of which is bounded in norm by b, and a maximum number of nearest neighbors $\mathcal{Z}$. In more detail, we define

$$\mathcal{L} = \sum_{X \in E} \mathcal{L}_X, \quad \mathcal{L}_X : \mathbb{R} \to \mathbb{L}_X(\mathcal{H}), \text{ piecewise continuous,} \tag{21}$$

$$b \coloneqq \sup_{s,X} \|\mathcal{L}_X(s)\|, \tag{22}$$

$$\mathcal{Z} \coloneqq \max_{X \in E} |\{Y \in E : Y \cap X \neq \emptyset\}|. \tag{23}$$

The parameters b and $\mathcal{Z}$ will determine the Lieb-Robinson speed and also the final results about the *spatial truncation*

$$\mathcal{L}_{\restriction V'} \coloneqq \sum_{X \subset V'} \mathcal{L}_X \tag{24}$$

of the Liouvillian $\mathcal{L}$ to some region $V' \subset V$. Now we are ready to state the Lieb-Robinson bound for this setting. Similar results on Liouvillians can be found in Refs. [48, 55]. The theorem is quite general and it might not be immediately obvious how statements about propagation of information are implied. But this will become clear in the next section.

Theorem 1 (Lieb-Robinson Bound [6][2])**.** *Let $\mathcal{L} : \mathbb{R} \to \mathbb{L}(\mathcal{H})$ be a local Liouvillian as specified in Eqn. (21) – (23) and $X, Y \subset V$. Then, for every $\mathcal{K}_Y \in \mathbb{L}_Y(\mathcal{H})$, $A_X \in \mathcal{B}_X(\mathcal{H})$, and $s \leq t$*

$$\|\mathcal{K}_Y \tau_{\mathcal{L}}(s,t)(A_X)\| \leq C \|\mathcal{K}_Y\| \|A_X\| \, \mathrm{e}^{v(t-s) - \mathrm{d}(X,Y)}, \tag{25}$$

[2] In Ref. [6] the bound is given for an arbitrary metric on the vertex set and the Liouvillians are allowed to have interaction range a in that metric. Our interaction graph distance d is induced by a metric on V for which $a = 1$.

where $v = \exp(1)b\mathcal{Z}$ and C is some constant depending polynomially on the size of the smaller of the two sets X and Y.

Remembering that the Liouvillian maps an observable to its time derivative. The theorem tells us that an evolved observable $\tau_{\mathcal{L}}(s,t)(A_X)$ remains basically unchanged when evolved with respect to a Liouvillian $\mathcal{K}_Y$ that is supported on a region a distance much larger than $v(t-s)$ away from X, i.e., that $\tau_{\mathcal{L}}(s,t)(A_X)$ is almost the identity outside the corresponding space-time cone. More intuitively, the Lieb-Robinson bound tells us that information travels with a velocity bounded by the *Lieb-Robinson speed* v of the considered lattice system. In the special case $\mathcal{K}_Y = i[B_Y, \cdot\,]$ for some $B_Y \in \mathcal{A}_Y(\mathcal{H})$, Eq. (25) yields a Lieb-Robinson bound in the more common form of an upper bound on the commutator $\|[B_Y, \tau_{\mathcal{L}}(s,t)(A_X)]\|$ (compare Refs. [48, 55]).

If a system is mixing in the sense that all states are driven towards a steady state then information encoded in the initial state gets lost at some point. This puts an upper bound on the distance over which information can propagate. Therefore, one might expect that there is some effective Lieb-Robinson speed that decreases in time. This is indeed true for certain systems with fluctuating disorder [10] and for a certain class of Liouvillian dynamics [14].

Finally let us mention that, the lattice can also be infinitely large (implied by the next theorem), but the restriction to finite-dimensional subsystems is not merely for simplicity of notation: For infinite-dimensional systems the situation can be quite different. For some anharmonic lattices [47], and other instances of strongly correlated models [62] Lieb-Robinson bounds can still be found, as well as for commutator-bounded operators [56]. Still, counterexamples to Lieb-Robinson bounds are known for models with infinite-dimensional constituents [19].

4 Consequences of Lieb-Robinson bounds

Lieb-Robinson bounds are fundamental for a plethora of statements concerning various properties of locally interacting systems. We first discuss immediate consequences as far as the dynamics of such systems is concerned. Next, we turn to implications for the classical simulation of time evolution. Finally, we discuss static properties that can be derived from Lieb-Robinson bounds.

4.1 Quasi-locality of quantum dynamics

The result of the last section suggests that the terms of the Liouvillian whose support is sufficiently far away from the support of an observable are irrelevant for the time evolution. More precisely, one should be able to spatially truncate the Liouvillian $\mathcal{L}$ to some region $V' \subset V$. If X is sufficiently far from the boundary of V', i.e., if $\mathrm{d}(X, V \setminus V')$ is larger than the radius $v\cdot(t-s)$ of the space time cone of $\tau_{\mathcal{L}}(s,t)(A_X)$, then the dynamics of A_X under the truncated Liouvillian $\mathcal{L}_{\restriction V'}$ and the original Liouvillian $\mathcal{L}$ should be very similar. In the next theorem we will see that this is indeed the case if the underlying interaction graph is of finite spacial dimension, which we define first. Let us denote the "sphere" around some subsystem $X \in E$ with radius n by

$$S_X(n) \coloneqq \{Y \in E : \mathrm{d}(Y, X) = n\}. \tag{26}$$

Then we say that an interaction graph E is of *spatial dimension* μ if there is a constant $M > 0$ that only depends on local properties of the interaction graph such that for all $X \in E$

$$|S_X(n)| \leq M n^{\mu-1}. \tag{27}$$

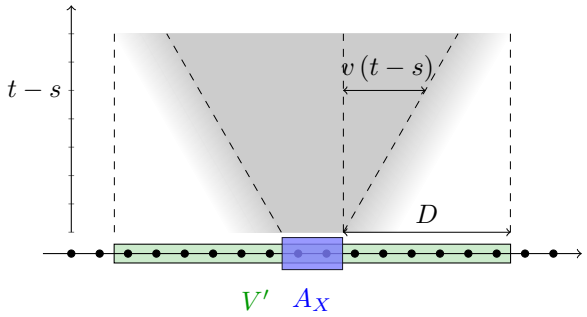

Fig. 2: The space time cone of an observable A_X in one spatial dimension. The truncation error scales exponentially in the distance $D - v \cdot (t-s)$.

For example, the interaction graph of next-neighbor Liouvillians on a μ-dimensional cubic lattice has dimension μ.

The bound from the following theorem is visualized in Fig. 2.

Theorem 2 (Quasi-locality of local Liouvillian dynamics [6])**.** *Let $\mathcal{L} : \mathbb{R} \to \mathbb{L}(\mathcal{H})$ be a local Liouvillian as specified in Eqs. (21) – (23) and let its interaction graph be of spatial dimension μ with the constant M as defined in Eq. (27). Then, for all $X \subset V' \subset V$ with $D \coloneqq \mathrm{d}(X, V \setminus V') \geq 2\mu - 1$, $A_X \in \mathcal{B}_X(\mathcal{H})$, and $s \leq t$*

$$\left\| \tau_{\mathcal{L}_{\restriction V'}}(s,t)(A_X) - \tau_{\mathcal{L}}(s,t)(A_X) \right\| \leq \frac{2M}{\mathcal{Z}} D^{\mu-1} \mathrm{e}^{v\cdot(t-s)-D} \left\| A_X \right\|, \tag{28}$$

where $v = \exp(1) b \mathcal{Z}$ is the Lieb-Robinson speed.

So, colloquially speaking, the full dynamics of local observables can be approximated with exponential accuracy by the dynamics of a sufficiently large subsystem. Of course, the size of the subsystem depends on the desired time span of the evolution. In particular the locality result makes an extension of time evolution to infinitely large lattices possible, i.e., it can be employed to rigorously define the thermodynamic limit.

Theorem 2 has two further immediate physical consequences, which can be seen as an interpretation of the result. For the rest of this section consider a lattice system with $V = \{1, \ldots, N\}$ and let ρ be a product state, i.e., $\rho = \bigotimes_{j=1}^{N} \rho_j$ where $\rho_j \in \mathcal{S}(\mathcal{H}_{\{j\}})$ for all j and moreover, let $X, Y \subset V$ such that $X \cap Y = \emptyset$.

Suppression of correlation functions: Consider two observables $A_X \in \mathcal{A}_X$ and $B_Y \in \mathcal{A}_Y$. Their correlation coefficient in state $\rho \in \mathcal{S}(\mathcal{H})$ is proportional to the *covariance*

$$\mathrm{cov}_\rho(A_X, B_Y) \coloneqq \langle A_X B_Y \rangle_\rho - \langle A_X \rangle_\rho \langle B_Y \rangle_\rho . \tag{29}$$

If ρ is a product state, $\mathrm{cov}_\rho(A_X, B_Y) = 0$. Now, Theorem 2 tells us that as long as $v \cdot (t-s) \ll \mathrm{d}(X,Y)/2$ the correlation coefficient of the time evolved observables will remain very small. More precisely, $\mathrm{cov}_\rho(\tau(s,t)(A_X), \tau(s,t)(B_Y))$ is upper bounded by $\exp(v \cdot (t-s) - \mathrm{d}(X,Y)/2)$ up to a constant factor. The measurement statistics of the two observables can show correlations only after the dynamics of the system had enough time to correlate the two regions X and Y (see Ref. [44] for a similar discussion in the context of Hamiltonian dynamics).

Speed of information propagation: Time evolution on a lattice system can also be thought of as a channel that one might want to use to send information from X to Y in the following

way: One party encodes a message by preparing at time s the part of the initial state ρ in the region X in a particular way, the other party tries to retrieve the message by measuring on region Y at time t. Lieb-Robinson bounds can be used to show that the amount of information that can be transferred in this way in a time span $t-s$ is exponentially suppressed if $\mathrm{d}(X,Y)$ is larger than $v\cdot(t-s)$. This can be made precise in the sense that the classical information capacity is exponentially small outside the cone, if the quantum many-body systems is used as a quantum channel [8]. In Ref. [11], the ballistic propagation of excitations and information propagation constrained by Lieb-Robinson bounds has been experimentally explored in systems of cold atoms.

4.2 Classical simulation of quantum dynamics

By classical simulation of quantum dynamics we mean the calculation of expectation values of local observables $\langle A_X\rangle_\rho(s,t)$, so that one could, for instance, plot them over time. If one tries to do that naively, i.e., by calculating the full propagator $\tau_{\mathcal{L}}(s,t)$ on a classical computer, one quickly runs into problems even with just having enough memory to store the propagator. For instance, if one has N subsystems with a local Hilbert space dimension of 2, then to completely specify the propagator in a naive way, one needs 2^{4N} complex numbers. Therefore, if one aims at classically simulating local observables one needs to come up with a smart simulation scheme that only deals with the information relevant for the simulation. We sketch two such schemes here:

Time evolution as (unitary) circuits: Here the quasi-locality Theorem 2 is of great help, since it already tells us that one can truncate the dynamics to a set V' containing the space time cone of the observable instead of considering the full system V. The arising error is exponentially small in the distance between the space time cone and the truncation. So the simulation cost does clearly not depend on the system size and the dynamics can hence be implemented efficiently in that. Of course, implementing the full simulation naively on V' is still by far not optimal. Famously, one can decompose the propagator $\tau_{\mathcal{L}_{\restriction V'}}(s,t)$ into products over short time steps and strictly local propagators, which is often called Trotter-decomposition [64]. At the heart of this approach is the following *product formula* that can be used to bound the error one makes by decomposing the propagator of a Liouvillian that is a sum of two Liouvillians $\mathcal{L}$ and $\mathcal{K}$ into the product of the propagators of these Liouvillians:

Theorem 3 (Trotter product formula [6, 34])**.** *Let*

$$\mathcal{L}=\sum_{X\in E}\mathcal{L}_X \tag{30}$$

be a Liouvillian with $\mathcal{L}_X\in\mathbb{L}_X$. Then there exist constants b and c that depend only on local properties of $\mathcal{L}$, and are in particular independent of the number of sites, such that for all $X\in E$ and operators A

$$\|\tau_{\mathcal{L}}(s,t)(A)-\tau_{\mathcal{L}}(s,t)\tau_{\mathcal{L}-\mathcal{L}_X}(s,t)(A)\|\leq c(t-s)^2\mathrm{e}^{b(t-s)}|E|\,\|A\|\,. \tag{31}$$

One can now decompose the time span $t-s$ into short time steps $[s_{j+1},s_j]$ and in each of these intervals approximate the propagator by a product of the strictly local propagators $\tau_{\mathcal{L}_X}(s_{j+1},s_j)$ for each edge X in the interaction graph of the Liouvillian. In other words, the full propagator can be approximated by a "circuit" of strictly local propagators. The number of time steps needed to reach a simulation with total error upper bounded by ϵ is proportional to $(t-s)^2|E|^2/\epsilon$ [34]. Of course, the above covers Hamiltonian dynamics as a special case. However, there one would rather apply similar ideas to the time evolution oper-

ator $\exp(-\mathrm{i}\,(t-s)H)$ rather than the propagator. In a variant of this circuit description for Hamiltonian dynamics in 1D, the time evolution operator can be approximated by a circuit of constant depth and time-dependent gates [50].

Time-dependent density-matrix renormalization group methods: A similar mindset is also fundamental for the simulation of time evolution using so-called tensor network states. The situation is particularly clear in 1D systems with sites $V = \{1, \ldots, N\}$ in pure states undergoing local Hamiltonian dynamics. If the initial state has a strong decay of initial correlations, then the time evolution can for short times be efficiently grasped in terms of *matrix-product states* (MPS) [22, 51, 59]. These are variational state vectors that are described by $O(d\,ND^2)$ variational parameters, where $D \in \mathbb{N}$ is a refinement parameter and d the dimension of the local Hilbert space. There are several variants of this approach, based on either a Trotter-decomposition [64] or a time-dependent variational principle [25]. Such schemes are subsumed under the term *time-dependent density matrix renormalization group method* (t-DMRG). At the heart of the functioning of t-DMRG is the insight that states generated by short time local Hamiltonian dynamics will have *low entanglement*. This can be formalized [18] in terms of so-called area laws [2, 20, 26, 65] that arise as a consequence of a Lieb-Robinson bound.

An area law is an upper bound on the entanglement of a state. More precisely, we say that a pure state satisfies an *area law* if for any region $R \subset V$ the (Rényi) entropy of the reduced state on R can be bounded by the size of the boundary of R, up to a constant. States of 1D systems satisfying an area law can be provably well approximated by matrix product states [68]. Indeed, t-DMRG simulates time evolution for short times to essentially machine precision. For long times, the entropy will in general grow too much, as then sites are in the space time cone of too many sites of the lattice, and an efficient simulation in terms of matrix-product states is hence [60] no longer possible [12, 61]. That is, the power of the t-DMRG approach can be rigorously grasped in terms of Lieb-Robinson bounds. For 1D local Liouvillian dynamics, variants of t-DMRG have also been proposed [66, 72], usually as variational principles over *matrix-product operators*, the mixed state analogues of matrix-product states, or by means of suitable sampling employing classical stochastic processes in Hilbert space [52].

4.3 Static properties derived from Lieb-Robinson bounds

Among the most important applications of Lieb-Robinson bounds are proof techniques related to static (time independent) properties of quantum lattice systems. Here we briefly mention some of them:

Clustering of correlations in Hamiltonian systems: One of the most relevant applications concerns the decay of correlations in the ground state of a local Hamiltonian with a spectral gap[3], first shown in Refs. [27, 31] and further generalized in Ref. [46]. The basic intuition underlying this intricate insight is that the spectral gap ΔE essentially defines a time scale in the system, which in turn can be related to a length scale.

Theorem 4 (Clustering of correlations in unique ground states [31, 43])**.** *Let $H \in \mathcal{A}(\mathcal{H})$ be a local Hamiltonian with a unique ground state ψ and a spectral gap $\Delta E > 0$ and $X, Y \subset V$. Then, for every $A_X \in \mathcal{B}_X(\mathcal{H})$ and $B_Y \in \mathcal{B}_Y(\mathcal{H})$*

$$|\mathrm{cov}_\psi(A_X, B_Y)| \leq C\,\|A_X\|\,\|B_Y\|\,\mathrm{e}^{-\mu\,\mathrm{d}(X,Y)}. \tag{32}$$

C and μ are constants both depending on ΔE. Moreover, C depends on the lattice geometry

[3] The spectral gap of a Hamiltonian ΔE is the difference between the ground state energy and the energy of the first exited state.

and the smaller of the surface areas of X and Y, and μ depends on the Lieb-Robinson speed.

The proof of this statement confirmed a long-standing conjecture in condensed-matter physics, that gapped Hamiltonian systems have exponentially clustering correlations in the ground state.

Clustering of correlations in Liouvillian systems: A similar intuition actually holds true for Liouvillian systems, where the role of the ground state of Hamiltonian systems is taken over by the stationary state. Clustering of correlations in local Liouvillian systems has first been considered in Ref. [55] and has been made rigorous and largely generalized in Ref. [33]: If a local Liouvillian is primitive (that is, if its stationary state has full rank) and has a spectral gap which is independent of the system size, then correlation functions between local observables again decay exponentially as a function of the distance between their supports.

Area laws of ground states of gapped Hamiltonians: It has been shown using Lieb-Robinson bounds that ground states of 1D local Hamiltonian systems with spectral gap $\Delta E > 0$ always satisfy an area law for the Rényi entropies (for a review, see Ref. [20]). This result has since been tightened [2] and area laws have also been shown for some instances of gapped higher-dimensional Hamiltonian systems [54]. It has also been shown that in 1D exponential clustering of correlations already implies an area law [7]. For local Liouvillians, general area laws (in terms of entropic measures suitable for mixed states) can be derived for stationary states [33], again using Lieb-Robinson bounds.

Approximating 1D ground states of gapped Hamiltonians with MPS: Since ground states of any 1D local Hamiltonian with a spectral gap $\Delta E > 0$ satisfy an area law for Rényi entropies they can be approximated [68] by matrix product states (MPS) in polynomial time [36]. This is used by the static *density-matrix renormalization group method* (DMRG) [58] (see also the chapter of Ors Legeza, Thorsten Rohwedder and Reinhold Schneider) for simulating ground state properties [59], which has led to a wealth of novel insights in condensed matter physics.

Higher-dimensional Lieb-Schultz-Mattis theorems: The Lieb-Schultz-Mattis theorem [1, 37] is an upper bound on the spectral gap of certain one-dimensional quantum spin systems. Using Lieb-Robinson bounds, a higher-dimensional Lieb-Schultz-Matthis theorem has been proven in Refs. [27, 29].

Stability and further properties of ground states: Lieb-Robinson bounds are one of the pillars of the formalism grasping the stability of ground states of a certain class of Hamiltonians (frustration-free Hamiltonians satisfying certain topological order conditions) under local perturbations. This has developed into a field of research in its own right, and we merely touch upon the topic here. Starting point is the concept of *quasi-adiabatic continuation* [30], which is a tool to connect dynamical properties of a Hamiltonian to static ones and relies on Lieb-Robinson bounds. Importantly, quasi-adiabatic continuation is a cornerstone of the proof of the stability of topological order under local perturbations [9] and related proofs of the stability of the spectral gap, of frustration-free Hamiltonians under general, quasi-local perturbations [41]. With similar tools, the stability of the area law for the entanglement entropy of the ground state can be proven [42, 65].

Stability of stationary states: Inspired by the stability results on Hamiltonian ground states, Lieb-Robinson bounds have also been used to prove the stability of stationary states of certain local Liouvillians [13, 33].

Structure of elementary excited states: The structure of elementary excited states has been explored using Lieb-Robinson bounds, which can be approximated by superimposing ground states to which local operators have been applied [24].

5 Fermionic Hamiltonians

While Lieb-Robinson bounds are usually stated for spin lattice system, they also hold for systems of fermions on a lattice. The situation is particularly simple for 1D systems with nearest neighbor coupling only, since in that case the Jordan-Wigner transform can be applied. In this section we first state a fermionic Lieb-Robinson bound and then introduce the Jordan-Wigner transform.

Again, as with spin lattice systems, we have an interaction (hyper)graph (V, E) but now work in the picture of second quantization, i.e., operators are given in terms of the fermionic creation and annhilation operators f_j and $f_k^\dagger$ for $j, k \in V$. These fermionic operators satisfy

$$\{f_j, f_k^\dagger\} = \delta_{j,k}, \tag{33}$$

where $\{A, B\} \coloneqq AB + BA$ is the anti-commutator. According to the *fermion number parity superselection rule* only observables that are even polynomials in the fermionic operators can occur in nature. A polynomial of fermionic operators is called *even* if it can be written as a linear combination of monomials, where each monomial is a product of an even number of fermionic operators from f_j and $f_k^\dagger$. Correspondingly, we denote the algebra of the parity preserving observables acting on a region $X \subset V$ by $\mathcal{G}_X$ for short. Now one can prove a fermionic Lieb-Robinson bound in the same way as Theorem 1 is proven:

Theorem 5 (Fermionic Lieb-Robinson bound)**.** *Let*

$$H = \sum_{X \in E} H_X \tag{34}$$

be a local time-dependent Hamiltonian with $H_X : \mathbb{R} \to \mathcal{G}_X$ and $\|H_X(r)\| \leq b$ for all $X \in E$ and $r \in \mathbb{R}$, τ its propagator, and $\mathcal{Z}$ the maximum number of nearest neighbors as defined in Eq. (23). Then, for every $A_X \in \mathcal{G}_X$, $B_Y \in \mathcal{G}_Y$ and $s, t \in \mathbb{R}$,

$$\|[B_Y, \tau(s,t)(A_X)]\| \leq C \|B_Y\| \|A_X\| \, \mathrm{e}^{v|t-s|-\mathrm{d}(X,Y)}, \tag{35}$$

where $v = \exp(1) b \mathcal{Z}$ and C is some constant depending polynomially on the size of the smaller of the two sets X and Y.

For the unphysical case where B_Y and A_X are observables that are odd polynomials in the fermionic operators one can still prove a similar Lieb-Robinson bound for the anti-commutator, providing a relevant proof-tool [28].

For the case of 1D systems with nearest neighbor interactions only, the analogy between fermionic and spin systems is even stronger in the sense that such systems can be mapped to each other by the Jordan-Wigner transform [32]. Note that a higher-dimensional variant has also been developed [67].

Consider a one-dimensional lattice with vertices $V = \{1, \ldots, N\}$. The Hilbert space of the spin-$1/2$ model on V is given by $\mathcal{H} \coloneqq \bigotimes_{j \in V} \mathcal{H}_j$ with $\mathcal{H}_j \cong \mathbb{C}^2$. We denote by $X_j, Y_j, Z_j \in \mathcal{A}_{\{j\}}$ the Pauli operators acting on site j of the spin chain. Then the Jordan-Wigner-Transformation is given by

$$f_j + f_j^\dagger = w_{2j-1} \coloneqq X_j \prod_{j'<j} Z_{j'} \tag{36}$$

$$\mathrm{i} f_j - \mathrm{i} f_j^\dagger = w_{2j} \coloneqq Y_j \prod_{j'<j} Z_{j'}, \tag{37}$$

where the $(w_j)_{j=1}^{2N}$ are called *Majorana operators*. The Majorana operators satisfy the anti-commutation relation $\{w_j, w_k\} = 2\delta_{j,k}$. It can be verified with elementary calculations that

$$f_j = \frac{1}{2}\left(w_{2j-1} - \mathrm{i}w_{2j}\right), \tag{38}$$

$$f_j^\dagger f_j = \frac{1}{2}\left(1 - \mathrm{i}w_{2j-1}w_{2j}\right), \tag{39}$$

as well as

$$Z_j = -\mathrm{i}w_{2j-1}w_{2j} = 2f_j^\dagger f_j - 1, \tag{40}$$

$$X_j = w_{2j-1}\prod_{j'<j} Z_{j'}, \qquad Y_j = w_{2j}\prod_{j'<j} Z_{j'}, \tag{41}$$

and

$$\forall j \leq k: \quad f_j^\dagger f_k = \frac{1}{4}S_j^+\Big(\prod_{j\leq j'<k} Z_{j'}\Big)S_k^-, \quad \text{where} \quad S_j^\pm \coloneqq X_j \pm \mathrm{i}Y_j. \tag{42}$$

Most importantly, as can be seen from Eq. (42), the Jordan-Wigner-Transformation preserves locality in the sense that a one-dimensional fermionic Hamiltonian with nearest neighbor or short range hopping and short range density-density interactions is mapped to a spin chain Hamiltonian with only short range interactions.

6 Conclusion

We have reviewed the Heisenberg picture for time-dependent Liouvillian dynamics in spin lattice systems. For this setting we have stated a Lieb-Robinson bound. Such bounds give rise to a plethora of statements about locally interacting systems which we have reviewed subsequently. Finally, we have explained the relevance for fermionic systems. We hope that this text serves as an introduction to Liouvillian dynamics on spin lattice systems and provides an overview of important consequences of Lieb-Robinson bounds.

Acknowledgments

We thank Earl T. Campbell, Mathis Friesdorf and Albert H. Werner for comments. We acknowledge support from the EU (Q-Essence, Raquel), the BMBF (QuOReP), the ERC (Taq), and the Studienstiftung des Deutschen Volkes.

References

[1] I. Affleck and E. H. Lieb, Lett. Math. Phys. **12**, 57 (1986).

[2] I. Arad, Z. Landau, and U. Vazirani, Phys. Rev. B **85**, 195145 (2012).

[3] C. Ates, B. Olmos, W. Li, and I. Lesanovsky, Phys. Rev. Lett. **109**, 233003 (2012).

[4] C.-E. Bardyn, M. A. Baranov, E. Rico, A. Imamoglu, P. Zoller, and S. Diehl, Phys. Rev. Lett. **109**, 130402 (2012).

[5] J. T. Barreiro, M. Müller, P. Schindler, D. Nigg, T. Monz, M. Chwalla, M. Hennrich, C. F. Roos, P. Zoller, and R. Blatt, Nature **470**, 486 (2011).

[6] T. Barthel and M. Kliesch, Phys. Rev. Lett. **105**, 010502 (2010).

[7] F. G. S. L. Brandão and M. Horodecki, Nat. Phys. **9**, 721 (2013).

[8] S. Bravyi, M. B. Hastings, and F. Verstraete, Phys. Rev. Lett. **97**, 050401 (2006).

[9] S. Bravyi, M. Hastings, and S. Michalakis, J. Math. Phys. **51**, 093512 (2010).

[10] C. K. Burrell, J. Eisert, and T. J. Osborne, Phys. Rev. A **80**, 052319 (2009).

[11] M. Cheneau, P. Barmettler, D. Poletti, M. Endres, P. Schauß, T. Fukuhara, C. Gross, I. Bloch, C. Kollath, and S. Kuhr, Nature **481**, 484 (2012).

[12] M. Cramer, C. M. Dawson, J. Eisert, and T. J. Osborne, Phys. Rev. Lett. **100**, 030602 (2008).

[13] T. S. Cubitt, A. Lucia, S. Michalakis, and D. Perez-Garcia, arXiv:1303.4744.

[14] B. Descamps and F. Verstraete, J. Math. Phys. **54**, 092202 (2013).

[15] S. Diehl, A. Micheli, A. Kantian, B. Kraus, H. P. Büchler, and P. Zoller, Nat. Phys. **4**, 878 (2008).

[16] S. Diehl, W. Yi, A. J. Daley, and P. Zoller, Phys. Rev. Lett. **105**, 227001 (2010).

[17] J. Eisert and T. Prosen, arXiv:1012.5013.

[18] J. Eisert and T. J. Osborne, Phys. Rev. Lett. **97**, 150404 (2006).

[19] J. Eisert and D. Gross, Phys. Rev. Lett. **102**, 240501 (2009).

[20] J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys. **82**, 277 (2010).

[21] C. Emary, C. Pöltl, A. Carmele, J. Kabuss, A. Knorr, and T. Brandes, Phys. Rev. B **85**, 165417 (2012).

[22] M. Fannes, B. Nachtergaele, and R. F. Werner, Commun. Math. Phys. **144**, 443 (1992).

[23] A. W. Glaetzle, R. Nath, B. Zhao, G. Pupillo, and P. Zoller, Phys. Rev. A **86**, 043403 (2012).

[24] J. Haegeman, S. Michalakis, B. Nachtergaele, T. J. Osborne, N. Schuch, and F. Verstraete, Phys. Rev. Lett. **111**, 080401 (2013).

[25] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pizorn, H. Verschelde, and F. Verstraete, Phys. Rev. Lett. **107**, 070601 (2011).

[26] M. B. Hastings, JSTAT P08024 (2007).

[27] M. B. Hastings, Phys. Rev. B **69**, 104431 (2004).

[28] M. B. Hastings, Phys. Rev. Lett. **93**, 126402 (2004).

[29] M. B. Hastings, Europhys. Lett. **70**, 824 (2005).

[30] M. B. Hastings and X.-G. Wen, Phys. Rev. B **72**, 045141 (2005).

[31] M. B. Hastings and T. Koma, Commun. Math. Phys. **265**, 781 (2006).

[32] P. Jordan and E. Wigner, Z. Phys. **47**, 631 (1928).

[33] M. Kastoryano and J. Eisert, J. Math. Phys. **54**, 102201 (2013).

[34] M. Kliesch, T. Barthel, C. Gogolin, M. Kastoryano, and J. Eisert, Phys. Rev. Lett. **107**, 120501 (2011).

[35] H. Krauter, C. A. Muschik, K. Jensen, W. Wasilewski, J. M. Petersen, J. I. Cirac, and E. S. Polzik, Phys. Rev. Lett. **107**, 080503 (2011).

[36] Z. Landau, U. Vazirani, and T. Vidick, arXiv:1307.5143.

[37] E. H. Lieb, T. D. Schultz, and D. C. Mattis, Ann. Phys. **16**, 407 (1961).

[38] E. H. Lieb and D. W. Robinson, Commun. Math. Phys. **28**, 251 (1972).

[39] G. Lindblad, Comm. Math. Phys. **48**, 119 (1976).

[40] D. Loss and E. V. Sukhorukov, Phys. Rev. Lett. **84**, 1035 (2000).

[41] S. Michalakis and J. Pytel, Commun. Math. Phys. **322**, 277 (2013).

[42] S. Michalakis, arXiv:1206.6900.

[43] B. Nachtergaele and R. Sims, in *New Trends in Mathematical Physics. Selected contributions of the XVth International Congress on Mathematical Physics*, edited by V. Sidoravicius (Springer, Heidelberg, 2009), pp. 591–614.

[44] B. Nachtergaele, Y. Ogata, and R. Sims, J. Stat. Phys. **124** (2006).

[45] B. Nachtergaele, R. Sims, IAMP News Bulletin, 22 (2010).

[46] B. Nachtergaele and R. Sims, Commun. Math. Phys. **265**, 119 (2006).

[47] B. Nachtergaele, H. Raz, B. Schlein, and R. Sims, Commun. Math. Phys. **286**, 1073 (2009).

[48] B. Nachtergaele, A. Vershynina, and V. A. Zagrebnov, Contemp. Math. **552**, 161 (2011).

[49] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[50] T. J. Osborne, Phys. Rev. Lett. **97**, 157202 (2006).

[51] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Quantum Inf. Comput. **7**, 401 (2007).

[52] H. Pichler, A. J. Daley, and P. Zoller, Phys. Rev. A **82**, 063605 (2010).

[53] C. Pineda, T. Barthel, and J. Eisert, Phys. Rev. A **81**, 050303(R) (2010).

[54] M. B. Plenio, J. Eisert, J. Dreissig, and M. Cramer, Phys. Rev. Lett. **94**, 060503 (2005).

[55] D. Poulin, Phys. Rev. Lett. **104**, 190401 (2010).

[56] I. Prémont-Schwarz, A. Hamma, I. Klich, and F. Markopoulou-Kalamara, Phys. Rev. A **81**, 040102(R) (2010).

[57] P. Schindler, M. Müller, D. Nigg, J. T. Barreiro, E. A. Martinez, M. Hennrich, T. Monz, S. Diehl, P. Zoller, and R. Blatt, Nat. Phys. **9**, 361 (2013).

[58] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).

[59] U. Schollwöck, Ann. Phys. **326**, 96 (2011).

[60] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **100**, 030504 (2008).

[61] N. Schuch, M. M. Wolf, K. G. H. Vollbrecht, and J. I. Cirac, New J. Phys. **10**, 033032 (2008).

[62] N. Schuch, S. K. Harrison, T. J. Osborne, and J. Eisert, Phys. Rev. A **84**, 032309 (2011).

[63] M. J. A. Schuetz, E. M. Kessler, J. I. Cirac, and G. Giedke, Phys. Rev. B **86**, 085322 (2012).

[64] H. F. Trotter, Proc. Am. Math. Soc **10**, 545 (1959).

[65] K. Van Acoleyen, M. Marien, and F. Verstraete, Phys. Rev. Lett. **111**, 170501 (2013).

[66] F. Verstraete, J. J. Garcia-Ripoll, and J. I. Cirac, Phys. Rev. Lett. **93**, 207204 (2004).

[67] F. Verstraete and J. I. Cirac, J. Stat. Mech. **09**, P09012 (2005).

[68] F. Verstraete and J. I. Cirac, Phys. Rev. B **73**, 094423 (2006).

[69] F. Verstraete, M. M. Wolf, and J. I. Cirac, Nat. Phys. **5**, 633 (2009).

[70] M. M. Wolf and J. I. Cirac, Commun. Math. Phys. **279**, 147 (2008).

[71] M. Žnidarič, B. Žunkovič, and T. Prosen, Phys. Rev. E **84**, 051115 (2011).

[72] M. Zwolak and G. Vidal, Phys. Rev. Lett. **93**, 207205 (2004).

[73] L. D. Site and V. Bach (eds.), *Many-Electron Approaches in Physics, Chemistry and Mathematics* (Springer: Mathematical Physics Studies, 2014).

2.2 Generalizing Hamiltonian complexity results to open quantum systems

Important methods and results concerning the time evolution of spin lattice systems are Lieb-Robinson [LR72] type bounds and also Trotter-Suzuki [Suz76, Tro59] type approximations [HDR90]. A Trotter-Suzuki approximation of unitary dynamics is the approximation of the time evolution operator by a composition of strictly local unitary operations. In the language of quantum information theory, this composition of local unitaries is a *quantum circuit* or, more generally, a *tensor network*, see Figure 2.1 for an example.

Lieb-Robinson bounds and Trotter-Suzuki type approximations allow for various schemes [DKSV04, Hae+11, Osb06, Vid04] to classically simulate real and imaginary time evolution of closed quantum systems efficiently in the system size. All known physical systems are, however, at lest weakly coupled to their environment (otherwise we wouldn't know about them). In the most simple case of Liouvillian dynamics, such as in many weak coupling situations, the interaction of the system with the environment is Markovian. More precisely, this means that the state of the system evolves according to a differential equation of motion, which is often called *master equation*. In order to represent proper quantum dynamics, the corresponding time evolution needs to be completely positive, or, equivalently [Lin76, WC08], the generator of such a continuous time evolution must have a certain normal form, called *Lindblad from* [Lin76]. Recently, such dynamics has been studied particularly intensely in the context of cold atoms in optical lattices [AOLL12, Die+08, DYDZ10], trapped ions [Bar+11, Sch+13], driven dissipative Rydberg gases [Gla+12], and macroscopic atomic ensembles [Kra+11]. Liouvillian dynamics also displays rich phenomena, that allow for dissipative state preparation [VWIC09],

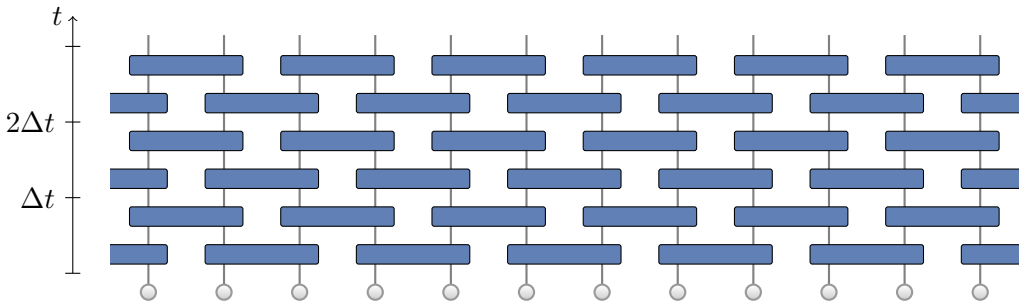

Figure 2.1: A quantum circuit originating from the Trotter-Suzuki approximation

$$\mathrm{e}^{-\mathrm{i}Ht} \approx \left(\prod\nolimits_j \mathrm{e}^{-\mathrm{i}h_{j,j+1}t/m}\right)^m$$

of the time evolution operator $\mathrm{e}^{-\mathrm{i}Ht}$, where $H = \sum_j h_{j,j+1}$ is a spin chain Hamiltonian with nearest-neighbour couplings $h_{j,j+1}$. Each box acting on sites j and $j+1$ corresponds to the strictly local time evolution operator $\mathrm{e}^{-\mathrm{i}\,h_{j,j+1}\,\Delta t}$. For such a spin chain with nearest neighbour interactions the circuit can be arranged in a brick layer structure.

dissipative phase transitions [Die+08], noise-driven criticality [EP10], non-equilibrium topological phase transitions [Bar+12], and precise timing of quantum operations [KWE13].

Considering the relevance of Liouvillian dynamics, it seems to be important to formulate Hamiltonian complexity results also for the open system setting. In the following, we first generalize Trotter-Suzuki type approximations to open quantum systems and discuss the physical implications. Then, we prove quasi-locality of Markovian time evolution, allowing for classical simulations of the local dynamics. This result provides two important building blocks for classical simulations of quantum systems. Moreover, the quasi-locality constitutes an important proof tool, e.g., for proving clustering of correlations in stationary states of rapidly mixing dynamics [KE13] and local topological quantum order [CLMPG13, KP13]. Combining quasi-locality with the result on rapid mixing [KT13], one can also obtain stability results such as the ones in Ref. [CLMPG13].

2.2.1 A dissipative Church-Turing theorem from generalizing the Trotter-Suzuki decomposition

This thesis is about complexity in simulations of quantum systems. Consequently, the first question we address is: What physical processes can one hope to efficiently simulate? The most powerful realistic situation that one can think of as a simulation, is the one of local dissipative quantum dynamics, as we will argue in the next paragraph. Such situations include the dissipative model of computation [VWIC09], which can efficiently simulate quantum computations, which, in turn, can efficiently simulate classical computations. The above question is also very much motivated by the *quantum complexity-theoretic Church-Turing thesis (QCCTT)* from computer science [KLM07]: "A quantum Turing machine [i.e. a quantum computer] can efficiently simulate any realistic model of computation". As such, the QCCTT is a vague statement that can neither be proven nor disproven. It serves as a working hypothesis in order to restrict to one specific model of computation, such as the unitary circuit model of quantum computation.

Here, we find reasonable assumptions under which we, nevertheless, can prove a Church-Turing type statement. First of all, we assume that there are only k-body interactions. This means that the Liouvillian can be decomposed into Liouvillians that only couple at most k subsystems for some constant k. Since we want to make a complexity theoretic statement, where one needs to be able to quantify the resources, we restrict to finitely many finite-dimensional subsystems. Next, we make the realistic assumption that the Liouvillian's time-dependence is piecewise continuous. Finally, we need to decide what a meaningful notion of "a system performing a computation" is. If a system would be allowed to exchange information with a potentially super powerful second system then we would hardly say that a computation is performed only by the resources of the system. This justifies that the system's time evolution

should be Markovian, i.e., described by a Liouvillian. In the following Publication [Kli+11a] we show that these assumptions are sufficient to simulate the dynamics efficiently in the unitary circuit model.

This result has various implications. First of all, it immediately follows that the model of dissipative quantum computation from Ref. [VWIC09] is equivalent to the unitary circuit model. Next, as we show on pages 40f, by lifting arguments from Ref. [PQSV11] to open quantum systems, the following limitation on efficient state generation follows: Starting in a product state, only exponentially few states can be approximately reached in polynomial time. Here, we fix some maximal approximation error $\epsilon > 0$ and count the states after also discretizing the state space with ϵ. Finally, in a similar way as Trotter-Suzuki [Suz76, Tro59] type approximations are used to simulate Hamiltonian dynamics, it is the hope that our dissipative Trotter formula proves useful in the open system setting. We will also elaborate on that further in Section 2.2.2, where we prove the *quasi-locality* of the dynamics for the case of finite ranged interactions.

On the next pages, we present the corresponding Publication [Kli+11a], which "opens up the potential impact of quantum computers to important applications in condensed-matter physics, quantum chemistry, and even biology" [Bro11].

Dissipative Quantum Church-Turing Theorem

M. Kliesch,[1,2] T. Barthel,[1,2] C. Gogolin,[1,2] M. Kastoryano,[3] and J. Eisert[1,2]

[1]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*

[2]*Institute for Physics and Astronomy, University of Potsdam, 14476 Potsdam, Germany*

[3]*Niels Bohr Institute, University of Copenhagen, 2100 Copenhagen, Denmark*

(Received 10 June 2011; published 12 September 2011)

We show that the time evolution of an open quantum system, described by a possibly time dependent Liouvillian, can be simulated by a unitary quantum circuit of a size scaling polynomially in the simulation time and the size of the system. An immediate consequence is that dissipative quantum computing is no more powerful than the unitary circuit model. Our result can be seen as a dissipative Church-Turing theorem, since it implies that under natural assumptions, such as weak coupling to an environment, the dynamics of an open quantum system can be simulated efficiently on a quantum computer. Formally, we introduce a Trotter decomposition for Liouvillian dynamics and give explicit error bounds. This constitutes a practical tool for numerical simulations, e.g., using matrix-product operators. We also demonstrate that most quantum states cannot be prepared efficiently.

DOI: 10.1103/PhysRevLett.107.120501 PACS numbers: 03.67.Ac, 02.60.Cb, 03.65.Yz, 89.70.Eg

One of the cornerstones of theoretical computer science is the Church-Turing thesis [1,2]. In its strong formulation it can be captured in the following way [3,4]: ''A probabilistic Turing machine can efficiently simulate any realistic model of computation.'' As such, it reduces any physical process—that can intuitively be thought of as a computational task in a wider sense—to what an elementary standard computer can do. Needless to say, in its strong formulation, the Church-Turing thesis is challenged by the very idea of a quantum computer, and hence by a fundamental physical theory that initially was thought to be irrelevant for studies of complexity. There are problems a quantum computer could efficiently solve that are believed to be intractable on any classical computer.

In this way, it seems that the strong Church-Turing thesis has to be replaced by a quantum version [2]. Colloquially speaking, the quantum Church-Turing thesis says that any process that can happen in nature that one could think of as being some sort of computation is efficiently simulatable.

Strong quantum Church-Turing thesis: Every quantum mechanical computational process can be simulated efficiently in the unitary circuit model of quantum computation.

Indeed, this notion of quantum computers being devices that can efficiently simulate natural quantum processes, being known under the name ''quantum simulation,'' is the topic of an entire research field initiated by the work of Feynman [5]. Steps towards a rigorous formulation have been taken by Lloyd [6] and many others [7].

Quite surprisingly, a very important class of physical processes appears to have been omitted in the quest for finding a sound theory of quantum simulation, namely, dissipative quantum processes. Such processes are particularly relevant since, in the end, every physical process is to some extent dissipative. If one aims at simulating a quantum process occurring in a lab, one cannot, however, reasonably require the inclusion of all modes of the environment to which the system is coupled into the simulation. Otherwise, one would always have to simulate all the modes of the environment, eventually of the entire Universe, rendering the task of simulation obsolete and futile. We argue that the most general setting in which one can hope for efficient simulatability is the one of Markovian dynamics [8] with arbitrary piecewise continuous time dependent control [9]. In any naturally occurring process the Liouvillian $\mathcal{L}$ determining the equation of motion

$$\frac{d}{dt}\rho(t) = \mathcal{L}_t(\rho(t)) \tag{1}$$

of the system state ρ is k-local. This means that the system is multipartite and $\mathcal{L}$ can be written as a sum of Liouvillians each acting nontrivially on at most k subsystems. In fact, all natural interactions are two-local in this sense. Since we are interested in processes which can be viewed as a computation, we assume that the subsystems are of fixed finite dimension. This is arguably the broadest class of natural physical processes that should be taken into account in a dissipative Church-Turing theorem and includes the Hamiltonian dynamics of closed systems as a special case.

In this work, we show the following.

(i) Every time evolution generated by a k-local time dependent Liouvillian can be simulated by a unitary quantum circuit with resources scaling polynomially in the system size N and simulation time τ.

(ii) As a corollary, we obtain that the dissipative model for quantum computing [11] can be reduced to the circuit model—proving a conjecture that was still open.

(iii) Technically, we show that the dynamics can be approximated by a Trotter decomposition, giving rise to a circuit of local channels, actually being reminiscent of the situation of unitary dynamics. In particular, in order to reach a final state that is only ϵ distinguishable from the exactly time evolved state, it will turn out to be sufficient to apply a circuit of Km local quantum channels, where

$$m = \left\lceil \max\left(\frac{2cK^2\tau^2}{\epsilon}, \frac{\tau b}{\ln 2}\right)\right\rceil \tag{2}$$

is the number of time steps, $K \leq N^k$ is the number of local terms in the Liouvillian, and b and c are constants independent of N, τ, K, and ϵ. Some obstacles of naive attempts to simulate dissipative dynamics are highlighted, and the specific role of the appropriate choice of norms is emphasized.

(iv) We also show that most quantum states cannot be prepared efficiently.

(v) In addition, the Trotter decomposition with our rigorous error bound is a practical tool for the numerical simulation of dissipative quantum dynamics on classical computers.

Setting.—We consider general quantum systems consisting of N subsystems of Hilbert-space dimension d. The dynamics is described by a quantum master equation (1) with a k-local Liouvillian of the form

$$\mathcal{L} = \sum_{\Lambda \subset [N]} \mathcal{L}_\Lambda, \tag{3}$$

where $[N] := \{1, 2, \ldots, N\}$ and $\mathcal{L}_\Lambda$ are *strictly* k-local Liouvillians. The subscript Λ means that the respective operator or superoperator acts nontrivially only on the subsystem Λ and we call an operator or superoperator *strictly* k-local if it acts nontrivially only on at most k subsystems. Each of the Liouvillians $\mathcal{L}_\Lambda$ can be written [10] in Lindblad form [12]

$$\mathcal{L}_\Lambda = -i[H_\Lambda, \cdot] + \sum_{\mu=1}^{d^k} \mathcal{D}[L_{\Lambda,\mu}], \tag{4}$$

where $\mathcal{D}[X](\rho) := 2X\rho X^\dagger - \{X^\dagger X, \rho\}$ and may depend on time piecewise continuously. In particular, we do not require any bound on the rate at which the Liouvillians may change.

The propagators $T_\mathcal{L}(t, s)$ are the family of superoperators defined by

$$\rho(t) = T_\mathcal{L}(t, s)(\rho(s)) \tag{5}$$

for all $t \geq s$. They are completely positive and trace preserving (CPT) and uniquely solve the initial value problem

$$\frac{d}{dt} T(t, s) = \mathcal{L}_t T(t, s), \qquad T(s, s) = \mathrm{id}, \tag{6}$$

where id denotes the identity map.

The main result, which is a bound on the error of the Trotter decomposition, will be somewhat reminiscent of the Trotter formula for time dependent Hamiltonian dynamics derived in Ref. [13]. The main challenge comes from the fact that we are dealing with superoperators rather than operators. The key to a meaningful Trotter decomposition is the choice of suitable norms for these superoperators. The physically motivated and strongest norm is the one arising from the operational distinguishability of two quantum states ρ and σ, which is given by the trace distance $\mathrm{dist}(\rho, \sigma) := \sup_{0 \leq A \leq \mathbb{1}} \mathrm{tr}[A(\rho - \sigma)]$. The trace distance coincides up to a factor of $1/2$ with the distance induced by the Schatten 1-norm $\|\cdot\|_1$, where the Schatten p-norm of a matrix A is $\|A\|_p := [\mathrm{tr}(|A|^p)]^{1/p}$. Therefore, we measure errors of approximations of superoperators with the induced operator norm, which is the so-called $(1 \to 1)$-norm. In general the $(p \to q)$-norm of a superoperator $T \in \mathcal{B}(\mathcal{B}(\mathcal{H}))$ is defined as [14]

$$\|T\|_{p\to q} := \sup_{\|A\|_p = 1} \|T(A)\|_q. \tag{7}$$

The difficulty in dealing with these norms lies in the fact that for $p < \infty$ the p-norm does not respect k-locality, e.g., $\|A \otimes \mathbb{1}_{n\times n}\|_1 = n\|A\|_1$. This problem is overcome by using the Lindblad form of the strictly k-local Liouvillians. In the end, all bounds can be stated in terms of the largest operator norm $\|X_t\|_\infty$ of the Lindblad operators $X \in \mathcal{L}_\Lambda$ of the strictly k-local terms. The notation $X \in \mathcal{L}_\Lambda$ means that X is one of the operators occurring in the Lindblad representation (4) of $\mathcal{L}_\Lambda$. From now on we assume that this largest operator norm a is everywhere bounded by a constant of order 1 and, in particular, independent of N, i.e., $a \in O(1)$.

Main result.—One can always approximate any dissipative dynamics generated by a k-local Liouvillian acting on N subsystems, even allowing for piecewise continuous time dependence, by a suitable Trotter decomposition. The error made in such a decomposition can be bounded rigorously.

Theorem 1 (Trotter decomposition of Liouvillian dynamics). Let $\mathcal{L} = \sum_{\Lambda\subset[N]} \mathcal{L}_\Lambda$ be a k-local Liouvillian that acts on N subsystems with local Hilbert-space dimension d. Furthermore, let $\mathcal{L}$ be piecewise continuous in time with the property that $a = \max_\Lambda \max_{X\in\mathcal{L}_\Lambda} \sup_{t\geq 0} \|X_t\|_\infty \in O(1)$. Then the error of the Trotter decomposition of a time evolution up to time τ into m time steps is

$$\left\| T_\mathcal{L}(\tau, 0) - \prod_{j=1}^{m} \prod_{\Lambda\subset[N]} T_{\mathcal{L}_\Lambda}\left(\tau\frac{j}{m}, \tau\frac{j-1}{m}\right)\right\|_{1\to 1} \leq \frac{cK^2\tau^2 e^{b\tau/m}}{m}, \tag{8}$$

where $c \in O(d^{2k})$, $b \in O(d^k)$, and $K \leq N^k$ is the number of strictly k-local terms $\mathcal{L}_\Lambda \neq 0$. $T_{\mathcal{L}_\Lambda}(\tau\frac{j}{m}, \tau\frac{j-1}{m})$ can be replaced by the propagator $T_{\mathcal{L}_\Lambda^{\mathrm{av}}}(\tau\frac{j}{m}, \tau\frac{j-1}{m}) = \exp(\tau\mathcal{L}_\Lambda^{\mathrm{av}}/m)$ of the average Liouvillian

$$\mathcal{L}_{\Lambda}^{\mathrm{av}} = \frac{m}{\tau}\int_{\tau(j-1)/m}^{\tau j/m} \mathcal{L}_{\Lambda}\, dt \tag{9}$$

without changing the scaling (8) of the error.

All constants are calculated explicitly in the Supplemental Material [15]. The supremum in a can be replaced by suitable time averages over the time steps such that $\|X_t\|_\infty$ can be large for small times. Before we turn to the proof of this result, we discuss important implications.

Implication 1 (Dissipative Church-Turing theorem). Time dependent Liouvillian dynamics can be simulated efficiently in the standard unitary circuit model.

Using the Stinespring dilation [16], each of the Km propagators $T_{\mathcal{L}_\Lambda}(\tau\frac{j}{m}, \tau\frac{j-1}{m})$ can be implemented as a unitary U_Λ^j acting on the subsystem Λ and an ancilla system of size at most d^{2k}. These unitaries can be decomposed further into circuits $\tilde{U}_\Lambda^j$ of at most $n = O(\log^\alpha(1/\epsilon_{\mathrm{SK}}))$ gates from a suitable gate set using the Solovay-Kitaev algorithm [17] with $\alpha < 4$ such that $\|U_\Lambda^j - \tilde{U}_\Lambda^j\|_\infty \leq \epsilon_{\mathrm{SK}}$. Note that for pure states, we have $\frac{1}{2}\|U|\psi\rangle\langle\psi|U^\dagger - \tilde{U}|\psi\rangle\times\langle\psi|\tilde{U}^\dagger\|_1 \leq \|U-\tilde{U}\|_\infty \leq \epsilon_{\mathrm{SK}}$ and the 1-norm is nonincreasing under partial trace. The full error is bounded by the error from the Trotter approximation (8) plus the one arising from the Solovay-Kitaev decomposition, in $(1\to 1)$-norm bounded by $Km\epsilon_{\mathrm{SK}}$.

At this point a remark on the appropriate degree of generality of the above result is in order. The proven result applies to dynamics under arbitrary piecewise continuous time dependent k-local Liouvillians. It does not include non-Markovian dynamics as often resulting from strong couplings. However, not only this result, but no dissipative Church-Turing theorem, can or should cover such a situation: Including highly non-Markovian dynamics would mean to also include extreme cases such as an evolution implementing a swap gate that could write the result of an incredibly complicated process happening in the huge environment into the system. In such an intertwined situation it makes only limited sense to speak of the time evolution of the system alone in the first place. On the other hand, in practical simulations of non-Markovian dynamics, where the influence of memory effects is known, pseudomodes can be included [18], thereby rendering the above results again applicable.

It has been shown recently [19] that the set of states that can be reached from a fixed pure reference state by k-local, time dependent Hamiltonian dynamics is exponentially smaller than the set of all pure quantum states. In fact, a more general statement holds true (see Supplemental Material [15]).

Implication 2 (Limitations of efficient state generation). Let $\mathcal{S}_\tau^\rho$ be the set of states resulting from the time evolution of an arbitrary initial state ρ under all possible (time dependent) k-local Liouvillians up to some time τ. For times τ that are polynomial in the system size, the relative volume of $\mathcal{S}_\tau^\rho$ (measured in the operational metric induced by the 1-norm) is exponentially small.

Finally, Theorem 1 also provides a rigorous error bound for the simulation of local time dependent Liouvillian dynamics on a classical computer. Even though classical simulation of quantum mechanical time evolution is generally believed to be hard in time, we have the following result.

Implication 3 (Simulation on classical computers). For fixed simulation time and efficiently evaluable initial states [20], dissipative dynamics can be simulated efficiently in the system size on classical computers, e.g., using a variant of time-dependent density matrix renormalization group.

This establishes a mathematically sound foundation for simulation techniques based on Trotter decomposition that have previously been used without proving that the approximation is actually possible; see, e.g., Ref. [21]. Recently, CPT maps like the local channels in the Trotter decomposition (8) have even been implemented in the lab [22].

Proof of theorem 1.—We now turn to the proof of the main result. First we will find $(1\to 1)$-norm estimates (i) for T and (ii) for T^- which will be used frequently. In the next step (iii) we derive a product formula, which we use iteratively (iv) to prove the Trotter decomposition. Finally, (v) we show how the second claim of the theorem concerning the approximation with the average Liouvillian can be proven. Throughout the proof we consider times $t \geq s \geq 0$.

(i) Because any CPT map T maps density matrices to density matrices, we have $\|T\|_{1\to 1} \geq 1$. In Ref. [14] it is shown that

$$\|T\|_{1\to 1} = \sup_{A=A^\dagger, \|A\|_1=1} \|T(A)\|_1 \tag{10}$$

for any CPT map T. Any self-adjoint operator $A = A_+ - A_-$ can, by virtue of its spectral decomposition, be written as the difference of a positive and negative part $A_\pm \geq 0$. Since T is CPT, $\|T(A_\pm)\|_1 = \mathrm{tr}(T(A_\pm)) = \|A\|_1$, hence $\|T\|_{1\to 1} \leq 1$, and finally $\|T\|_{1\to 1} = 1$.

(ii) For any Liouvillian $\mathcal{K}$ the propagator $T_{\mathcal{K}}(t,s)$ is invertible and the inverse $T_{\mathcal{K}}^-(t,s) = (T_{\mathcal{K}}(t,s))^{-1}$ is the unique solution of

$$\frac{d}{dt}T^-(t,s) = -T^-(t,s)\mathcal{K}_t, \qquad T^-(s,s) = \mathrm{id}. \tag{11}$$

From the representation of T^- as a reversely time-ordered exponential, the inequality

$$\|T_{\mathcal{K}}^-(t,s)\|_{1\to 1} \leq \exp\left(\int_s^t \|\mathcal{K}_r\|_{1\to 1}\, dr\right) \tag{12}$$

follows. This can be proved rigorously with the ideas from Ref. [23] (see Supplemental Material [15]).

For the case where $\mathcal{K}$ is strictly k-local, we use its Lindblad representation and the inequality $\|A\rho B\|_1 \leq \|A\|_\infty\|\rho\|_1\|B\|_\infty$ to establish $\|\mathcal{K}\|_{1\to 1} \in O(d^k)$ and hence $\|T_{\mathcal{K}}^-(t,s)\|_{1\to 1} \leq e^{b(t-s)}$, with $b \in O(d^k)$.

(iii) In the first step we use similar techniques as the ones being used for the unitary case [13] where differences of time evolution operators are bounded in operator norm by commutators of Hamiltonians. Applying the fundamental theorem of calculus twice, one can obtain for any two Liouvillians $\mathcal{K}$ and $\mathcal{L}$

$$\begin{aligned} T_{\mathcal{K}+\mathcal{L}}(t,s) &- T_{\mathcal{K}}(t,s)T_{\mathcal{L}}(t,s) \\ &= T_{\mathcal{K}}(t,s)T_{\mathcal{L}}(t,s)\int_s^t T_{\mathcal{L}}^-(r,s)\int_s^r \frac{d}{du}(T_{\mathcal{K}}^-(u,s) \\ &\quad \times \mathcal{L}_r T_{\mathcal{K}}(u,s))T_{\mathcal{K}}^-(r,s)T_{\mathcal{K}+\mathcal{L}}(r,s)dudr \\ &= \int_s^t\int_s^r T_{\mathcal{K}}(t,s)T_{\mathcal{L}}(t,r)T_{\mathcal{K}}^-(u,s) \\ &\quad \times [\mathcal{K}_u, \mathcal{L}_r]T_{\mathcal{K}}^-(r,u)T_{\mathcal{K}+\mathcal{L}}(r,s)dudr. \end{aligned} \tag{13}$$

In the next step we take the $(1 \to 1)$-norm of this equation, use the triangle inequality, employ submultiplicativity of the norm, and use (i) and (ii) to obtain $\int_s^t\int_s^r \|[\mathcal{K}_u, \mathcal{L}_r]\|_{1\to 1}dudr$ as an upper bound. In the case where $\mathcal{K}$ and $\mathcal{L}$ are strictly k-local $\|[\mathcal{K}_u, \mathcal{L}_r]\|_{1\to 1} \in O(d^{2k})$, which follows by the same arguments used in (ii) to bound $\|\mathcal{K}\|_{1\to 1}$. In the case where $\mathcal{L}$ is only k-local with K terms, $\|[\mathcal{K}_u, \mathcal{L}_r]\|_{1\to 1}$ is increased by at most the factor K such that

$$\|T_{\mathcal{K}+\mathcal{L}}(t,s) - T_{\mathcal{K}}(t,s)T_{\mathcal{L}}(t,s)\|_{1\to 1} \in O((t-s)^2 e^{b(t-s)}d^{2k}K). \tag{14}$$

(iv) The propagator can be written as

$$T_{\mathcal{L}}(\tau,0) = \prod_{j=1}^{m} T_{\mathcal{L}}(\tau j/m, \tau(j-1)/m). \tag{15}$$

Using the inequality

$$\|T_1T_2 - \tilde{T}_1\tilde{T}_2\| \le \|T_1\|\|T_2 - \tilde{T}_2\| + \|T_1 - \tilde{T}_1\|\|\tilde{T}_2\| \tag{16}$$

and Eq. (14) iteratively, one can establish the result as stated in Eq. (8).

(v) For any strictly k-local Liouvillian $\mathcal{K}$ the propagator $T_{\mathcal{K}}(t,s)$ can be approximated by the propagator of the average Liouvillian,

$$\left\| T_{\mathcal{K}}(t,s) - \exp\left(\int_s^t \mathcal{K}_r dr\right)\right\|_{1\to 1} = \frac{1}{3}b(t-s)^2. \tag{17}$$

This can be shown using the techniques described above by lifting the proof from Ref. [19] to the dissipative case (see Supplemental Material [15]). A comparison of Eq. (17) with Eq. (14) shows that the error introduced by using the average Liouvillian is small compared to the error introduced by the product decomposition and does not change the scaling of the error.

Conclusion.—In this work we show that under reasonable assumptions the dynamics of open quantum systems can be simulated efficiently by a circuit of local quantum channels in a Trotter-like decomposition. This channel circuit can further be simulated by a unitary quantum circuit with polynomially many gates from an arbitrary universal gate set. As a corollary it follows that the dissipative model of quantum computation is no more powerful than the standard unitary circuit model. Furthermore, the result implies that k-local Liouvillian dynamics can be simulated efficiently in the system size on a classical computer. It also shows that systems considered in the context of dissipative phase transitions [11,24] can be simulated in both of the above senses. The result can be seen as a quantum Church-Turing theorem in the sense that under reasonable and necessary requirements any general time evolution of an open quantum system can be simulated efficiently on a quantum computer.

This work was supported by the EU (Qessence, Minos, COQUIT, Compas), the BMBF (QuOReP), the EURYI, the Niels Bohr International Academy, the German National Academic Foundation, and the Perimeter Institute. We would like to thank M. P. Müller and T. Prosen for discussions.

[1] A. M. Turing, Proc. London Math. Soc. **42**, 230 (1937); A. Church, Ann. Math. **33**, 346 (1932).
[2] E. Bernstein, and U. Vazirani, SIAM J. Comput. **26**, 1411 (1997).
[3] P. Kaye, R. Laflamme, and M. Mosca, *An Introduction to Quantum Computing* (Oxford University Press, Oxford, 2007).
[4] *The Undecidable, Basic Papers on Undecidable Propositions, Unsolvable Problems And Computable Functions*, edited by M. Davis (Raven Press, New York, 1965).
[5] R. Feynman, Int. J. Theor. Phys. **21**, 467 (1982).
[6] S. Lloyd, Science **273**, 1073 (1996).
[7] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, SIAM J. Comput. **37**, 166 (2007); D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders, Commun. Math. Phys. **270**, 359 (2007).
[8] T. S. Cubitt, J. Eisert, and M. M. Wolf, arXiv:0908.2128.
[9] For finite dimensional systems this means mathematically that $T(t,s)$ defined in Eq. (5) satisfies (i) $T(t,r)T(r,s) = T(t,s)\ \forall\ t \ge r \ge s \ge 0$ and (ii) $\lim_{\epsilon\to 0}\|T(t+\epsilon,t) - \mathrm{id}\| = 0\ \forall\ t \ge 0$. Then T is generated by a time dependent Liouvillian $\mathcal{L}_t$ [10].
[10] M. M. Wolf and J. I. Cirac, Commun. Math. Phys. **279**, 147 (2008).
[11] F. Verstraete, M. M. Wolf, and J. I. Cirac, Nature Phys. **5**, 633 (2009).
[12] G. Lindblad, Commun. Math. Phys. **48**, 119 (1976).
[13] J. Huyghebaert and H. De Raedt, J. Phys. A **23**, 5777 (1990).
[14] J. Watrous, Quantum Inf. Comput. **5**, 58 (2005).
[15] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.107.120501 for further details on the presented proofs and details on Implication 2.

[16] W. F. Stinespring, Proc. Am. Math. Soc. **6**, 211 (1955); V. Paulsen, *Completely Bounded Maps and Operator Algebras* (Cambridge University Press, Cambridge, England, 2002).

[17] C. M. Dawson and M. A. Nielsen, Quantum Inf. Comput. **6**, 81 (2006); A. Yu. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation* (American Mathematical Society, Providence, RI, 2002), p. 257.

[18] A. Imamoglu, Phys. Rev. A **50**, 3650 (1994); H. P. Breuer, Phys. Rev. A **70**, 012106 (2004).

[19] D. Poulin, A. Qarry, R. D. Somma, and F. Verstraete, Phys. Rev. Lett. **106**, 170501 (2011).

[20] We call states efficiently evaluable if they allow for an efficient evaluation of the expectation values of all strictly k'-local observables for a given $k' \in O(1)$. The stated implication follows from the fact that the application of Trotter circuits to such observables yields again strictly k''-local observables with $k'' \in O(1)$ due to a light cone effect.

[21] L. L. Halcomb and D. J. Diestler, J. Chem. Phys. **121**, 3393 (2004); R. Kapral, Annu. Rev. Phys. Chem. **57**, 129 (2006); D. M. Kernan, G. Ciccotti, and R. Kapral, J. Phys. Chem. B **112**, 424 (2008); G. Benenti, G. Casati, T. Prosen, D. Rossini, and M. Znidaric, Phys. Rev. B **80**, 035110 (2009).

[22] J. Barreiro *et al.*, Nature (London) **470**, 486 (2011).

[23] J. D. Dollard and C. N. Friedman, J. Math. Phys. (N.Y.) **18**, 1598 (1977).

[24] S. Diehl, A. Micheli, A. Kantian, B. Kraus, H. P. Büchler, and P. Zoller, Nature Phys. **4**, 878 (2008); J. Eisert and T. Prosen, arXiv:1012.5013.

Supplementary Material
(A dissipative quantum Church-Turing theorem)

M. Kliesch,[1,2] T. Barthel,[1,2] C. Gogolin,[1,2] M. Kastoryano,[3] and J. Eisert[1,2]

[1]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*
[2]*Institute for Physics and Astronomy, University of Potsdam, 14476 Potsdam, Germany*
[3]*Niels Bohr Institute, University of Copenhagen, 2100 Copenhagen, Denmark*

APPENDIX

In this appendix we elaborate on some of the technical aspects of our results and give explicit expressions for all involved constants. First, we give a detailed derivation of the error caused by the Trotter approximation for the time evolution under a time dependent k-local Liouvillian. Along the way we also derive a completely general bound for the Trotter error for arbitrary (not necessarily k-local) time dependent Liouvillians, which we don't need directly for the statements made in the paper, but which could be of interest independently as the bounds of our more specialized theorem is not optimal in certain situations. Secondly, we present a detailed derivation of the error that is made when the time evolution under the time dependent Liouvillian is replaced by that of the average Liouvillian on a small time step. Finally, we prove results on the scaling behavior of ϵ-nets used in the end of the paper to argue that only an exponentially small subset of states can be prepared with time dependent k-local Liouvillian dynamics in polynomial time from a fixed reference state. Our argument lifts the considerations from Ref. [1] to the space of density matrices and the physically relevant trace distance.

TROTTER APPROXIMATION FOR TIME DEPENDENT LIOUVILLIANS

We start by giving a detailed proof that for short time intervals it is possible to approximate the time evolution of a k-local time dependent Liouvillian $\mathcal{K}+\mathcal{L}$ by splitting off a strictly k-local part $\mathcal{K}$ and performing the time evolution under $\mathcal{L}$ and $\mathcal{K}$ sequentially.

Theorem 2 (Product decomposition of propagators). *Let $\mathcal{L}$ and $\mathcal{K}$ be two time dependent Liouvillians that act on the same quantum system of N subsystems with local Hilbert space dimension d. Furthermore, let $\mathcal{K}$ be strictly k-local and let $\mathcal{L}$ be k-local consisting of K strictly k-local terms $\mathcal{L}_\Lambda$. For $t \geq s$ the Trotter error is given by*

$$\left\|T_{\mathcal{K}+\mathcal{L}}(t,s) - T_{\mathcal{K}}(t,s)T_{\mathcal{L}}(t,s)\right\|_{1\to 1} \leq (t-s)^2 \mathrm{e}^{b(t-s)} c\, K\,, \tag{1}$$

where

$$\begin{aligned} b &= 2a^2(2+4d^k),\\ c &= 2a^2 + 8a^3 d^k + 16a^4 d^{2k},\\ a &= \max_\Lambda \max_{X\in\mathcal{K}\cup\mathcal{L}_\Lambda} \sup_{s\leq v\leq t} \|X_v\|_\infty\,. \end{aligned}$$

We will use this theorem iteratively to bound the error caused by decomposing the propagators of arbitrary k-local Liouvillians into the propagators of the individual strictly k-local terms.

The proof of this theorem can be presented most conveniently as a series of Lemmas. From the main text (point (i) in the proof of the Theorem, page 3) we already know that completely positive and trace preserving (CPT) maps are contractive:

Lemma 3 (Contraction property of the propagator)**.** *Let T be a CPT map. Then $\|T\|_{1\to 1} = 1$.*

We also need to bound the inverse propagator.

Lemma 4 (Backward time evolution)**.** *For $t \geq s$:*

(i) $T_{\mathcal{L}}(t,s)$ is invertible and its inverse is $T^-_{\mathcal{L}}(t,s)$ as defined by Eq. (11) in the main text.

(ii) If the Liouvillian $\mathcal{L}$ is piecewise continuous in time then

$$\left\|T^-_{\mathcal{L}}(t,s)\right\|_{1\to 1} \leq \exp\Big(\int_s^t \|\mathcal{L}_r\|_{1\to 1}\,\mathrm{d}r\Big)\,. \tag{2}$$

Proof. First, we consider the case where $\mathcal{L}$ is continuous in time and use the theory presented in Ref. [2] and in particular the "properties" which are proven in this reference. The *product integral* of $\mathcal{L}$ is defined analogously to the Riemann integral,

$$\prod_s^t \exp(\mathcal{L}_r\,\mathrm{d}r) \coloneqq \lim_{\Delta r_j\to 0\,\forall j} \prod_{j=1}^J \exp(\mathcal{L}_{r_j}\Delta r_j)\,, \tag{3}$$

where $\prod_{j=1}^J X_j \coloneqq X_J X_{J-1}\ldots X_1$. Since $T_{\mathcal{L}}(t,s)$ solves the initial value problem in Eq. (6) from the main text, $T_{\mathcal{L}}(t,s) = \prod_s^t \exp(\mathcal{L}_r\,\mathrm{d}r)$ which is exactly the statement of property 1.
(i) Property 3 precisely states that a product integral is invertible. It is not hard to see that the inverse of $T_{\mathcal{L}}(t,s)$ solves the initial value problem (11) from the main text.
(ii) The inverse propagator is

$$T^-_{\mathcal{L}}(t,s) = \left(\prod_s^t \exp(\mathcal{L}_r\,\mathrm{d}r)\right)^{-1}\,. \tag{4}$$

Since matrix inversion is continuous,

$$T^-_{\mathcal{L}}(t,s) = \lim_{\Delta r_j\to 0\,\forall j} \prod_{j=J}^1 \exp(-\mathcal{L}_{r_j}\Delta r_j). \tag{5}$$

We call this the *reversely ordered product integral* and use the convention $\prod_{j=J}^{1} X_j \coloneqq X_1 X_2 \dots X_J$. Using the submultiplicativity of the $(1 \to 1)$-norm and the triangle inequality we obtain from Eq. (5)

$$\left\|T_{\mathcal{L}}^{-}(t,s)\right\|_{1\to1} \leq \lim_{\Delta r_j \to 0\, \forall j} \prod_{j=J}^{1} \exp(\left\|\mathcal{L}_{r_j}\right\|_{1\to1} \Delta r_j) \quad (6)$$

$$= \exp(\lim_{\Delta r_j \to 0\, \forall j} \sum_{j=1}^{J} \left\|\mathcal{L}_{r_j}\right\|_{1\to1} \Delta r_j) \quad (7)$$

The definition of the Riemann integral finishes the proof for the continuous case.

If $\mathcal{L}$ is only piecewise continuous in time then (i) and (ii) hold for all the intervals where $\mathcal{L}$ is continuous and from that and the composition property $T_{\mathcal{L}}(u,v)T_{\mathcal{L}}(v,w) = T_{\mathcal{L}}(u,w)$ $(u \geq v \geq w)$ it follows that (i) and (ii) hold on the whole time interval $[s,t]$. □

With these tools at hand we can now prove a bound on the Trotter error of two arbitrary (not necessarily k-local) time dependent Liouvillians.

Theorem 5 (General Trotter error). *For two arbitrary time dependent Liouvillians $\mathcal{K}$ and $\mathcal{L}$ the Trotter error is given by*

$$\|T_{\mathcal{K}+\mathcal{L}}(t,s) - T_{\mathcal{K}}(t,s)T_{\mathcal{L}}(t,s)\|_{1\to1} \leq \int_s^t \int_s^r \|[\mathcal{K}_u, \mathcal{L}_r]\|_{1\to1}\, \mathrm{d}u\, \mathrm{d}r\, \mathrm{e}^{2\int_s^t \|\mathcal{K}_v\|_{1\to1}\, \mathrm{d}v} \quad (8)$$

$$\leq \tfrac{1}{2}(t-s)^2 \sup_{t\geq r\geq u\geq s} \|[\mathcal{K}_u, \mathcal{L}_r]\|_{1\to1}\, \mathrm{e}^{2(t-s)\sup_{t\geq v\geq s}\|\mathcal{K}_v\|_{1\to1}}\,. \quad (9)$$

Proof. We use a similar argument as in Ref. [3]. With the fundamental theorem of calculus we obtain

$$\begin{aligned}
T_{\mathcal{L}}^{-}(t,s)T_{\mathcal{K}}^{-}(t,s)T_{\mathcal{K}+\mathcal{L}}(t,s) - \mathrm{id} &= \int_s^t \partial_r \big(T_{\mathcal{L}}^{-}(r,s)T_{\mathcal{K}}^{-}(r,s)T_{\mathcal{K}+\mathcal{L}}(r,s)\big)\, \mathrm{d}r \\
&= \int_s^t T_{\mathcal{L}}^{-}(r,s)[T_{\mathcal{K}}^{-}(r,s), \mathcal{L}_r]T_{\mathcal{K}+\mathcal{L}}(r,s)\, \mathrm{d}r \\
&= \int_s^t T_{\mathcal{L}}^{-}(r,s)\big(T_{\mathcal{K}}^{-}(r,s)\mathcal{L}_r T_{\mathcal{K}}(r,s) - \mathcal{L}_r\big)T_{\mathcal{K}}^{-}(r,s)T_{\mathcal{K}+\mathcal{L}}(r,s)\, \mathrm{d}r \\
&= \int_s^t T_{\mathcal{L}}^{-}(r,s) \int_s^r \frac{\mathrm{d}}{\mathrm{d}u}\Big(T_{\mathcal{K}}^{-}(u,s)\mathcal{L}_r T_{\mathcal{K}}(u,s)\Big)\mathrm{d}u\, T_{\mathcal{K}}^{-}(r,s)T_{\mathcal{K}+\mathcal{L}}(r,s)\, \mathrm{d}r \\
&= \int_s^t \int_s^r T_{\mathcal{L}}^{-}(r,s)T_{\mathcal{K}}^{-}(u,s)[\mathcal{L}_r, \mathcal{K}_u]T_{\mathcal{K}}(u,s)T_{\mathcal{K}}^{-}(r,s)T_{\mathcal{K}+\mathcal{L}}(r,s)\, \mathrm{d}u\, \mathrm{d}r\,.
\end{aligned}$$

Multiplying with $T_{\mathcal{K}}(t,s)T_{\mathcal{L}}(t,s)$ from the left yields

$$T_{\mathcal{K}+\mathcal{L}}(t,s) - T_{\mathcal{L}}(t,s)T_{\mathcal{K}}(t,s) = \int_s^t \int_s^r T_{\mathcal{K}}(t,s)T_{\mathcal{L}}(t,r)T_{\mathcal{K}}^{-}(u,s)[\mathcal{L}_r, \mathcal{K}_u]T_{\mathcal{K}}^{-}(r,u)T_{\mathcal{K}+\mathcal{L}}(r,s)\, \mathrm{d}u\, \mathrm{d}r. \quad (10)$$

With submultiplicativity of the $(1 \to 1)$-norm and the bounds on the norms of the forward and backward propagators from lemma 3 and 4 the result follows. □

To complete the proof of theorem 2 one needs to bound the norms $\|[\mathcal{L}_r, \mathcal{K}_u]\|_{1\to1}$ and $\|\mathcal{K}_r\|_{1\to1}$ in (9) for the special case that $\mathcal{K}$ is strictly k-local and $\mathcal{L}$ is k-local with K strictly k-local terms.

Lemma 6. *Let $\mathcal{K}$ and $\mathcal{L}$ be two Liouvillians which act on the same operator space of N subsystems with local Hilbert space dimension d. Furthermore, let $\mathcal{K}$ be strictly k-local and $\mathcal{L}$ be k-local consisting of K strictly k-local terms $\mathcal{L}_\Lambda$. Then*

$$2\,\|\mathcal{K}_v\|_{1\to1} \leq b_v \quad (11)$$

and

$$\tfrac{1}{2}\,\|[\mathcal{L}_r, \mathcal{K}_u]\|_{1\to1} \leq c_{r,u}K\,, \quad (12)$$

where $b_v = 4a_v + 8d^k a_v^2$, $c_{r,u} = 2a_r a_u + 4(a_r a_u^2 + a_r^2 a_u)d^k + 16 a_r^2 a_u^2 d^{2k}$, and $a_t = \max_\Lambda \max\{\|X_t\|_\infty : X \in \mathcal{K} \cup \mathcal{L}_\Lambda\}$.

Proof. First, let both Liouvillians be strictly k-local. Hence each of them can be written with at most d^k Lindblad operators. Let the Lindblad representations of $\mathcal{K}$ and $\mathcal{L}$ be

$$\mathcal{K} = -\mathrm{i}[G, \cdot] + \sum_{\nu=1}^{d^k} \mathcal{D}[K_\nu] \quad (13)$$

and

$$\mathcal{L} = -\mathrm{i}[H,\cdot] + \sum_{\mu=1}^{d^k} \mathcal{D}[L_\mu]\,, \tag{14}$$

where $\mathcal{D}[X](\rho) \coloneqq 2X\rho X^\dagger - \{X^\dagger X, \rho\}$. Inequality (11) follows from counting the number of terms in (13) and using that $\|A\rho B\|_1 \leq \|A\|_\infty \|\rho\|_1 \|B\|_\infty$. Similarly, by writing out the commutator $[\mathcal{K},\mathcal{L}]$ and using the above representations one can verify that $[\mathcal{K}_r,\mathcal{L}_u] \leq 2a_r a_u + 4(a_r a_u^2 + a_r^2 a_u)d^k + 16a_r^2 a_u^2 d^{2k}$. If $\mathcal{L} = \sum_{\Lambda\subset[N]} \mathcal{L}_\Lambda$ is k-local with K terms the bound is increased by at most a factor of K. □

Theorem 2 follows as a corollary of theorem 5 and lemma 6 by inserting the suprema of the bounds (11) and (12) into Eq. (9). Instead of using suprema in the step from Eq. (8) to Eq. (9) one can take averages over b_v and $c_{r,u}$ to obtain a better, but more complicated bound. One can also improve the scaling of the error with the size of the time steps by using higher order Trotter schemes as in Ref. [3] (time dependent case) or Ref. [4] (time constant case).

APPROXIMATION BY THE AVERAGE LIOUVILLIAN

In the product formula in our theorem 1 in the main text one can replace the time ordered integrals $T_{\mathcal{L}_\Lambda}(t,s)$ by ordinary exponentials of the time averaged Liouvillians. This is not essential to our argument concerning the quantum Church-Turing thesis, but makes the result more useful for applications. The additional error caused by doing this is bounded in the following theorem:

Theorem 7 (Approximation by the average Liouvillian). *Let $\mathcal{K}$ be a strictly k-local Liouvillian acting on an operator space with local Hilbert space dimension d. Then for any $t \geq s$*

$$\|T_\mathcal{K}(t,s) - \exp((t-s)\mathcal{K}^{\mathrm{av}})\|_{1\to1} \leq \tfrac{1}{3}b(t-s)^2\,, \tag{15}$$

where the average Liouvillian

$$\mathcal{K}^{\mathrm{av}} \coloneqq \frac{1}{t-s}\int_s^t \mathcal{K}_r\,\mathrm{d}r \tag{16}$$

is indeed a Liouvillian, $b = 2a^2(2+4d^k)$, and $a = \max_{X_t\in\mathcal{K}} \sup_t \|X_t\|_\infty$.

Proof. We lift the proof from Ref. [1] to the dissipative setting. Let $t \geq s$ be fixed. Applying the fundamental theorem of calculus and the definition of $\mathcal{K}^{\mathrm{av}}$, we obtain

$$\begin{aligned}
T_{\mathcal{K}^{\mathrm{av}}}(t,s) - T_\mathcal{K}(t,s) &= -T_\mathcal{K}(t,s)\int_s^t T_\mathcal{K}^-(u,s)\,(\mathcal{K}_u - \mathcal{K}^{\mathrm{av}})\,T_{\mathcal{K}^{\mathrm{av}}}(u,s)\,\mathrm{d}u \\
&= -\frac{1}{t-s}\int_s^t\int_s^t T_\mathcal{K}(t,u)\,(\mathcal{K}_u - \mathcal{K}_r)\,T_{\mathcal{K}^{\mathrm{av}}}(u,s)\,\mathrm{d}r\,\mathrm{d}u \\
&= -\frac{1}{t-s}\int_s^t\int_s^t \Big(T_\mathcal{K}(t,u)\mathcal{K}_u T_{\mathcal{K}^{\mathrm{av}}}(u,s) - T_\mathcal{K}(t,r)\mathcal{K}_u T_{\mathcal{K}^{\mathrm{av}}}(r,s)\Big)\mathrm{d}r\,\mathrm{d}u\,.
\end{aligned}$$

The inequality in Eq. (16) from the main text yields

$$\begin{aligned}
\|T_{\mathcal{K}^{\mathrm{av}}}(t,s) - T_\mathcal{K}(t,s)\|_{1\to1} \leq \frac{1}{t-s}\int_s^t\int_s^t \|\mathcal{K}_u\|_{1\to1}\Big(&\|T_\mathcal{K}(t,u) - T_\mathcal{K}(t,r)\|_{1\to1}\,\|T_{\mathcal{K}^{\mathrm{av}}}(u,s)\|_{1\to1} \\
&+ \|T_\mathcal{K}(t,r)\|_{1\to1}\,\|T_{\mathcal{K}^{\mathrm{av}}}(u,s) - T_{\mathcal{K}^{\mathrm{av}}}(r,s)\|_{1\to1}\Big)\mathrm{d}r\,\mathrm{d}u\,.
\end{aligned} \tag{17}$$

From $T_\mathcal{K}(u,s) - T_\mathcal{K}(r,s) = -\int_r^u T_\mathcal{K}(v,s)\mathcal{K}_v\,\mathrm{d}v$, lemma 3 and the submultiplicativity of the norm we know that for $t \geq u,r \geq s$

$$\|T_\mathcal{K}(u,s) - T_\mathcal{K}(r,s)\|_{1\to1} \leq \left|\int_u^r \|\mathcal{K}_v\|_{1\to1}\,\mathrm{d}v\right| \tag{18}$$

and similarly for $\mathcal{K}^{\mathrm{av}}$. With (17) we obtain

$$\begin{aligned}
&\|T_{\mathcal{K}^{\mathrm{av}}}(t,s) - T_\mathcal{K}(t,s)\|_{1\to1} \\
&\qquad\leq 2\int_s^t\int_s^t \left|\int_u^r \|\mathcal{K}_v\|_{1\to1}\,\mathrm{d}v\right|\mathrm{d}r\,\mathrm{d}u\,.
\end{aligned} \tag{19}$$

It remains to show that $\mathcal{K}^{\mathrm{av}}$ is a Liouvillian, i.e., that $\exp(t\mathcal{K}^{\mathrm{av}})$ is a CPT map for all $t \geq 0$. First of all, finite sums of Liouvillians are Liouvillians. Furthermore, limits of se-

quences of Liouvillians are Liouvillians since the exponential function is continuous and the set of CPT maps is closed. □

EFFICIENTLY PREPARABLE STATES CONSTITUTE AN EXPONENTIALLY SMALL SUBSET OF STATE SPACE

In the following we will argue that for every fixed initial state, the time evolution for a time interval of length τ under any (possibly time dependent) k-local Liouvillian yields a state that lies inside of one of N_T ϵ-balls in trace distance. For times τ which are polynomial in the system size, N_T is exponentially smaller than the cardinality of any ϵ-net (in trace distance) that covers the state space $\mathcal{S}$. The case of Hamiltonian dynamics and state vectors is investigated in Ref. [1]. It will be convenient to use the Bachmann-Landau symbols O and Ω for asymptotic upper and lower bounds up to constant factors.

By using theorem 1 of the main text, which provides an error bound for the Trotter approximation of a Liouvillian time evolution, together with the Stinespring dilation [5] and the Solovay-Kitaev algorithm [6], one obtains the following:

Theorem 8 (Number of channel circuits). *The propagator from time* 0 *to time* τ*, generated by any* k*-local time dependent Liouvillian acting on* N *subsystems with local Hilbert space dimension* $d \in \mathrm{O}(1)$ *can be approximated in* $(1 \to 1)$*-norm to accuracy* $\epsilon > 0$ *with one out of* N_T *channel circuits, where*

$$\log(N_T) \in \mathrm{O}\left(\frac{N^{3k+2}\tau^4}{\epsilon^5}\right) . \tag{20}$$

Proof. According to theorem 1 of the main text, the propagator $T_{\mathcal{L}}(\tau,0)$ of the Liouvillian time evolution can be approximated by a circuit $\prod_{j=1}^{m}\prod_{\Lambda\subset[N]} T_\Lambda^j$ of at most $N^k m$ strictly k-local channels T_Λ^j to precision ϵ_1 in $(1\to 1)$-norm, where according to Eq. (2) from the main text, $m = 2cN^{2k}\tau^2/\epsilon_1$. We have assumed that $2\ln(2)cN^{2k}\tau/\epsilon_1 \geq b$ where c and b are given explicitly in theorem 2 and depend only on strictly local properties of the Liouvillian. Employing the Stinespring dilation [5] for each of the channels T_Λ^j one obtains a circuit of at most $N^k m$ strictly $3k$-local unitary gates U_Λ^j. Each U_Λ^j acts on an enlarged system composed of the d^k-dimensional original subsystem and an ancilla system of dimension d^{2k}. One can use the Solovay-Kitaev algorithm [6] to approximate every single gate U_Λ^j of the unitary circuit by a circuit $\tilde{U}_\Lambda^j$ of one- and two-qubit gates from a universal gate set of cardinality $n_{\mathrm{SK}} \in \mathrm{O}(1)$, e.g., $n_{\mathrm{SK}} = 3$. With $N_{\mathrm{SK}} = c_{\mathrm{SK}}\log^\alpha(1/\epsilon_{\mathrm{SK}})$ of those n_{SK} standard gates, each unitary U_Λ^j can be approximated to accuracy ϵ_{SK} introducing a total error $\epsilon_2 = N^k m\epsilon_{\mathrm{SK}}$. The constant c_{SK} depends on d^{3k}.

Consequently, we have for the dilation U of $\prod_{j=1}^{m}\prod_{\Lambda\subset[N]} T_\Lambda^j$ an approximation $\tilde{U}$ with operator norm accuracy ϵ_2, given by a unitary circuit of $N_{\text{All gates}} = N_{\mathrm{SK}}N^k m$ standard gates from the universal gate set. Note that for any pure state $|\psi\rangle$, we have $\frac{1}{2}\|U\,|\psi\rangle\langle\psi|\,U^\dagger, \tilde{U}\,|\psi\rangle\langle\psi|\,\tilde{U}^\dagger\|_1 \leq \|U-\tilde{U}\|_\infty$ and the 1-norm is non-increasing under partial trace. Tracing out the ancillas, we obtain an approximation $\tilde{T}$ of $T_{\mathcal{L}}(\tau,0)$ with error $\|T_{\mathcal{L}}(\tau,0)-\tilde{T}\|_{1\to 1} \leq \epsilon = \epsilon_1 + 2\epsilon_2$. The total number of different channels $\tilde{T}$, which can arise in this way from the chosen universal gate set, is $N_T \leq {n_{\mathrm{SK}}}^{N_{\text{All gates}}}$, i.e., for given c,τ,k,N and d, a number of N_T standard gates are enough to approximate any $T_{\mathcal{L}}(\tau,0)$ in $(1\to 1)$-norm to accuracy ϵ.

To conclude, we bound the order of N_T.

$$\begin{aligned}\log(N_T) &\leq N_{\text{All gates}}\log n_{\mathrm{SK}}\\ &= c_{\mathrm{SK}}\log^\alpha\left(\frac{2cN^{3k}\tau^2}{\epsilon_1\epsilon_2}\right)\frac{2cN^{3k}\tau^2}{\epsilon_1}\log n_{\mathrm{SK}}\\ &< c_{\mathrm{SK}}(3k)^\alpha\log^\alpha\left(\frac{2cN\tau}{\epsilon_1\epsilon_2}\right)\frac{2cN^{3k}\tau^2}{\epsilon_1}\log n_{\mathrm{SK}}\,.\end{aligned} \tag{21}$$

Since we are interested in the scaling of $\log(N_T)$ for large N and small ϵ_1,ϵ_2 we can assume that the argument of the logarithm is larger than 18 and use that $\log_2^4(x) < x^2$ for $x \geq 18$ to obtain

$$\log(N_T) < C\frac{N^{3k+2}\tau^4}{\epsilon_1^3\epsilon_2^2} \tag{22}$$

with $C = c_{\mathrm{SK}}(3k)^\alpha(2c)^3\log n_{\mathrm{SK}}$. □

The above theorem shows that the time evolution under a k-local Liouvillian can be approximated by one out of N_T many circuits to accuracy ϵ. The states that can be reached by any k-local Liouvillian time evolution, starting from a fixed initial state, are hence all contained in the union of N_T ϵ-balls (in 1-norm) around the output states of these circuits.

Let us now determine whether those ϵ-balls can possibly cover the whole state space. For this purpose we introduce ϵ-nets. We consider a D-dimensional Hilbert space $\mathcal{H}$ and denote

(i) the set of state vectors, i.e., the set of normalized vectors in $\mathcal{H}$ by $P \subset \mathcal{H}$,
(ii) the set of density matrices by $\mathcal{S} \subset \mathcal{B}(\mathcal{H})$, and
(iii) the set of rank one projectors by $\mathcal{P} \subset \mathcal{S}$.

For an arbitrary subset $R \subset \mathcal{B}(\mathcal{H})$ and some $\epsilon > 0$ we call a finite subset $\mathcal{N}_\epsilon^p(R) \subset R$ satisfying

$$\forall a \in R\ \exists b \in \mathcal{N}_\epsilon^p(R) : \|a-b\|_p \leq \epsilon \tag{23}$$

an ϵ*-net for* R *in (Schatten)* p*-norm*. Furthermore, we call an ϵ-net $\hat{\mathcal{N}}_\epsilon^p(R)$ *optimal* if any other set $X \subset R$ with smaller cardinality $|X| < |\hat{\mathcal{N}}_\epsilon^p(R)|$ cannot be an ϵ-net for R in p-norm. Similarly, we define ϵ-nets $\mathcal{N}_\epsilon^{\mathrm{HS}}(P) \subset P$ for state vectors in Hilbert space norm and, as before, we denote optimal ϵ-nets by $\hat{\mathcal{N}}_\epsilon^{\mathrm{HS}}(P)$.

In Ref. [7] it was shown that for the set of state vectors of a D-dimensional quantum system there exist ϵ-nets of cardinality at most $|\mathcal{N}_\epsilon^{\mathrm{HS}}(P)| \leq (5/(2\epsilon))^{2D}$. As the Hilbert space distance upper bounds [8] the trace distance,

$$\begin{aligned}|\,|\psi\rangle - |\phi\rangle\,|_2 &\geq \tfrac{1}{2}\,\|\,|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|\,\|_1\\ &= \mathrm{dist}(\,|\psi\rangle\langle\psi|\,,\,|\phi\rangle\langle\phi|\,)\,,\end{aligned} \tag{24}$$

this also implies the existence of ϵ-nets for $\mathcal{P}$ in p-norm of cardinality $|\mathcal{N}_\epsilon^p(\mathcal{P})| \leq (5/\epsilon)^{2D}$ for any $p \geq 1$. By comparing the volume of the ϵ-balls with the volume of the whole set of state vectors one can see that for state vectors this construction is essentially optimal.

Lemma 9. *For a D-dimensional quantum system*

$$|\mathcal{N}_\epsilon^{\mathrm{HS}}(P)| \in \Omega(\left(\frac{1}{\epsilon}\right)^{2D-1}) \cap \mathrm{O}(\left(\frac{5}{2\epsilon}\right)^{2D}) \,. \tag{25}$$

Proof. The set of state vectors in a D-dimensional Hilbert space is isomorphic to a $(2D-1)$-sphere with radius 1 in $(2D)$-dimensional real Euclidean space such that the Hilbert space norm $|\cdot|_2$ on state vectors coincides with the Euclidean norm in $\mathbb{R}^{2D}$. The surface area of a $(n-1)$-sphere of radius r is $S_{n-1}(r) = nC_n r^{n-1}$, where $C_n = \pi^{n/2}/\Gamma(n/2+1)$ and Γ is the Euler gamma function. The set of states within Hilbert space distance ϵ to a given state is a spherical cap on that sphere with opening angle $4\arcsin(\epsilon/2)$. For $\epsilon \ll 1$, the area of such a cap is approximately equal to the volume of a $(2D-1)$-ball of radius ϵ. In fact, a more detailed analysis reveals that for $D = 3$ the two are exactly identical and for $D > 3$ the cap is always smaller than the $(2D-1)$-ball. The volume of an n-ball of radius r is $V_n(r) = C_n r^n$. Thus for $D \geq 3$,

$$\begin{aligned}\left(\frac{5}{2\epsilon}\right)^{2D} &\geq |\mathcal{N}_\epsilon^{\mathrm{HS}}(P)| \geq \frac{S_{2D-1}(1)}{V_{2D-1}(\epsilon)} = \frac{2DC_{2D}}{C_{2D-1}\epsilon^{2D-1}} \\ &= 2\sqrt{\pi}\frac{\Gamma(D+1/2)}{\Gamma(D)}\left(\frac{1}{\epsilon}\right)^{2D-1} \geq \frac{15\pi}{8}\left(\frac{1}{\epsilon}\right)^{2D-1},\end{aligned}$$

where the first inequality follows from Ref. [7]. □

This is essentially the argument used in Ref. [1] to establish that Hilbert space is a "convenient illusion". However, the lower bound on $|\hat{\mathcal{N}}_\epsilon^{\mathrm{HS}}(P)|$ does not immediately imply a lower bound on $|\hat{\mathcal{N}}_\epsilon^p(\mathcal{P})|$ (and hence also not for $|\hat{\mathcal{N}}_\epsilon^p(\mathcal{S})|$) for any $p \geq 1$. In particular, there are states with distance 2 in Hilbert space norm and distance 0 in any of the p-norms, namely, any pair of state vectors $\{|\psi\rangle, -|\psi\rangle\}$.

We now show that a similar lower bound as in the last lemma holds for the size of optimal ϵ-nets for $\mathcal{P}$ and $\mathcal{S}$ in p-norm.

Lemma 10. *For $p \in \{1, 2\}$*

$$|\hat{\mathcal{N}}_\epsilon^p(\mathcal{S})| \geq |\hat{\mathcal{N}}_{2\epsilon}^p(\mathcal{P})| \in \Omega(\left(\frac{1}{4\epsilon}\right)^{2D-3}). \tag{26}$$

Proof. For a given state vector $|\psi\rangle$ it will be convenient to use the notation $\psi := |\psi\rangle\langle\psi|$.

We start to prove the first inequality. Fix $p \in \{1, 2\}$. There is a family $\{\rho_j\} \subset \hat{\mathcal{N}}_\epsilon^p(\mathcal{S})$ such that their ϵ-neighborhoods in p-norm cover $\mathcal{P}$ and such that for each ρ_j there exists a rank-1 projector $\psi_j \in \mathcal{P}$ satisfying $\|\rho_j - \psi_j\|_p \leq \epsilon$. Then $\{\psi_j\}$ is a (2ϵ)-net for $\mathcal{P}$ in p-norm with $|\hat{\mathcal{N}}_\epsilon^p(\mathcal{S})| \geq |\{\psi_j\}| \geq |\hat{\mathcal{N}}_{2\epsilon}^p(\mathcal{P})|$.

From $\|\cdot\|_1 \geq \|\cdot\|_2$ it follows that $|\mathcal{N}_\epsilon^2(\mathcal{P})| \leq |\mathcal{N}_\epsilon^1(\mathcal{P})|$. Hence it remains to prove the lower bound for $|\mathcal{N}_{2\epsilon}^2(\mathcal{P})|$ in (26). For this we construct an ϵ'-net $\mathcal{N}_{\epsilon'}^{\mathrm{HS}}(P)$ for state vectors in Hilbert space norm from a (2ϵ)-net $\mathcal{N}_{2\epsilon}^2(\mathcal{P})$. For every element $\psi_j \in \mathcal{N}_{2\epsilon}^2(\mathcal{P})$ we fix an eingenvalue-1 eigenvector $|\psi_j\rangle$. Using the $(\epsilon^2/2)$-net $\mathcal{N}_{\epsilon^2/2}^1([0,1[) = \{\epsilon^2, 2\epsilon^2, \ldots, \lceil 1/\epsilon^2\rceil\epsilon^2\}$ for $[0,1[$ with cyclic boundary conditions we define the set

$$\mathcal{N}_{\epsilon'}^{\mathrm{HS}}(P) = \{\mathrm{e}^{2\pi\mathrm{i}\delta}\,|\psi\rangle : \delta \in \mathcal{N}_{\epsilon^2/2}^1([0,1[),\ |\psi\rangle \in \{|\psi_j\rangle\}\}\,. \tag{27}$$

This is an ϵ'-net for P and we will find an expression for ϵ' in terms of ϵ.

Let $|\phi\rangle \in P$. Then there exists a state vector $|\psi\rangle \in \{|\psi_j\rangle\}$ such that

$$\begin{aligned}(2\epsilon)^2 &\geq \|\phi - \psi\|_2^2 = 2 - 2|\langle\phi|\psi\rangle|^2 \\ &\geq 2 - 2|\langle\phi|\psi\rangle|\,,\end{aligned}$$

and a $\delta \in \mathcal{N}_{\epsilon^2/2}^1([0,1[)$ such that

$$\big|\,|\langle\phi|\psi\rangle| - \mathrm{Re}(e^{2\pi\mathrm{i}\delta}\langle\phi|\psi\rangle)\big| < (2\epsilon)^2\,.$$

Together this yields

$$3(2\epsilon)^2 > 2 - 2\mathrm{Re}(\mathrm{e}^{2\pi\mathrm{i}\delta}\langle\phi|\psi\rangle) = |\,|\phi\rangle - \mathrm{e}^{2\pi\mathrm{i}\delta}|\psi\rangle\,|_2^2\,.$$

Since $\mathrm{e}^{2\pi\mathrm{i}\delta}|\psi\rangle \in \mathcal{N}_{\epsilon'}^{\mathrm{HS}}(P)$, we can choose $\epsilon' = 4\epsilon > \sqrt{12}\epsilon$ to make $\mathcal{N}_{\epsilon'}^{\mathrm{HS}}(P)$ a (4ϵ)-net. From the definition (27) of $\mathcal{N}_{\epsilon'}^{\mathrm{HS}}(P)$ we can bound its cardinality

$$\begin{aligned}|\mathcal{N}_{4\epsilon}^{\mathrm{HS}}(P)| &= |\mathcal{N}_{\epsilon^2}^1([0,1[)|\,|\{|\psi_i\rangle\}| \\ &< \lceil 1/\epsilon^2\rceil|\mathcal{N}_{2\epsilon}^2(\mathcal{P})|\,,\end{aligned} \tag{28}$$

where we have used that by construction $|\{|\psi_i\rangle\}| = |\mathcal{N}_{2\epsilon}^2(\mathcal{P})|$. Finally, as the described construction works for any (2ϵ)-net $\mathcal{N}_{2\epsilon}^2(\mathcal{P})$, we obtain

$$\lceil 1/\epsilon^2\rceil\ |\hat{\mathcal{N}}_{2\epsilon}^2(\mathcal{P})| > |\hat{\mathcal{N}}_{4\epsilon}^{\mathrm{HS}}(P)| \tag{29}$$

and lemma 9 finishes the proof. □

Combining theorem 8 and lemma 10, we arrive at the following theorem:

Theorem 11 (Limitations of efficient state generation). *For every fixed initial state, the time evolution for a time interval of length τ under any k-local Liouvillian acting on N subsystems with local Hilbert space dimension d yields a state that lies inside one of N_T ϵ-balls in 1-norm with $\log(N_T) \in \mathrm{O}\left(N^{3k+2}\tau^4/\epsilon^5\right)$. For times τ polynomial in the system size N, this is asymptotically exponentially smaller than $\log|\hat{\mathcal{N}}_\epsilon^1(\mathcal{S})| \in \Omega(-2d^N)$ where $|\hat{\mathcal{N}}_\epsilon^1(\mathcal{S})|$ is the cardinality of an optimal ϵ-net in 1-norm that covers the state space $\mathcal{S}$.*

[1] D. Poulin, A. Qarry, R. D. Somma, and F. Verstraete, Phys. Rev. Lett. **106**, 170501 (2011).

[2] J. D. Dollard and C. N. Friedman, J. Math. Phys. **18**, 1598 (1977).
[3] J. Huyghebaert and H. De Raedt, J. Phys. A **23**, 5777 (1990).
[4] M. Suzuki, Phys. Lett. A, **146**, 319 (1990).
[5] W. F. Stinespring, Proc. Amer. Math. Soc. **6**, 211 (1955); V. Paulsen, *Completely Bounded Maps and Operator Algebras* (Cambridge University Press, 2002).
[6] C. M. Dawson and M. A. Nielsen, Quant. Inf. Comp. **6**, 81 (2006); A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, Am. Math. Soc. (2002).
[7] P. Hayden, D. W. Leung, and A. Winter, Comm. Math. Phys. **265**, 95 (2006).
[8] P. Hayden, D. W. Leung, P. W. Shor, and A. Winter, Commun. Math. Phys. **250**, 371 (2004).

Erratum: Dissipative Quantum Church-Turing Theorem [Phys. Rev. Lett. 107, 120501 (2011)]

M. Kliesch, T. Barthel, C. Gogolin, M. Kastoryano, and J. Eisert

(Received 21 August 2012; published 13 September 2012)

DOI: 10.1103/PhysRevLett.109.119904 PACS numbers: 03.67.Ac, 02.60.Cb, 03.65.Yz, 89.70.Eg, 99.10.Cd

Implication 3 in Ref. [1] is concerned with the classical simulation of k-local Markovian open system dynamics. While a slightly weaker statement compared to the one given is true [2], as can be proven using Lieb-Robinson bounds, this result is, strictly speaking, not implied by the argument given in Ref. [1].

In the physically important case where the local Liouvillian terms $\mathcal{L}_\Lambda$ act on k neighboring subsystems of a regular lattice of finite spatial dimension, the time evolution is quasilocal [2]. This means that, up to an exponentially small error, the diameter of the support of any evolved local observable grows at most linearly in time, which is reminiscent of a light cone effect. Consequently, the evolution of the local observable can be approximated to arbitrary precision by applying the propagator of a spatially truncated version of the Liouvillian. Hence, the time evolution can be simulated on classical computers with a cost that is independent of the system size.

Our original argument leading to implication 3 is wrong because it turns out that a truncation that is solely based on the causal structure of the channel circuit originating from the Trotter approximation does not yield quasilocal dynamics and alone does not establish an efficient classical simulation scheme.

We would like to thank T. J. Osborne for kindly pointing out this error to us.

[1] M. Kliesch, T. Barthel, C. Gogolin, M. Kastoryano, and J. Eisert, Phys. Rev. Lett. **107**, 120501 (2011).
[2] T. Barthel and M. Kliesch, Phys. Rev. Lett. **108**, 230504 (2012).

2.2.2 Lieb-Robinson bounds and quasi-locality of time evolution

In the last section we have seen that dissipative dynamics can be simulated efficiently on a quantum computer. This motivates the question, under what assumptions and in what sense Markovian dynamics can be simulated efficiently even on classical computers. It is strongly believed that the dynamics, even when restricted to local observables, cannot be simulated efficiently in time[2]. The local structure from the dissipative Trotter trotter circuits from the last publication is a "brick layer" structure (as in Figure 2.1 and Figure 1c on page 109). For one such circuit and a local observable only "bricks" insides the observables cone can influence the expectation value. This already suggests that one can simulate time evolved local observables efficiently in the system size. A proper error analysis, however, reveals that this implication does not hold in general. But nevertheless, we use Lieb-Robinson bound techniques to show that the error made when truncating the Liouvillian some distance away from the observable's Lieb-Robinson cone is exponentially small in that distance. This indeed implies that local dynamics can be simulated efficiently in the system size on classical computers. Technically, we slightly generalize a dissipative Lieb-Robinson bound from Ref. [Pou10] (see also Ref. [NVZ11]). The bound on the speed of propagation of correlations gives rise to a space-time cone for the considered observable or subsystem. This allows us to prove that one can truncate the Liouvillian at some distance away from the space-time cone and obtain the same time evolution up to an error exponentially small in that distance. Hence, the simulation cost is independent of the system size and is bounded inverse polynomially in the error. This answers the question from a purely complexity theoretic point of view. However, exploiting this in an exact diagonalization approach directly requires unfeasible many resources. In order to reduce the computation cost one can Trotter-approximate (see Publication [Kli+11a]) the truncated Liouvillian, as a first step. For the case without dissipation, these two approximations are the basic building blocks of *density matrix renormalization group (DMRG)* methods [ECP10, Eis13, Oru13, PGVWC07, Sch11, Sch13, VMC08]. With these methods one performs further approximation steps that exploit the fact that the generated states are close to a low-dimensional submanifold of state space. One remaining problem for dissipative systems is that these further approximations only preserve positivity of the time evolution if there is no dissipation present. This is not only a technical issue, but comes from a fundamental problem that we explain in detail in Section 2.4.2.

For the sake of completeness, the corresponding Publication [BK12] on quasi-locality of Markovian time evolution is presented in Appendix B.

[2] otherwise BQP=BPP, i.e., one could perform quantum computations classically efficiently

2.3 Locality of temperature from a new thermal Lieb-Robinson type bound

In this section we will see that, also when considering high temperature states, a locality structure allows for statements reminiscent of the quasi-locality of time evolution and the resulting simulatability result. It will turn out that the developed methods are closely interconnected with our physical understanding of temperature.

Basic statistical mechanics teaches us that temperature is *intensive*, or, in other words, a *local* quantity. This concept works perfectly well for non-interacting systems. But once a subsystem of interest interacts with its environment, the intensivity of temperature breaks down [HM05]: Interactions generate correlations that lead to noticeable deviations of the state of the subsystem from a thermal state. Hence, given only a subsystem state, there is no canonical way to assign a temperature to the subsystem. We call this the *locality of temperature problem*. Ultimately, this is a standard problem in quantum mechanics, as it is about studying the classical concept of temperature in the limit of small length scales. In the light of recent advances in the foundations of statistical mechanics [GLTZ06, LPSW09, RGE12], suggestions for small thermal machines working in the quantum regime [LPS10, ME12a], and experimental achievements of, e.g., thermometers on the nano-scale [GB02], a better understanding of the limitations of the concept of temperature on small scales is a pressing issue.

Progress in that and also other directions is hindered by a lack of mathematical tools exploiting the locality structure in thermal states. As we already saw, a lot is known about ground states [Has06, HW05, LVV13, MZ13, NS06], as powerful mathematical tools such as Lieb-Robinson bounds (see Publication [KGE14b] on pages 13ff) and the detectability lemma [AALV09, AAVL11] are available. For thermal states a lot less is known, in particular, when it comes to finite systems, which are not one-dimensional. The available explicit clustering results, for instance, are of the following type: In the limit of the distance between the supports of two observables being infinitely large, the observables become uncorrelated (w.r.t. the covariance) [PY95]. But physical intuition tells us that long-range correlations are rare and one should expect systems to have a finite correlation length in many situations. It is known, however, that thermal states at high temperature can be approximated by states of matrix product form with sub-exponentially many parameters in the system size [Has06]. While the implications seem to be effectively restricted to one-dimensional systems, this work develops an important cluster expansion technique that exploits the system's locality structure.

In the following Publication [Kli+14] we build on these methods and develop them further. Moreover, we prove a perturbation formula and establish a strong version of exponential clustering of correlations for high temperature states on fermionic and spin lattices. These results have a number of implications: First of all, it turns out that temperature is intensive on a given length scale if and only if correlations are negligible on that length scale. One important

point is, to identify the right correlation measure, which we call the *averaged generalized covariance*. After showing that above a critical temperature the *generalized covariance* decays exponentially, this implies the intensivity of high temperatures. The critical temperature only depends on the interaction strength and local geometry of the lattice. Therefore, this provides an upper bound on physical critical temperatures for thermal phase transitions with long-range order. Our results also imply that the generalized covariance is a measure of stability of thermal states against distant Hamiltonian perturbations. For high temperatures, we obtain a thermal Lieb-Robinson type bound (see, e.g., Figure 4 on page 51) and, as a rigorous computational consequence, thermal states can be locally simulated efficiently on classical computers. For smaller temperatures, our perturbation formula can provide a guideline of how the parameters should be chosen in classical simulations, such as Monte Carlo simulations [TATW03]. We also comment on relations to known results, such as actual physical critical temperatures, foundations of statistical mechanics [GLTZ06, LPSW09, RDO08, RGE12], matrix product operator approximations [Has06], and local topological quantum order [CLMPG13] (see pages 51ff). One more connection to the previous Section 2.2 is that the results presented here are also applicable to stationary states of Liouvillian time evolution with Davies generators [Dav79, Spo78] of commuting Hamiltonians. In this case, the Liouvillian is local and has the Hamiltonian's thermal state as unique stationary state.

In fact, the following Publication [Kli+14] has already proven useful to further develop MPO approximations of thermal states [MSVC14]. It is also worth mentioning a second work [MAMW13], slightly more recent than ours, on thermal states of lattice systems that also uses the locality structure.

Locality of Temperature

M. Kliesch,[1,*] C. Gogolin,[1] M. J. Kastoryano,[1] A. Riera,[1,2] and J. Eisert[1]

[1]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*

[2]*Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Am Mühlenberg 1, 14476 Potsdam-Golm, Germany*

(Received 17 September 2013; revised manuscript received 5 March 2014; published 31 July 2014)

This work is concerned with thermal quantum states of Hamiltonians on spin- and fermionic-lattice systems with short-range interactions. We provide results leading to a local definition of temperature, thereby extending the notion of "intensivity of temperature" to interacting quantum models. More precisely, we derive a perturbation formula for thermal states. The influence of the perturbation is exactly given in terms of a generalized covariance. For this covariance, we prove exponential clustering of correlations above a universal critical temperature that upper bounds physical critical temperatures such as the Curie temperature. As a corollary, we obtain that above the critical temperature, thermal states are stable against distant Hamiltonian perturbations. Moreover, our results imply that above the critical temperature, local expectation values can be approximated efficiently in the error and the system size.

DOI: 10.1103/PhysRevX.4.031019

Subject Areas: Condensed Matter Physics, Quantum Physics, Statistical Physics

I. INTRODUCTION

The ongoing miniaturization of devices, with structures reaching the nanoscale, has lead to the development of extremely small thermometers [1,2], some of which are so small that they can only be read out with powerful electron microscopes [3]. Even small thermal machines working in the quantum regime have been suggested [4,5]. In order to understand the working of such devices, it is necessary to formulate a theory of statistical mechanics and thermodynamics at the microscopic and mesoscopic scales. A prerequisite for such a formulation is a good understanding of the limitations of the concept of temperature at small scales.

The problem with assigning locally a temperature to a small subsystem of a globally thermal system is the following: Interactions between the subsystem and its environment that generate correlations can lead to noticeable deviations of the state of the subsystem from a thermal state (see Fig. 1). Hence, given only a subsystem state, there is no canonical way to assign a temperature to the subsystem. We call this the locality-of-temperature problem.

The first steps toward a solution of the locality-of-temperature problem have been taken in Refs. [6–8], and more recently, within the mind-set of quantum information theory, in Ref. [9]. The general locality-of-temperature problem is, however, still open. In this work, we conclusively solve it for spin- and fermionic-lattice systems.

More precisely, we first show that the locality-of-temperature problem is equivalent to a decay of correlations measured by an averaged generalized covariance that precisely captures the response of expectation values to perturbations of the Hamiltonian. We expect the corresponding equality to be useful for applications beyond the scope of this article.

We then provide conditions under which the generalized covariance decays exponentially with the distance, including a detailed analysis of the preasymptotic, and of the finite-size regime. In particular, this exponential decay holds above a universal critical temperature that only depends on the "connectivity" of the underlying graph of the model and is an upper bound on physically relevant critical temperatures such as the Curie temperature.

While, in the low-temperature regime, quantum lattice models exhibit a great diversity of phases, many of which involve the emergence of long-range or topological order [10], in the high-temperature regime, exponential clustering

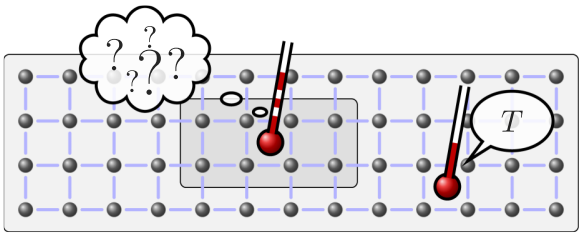

FIG. 1. The locality-of-temperature problem: Subsystems of thermal states are themselves, in general, not in a state with a locally well-defined temperature. Down to which length scale can temperature be an intensive quantity?

*martin.kliesch@gmail.com

of correlations is expected. Our rigorous results help to delineate the boundary between these two regimes. They build upon and go significantly beyond previous results on the clustering of correlations in classical systems [11], for quantum gases [12], i.e., translation-invariant Hamiltonians in the continuum, and cubic lattices [13–15].

Mathematically, we significantly contribute to the problem of whether and under which precise conditions thermal quantum states are stable against distant Hamiltonian perturbations. Stability of thermal states is particularly relevant in the broader scheme of phase transitions in classical and quantum lattice models [14,16] as well as for the foundations of statistical mechanics and the equilibration and thermalization behavior of closed quantum systems [17–25]. In the light of the recent surge of interest in these topics, developing a better understanding of the properties of thermal states has become a timely issue.

A major obstacle to progress on some of the most interesting open questions in this context, such as equilibration time scales in closed quantum systems, is the limited set of mathematical tools available for exploiting the structure of locally interacting Hamiltonians [25]. Our results are among the first that explicitly exploit properties of local Hamiltonians, without being limited to very specific models.

For quantum Monte Carlo simulations [26], our results provide a guideline as to how large the finite system size has to be taken in order to be able to sample from the right partition function and, conversely, to identify observables that are best suited to detect long-range correlations.

In fact, our results are reminiscent of known statements about ground states. If a Hamiltonian has a unique ground state and is gapped, correlations in its ground state cluster exponentially and faraway regions become essentially uncorrelated. This clustering of correlations is rigorously proven using information-theory-inspired methods such as Lieb-Robinson bounds and quasiadiabatic continuation [27–29]. These rigorous results allow for certified algorithms that efficiently approximate ground states of gapped Hamiltonians on classical computers [30]. In the same spirit, we are able to show that an exponential decay of correlations renders thermal states locally efficiently simulatable.

The rest of this paper is structured as follows: In Sec. II, we formulate the precise setting and explain the main results and their implications. In Sec. III, we discuss connections to known results on phase transitions, thermalization in closed quantum systems, and matrix product operator approximations. Then, in Sec. IV, we discuss basic properties of the generalized covariance, explain how our results can be made applicable to finite-range k-body interactions, and state the results for fermionic lattices. We proceed with proving all theorems in Sec. V and conclude in Sec. VI. In the Appendix, we provide a detailed proof of two bounds on truncated cluster expansions, one of which is an important ingredient to the proof of clustering of correlations.

II. SETTING AND MAIN RESULTS

In this section, we introduce the setting, state the locality-of-temperature problem more formally, and state our results.

A. Perturbation formula for thermal states

As the first result, we state a perturbation formula, which is a general statement about the response of the expectation value of an observable in the thermal state, upon changes in the system Hamiltonian. It does not make any reference to the locality structure of the Hamiltonian but turns out to be especially useful when correlations between local observables decay rapidly with distance.

Throughout the paper, we assume the Hilbert space to be finite dimensional [31] and denote the thermal state, or Gibbs state, of a Hamiltonian H at inverse temperature β by

$$g(\beta) := \frac{e^{-\beta H}}{Z(\beta)}, \tag{1}$$

with $Z(\beta) := \mathrm{Tr}(e^{-\beta H})$ being the partition function. If we mean the thermal state or partition function of a different Hamiltonian H', we write $g[H'](\beta)$ or $Z[H'](\beta)$.

We measure correlations by the (generalized) covariance that we define for any two operators A and A', full-rank quantum state ρ, and parameter $\tau \in [0,1]$ as

$$\mathrm{cov}_\rho^\tau(A,A') := \mathrm{Tr}(\rho^\tau A \rho^{1-\tau} A') - \mathrm{Tr}(\rho A)\mathrm{Tr}(\rho A'). \tag{2}$$

We discuss various properties of this covariance and generalizations to arbitrary-rank quantum states in Sec. IV A.

The generalized covariance appears naturally in our first theorem about the response of expectation values to perturbations. More precisely, when we are given an unperturbed Hamiltonian H_0 and a perturbed Hamiltonian H, then the difference of expectation values in the corresponding thermal states is captured by that covariance.

Theorem 1 (Perturbation formula).—Let H_0 and H be Hamiltonians acting on the same Hilbert space. For $s \in [0,1]$, define the interpolating Hamiltonian by $H(s) := H_0 + s(H - H_0)$ and denote its thermal state by $g_s := g[H(s)]$. Then,

$$\begin{aligned}&\mathrm{Tr}[A g_0(\beta)] - \mathrm{Tr}[A g(\beta)] \\ &\quad = \beta \int_0^1 d\tau \int_0^1 ds\, \mathrm{cov}_{g_s(\beta)}^\tau (H - H_0, A)\end{aligned} \tag{3}$$

for any operator A.

The proof of the theorem, which is presented in Sec. V A, relies on the fundamental theorem of calculus and Duhamel's formula. We refer to the double integral over

the covariance in Eq. (3) as the averaged (generalized) covariance.

B. Spin-lattice systems

In the remainder of this work, we will be concerned with spin- and fermionic-lattice systems. We will only write out everything for spin systems and then later, in Secs. IV C and V C, explain the necessary modifications for fermionic systems. In the case of spin-lattice systems, the Hilbert space is given by $\mathcal{H} = \bigotimes_{x\in V} \mathcal{H}_x$, where V is called the vertex set and is assumed to be finite. To make the presentation more accessible, many of the following definitions are highlighted in Fig. 2. A local Hamiltonian with interaction (hyper)graph (V, E) is a sum

$$H = \sum_{\lambda\in E} h_\lambda \tag{4}$$

of local Hamiltonian terms h_λ acting on $\mathcal{H}$. The (hyper) edge set E is the set of supports $\lambda = \mathrm{supp}(h_\lambda) \subset V$ of the local terms h_λ. For any subset of edges $F \subset E$, we denote by $H_F \coloneqq \sum_{\lambda\in F} h_\lambda$ the Hamiltonian only containing the interactions in F, and for any subsystem $B \subset V$, we define the truncated Hamiltonian to be $H_{\restriction B} \coloneqq H_{E(B)}$, where $E(B) \subset \{\lambda \in E : \lambda \subset B\}$ is the restricted edge set and we take $H_{\restriction B}$ to be an operator on the Hilbert space $\mathcal{H}_B \coloneqq \bigotimes_{x\in B} \mathcal{H}_x$.

Given some subsystem $S \subset V$, there are two natural thermal states associated with it.

(i) $g_{\restriction S}(\beta) \coloneqq g[H_{\restriction S}](\beta)$ denotes the thermal state of S alone, i.e., the thermal state of the truncated Hamiltonian $H_{\restriction S}(\beta)$.

(ii) $g^S(\beta) \coloneqq \mathrm{Tr}_{S^c}[g(\beta)]$ denotes the full thermal state reduced to S.

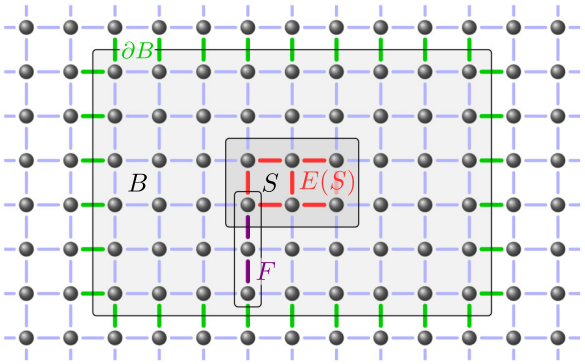

FIG. 2. A 2D square lattice: The boxes indicate subsystems $S \subset B \subset V$. The edges in S are $E(S)$, boundary edges of B are ∂B, and F is a shortest path connecting S and ∂B; hence, $d(S, \partial B) = |F| = 2$. The set of edges $E(S)$ is an example for an animal of size $|E(S)| = 7$, while ∂B is not connected and hence not an animal.

For a noninteracting Hamiltonian, these two states coincide, but, in general, this is not the case due to correlations between S and its environment. This discrepancy raises the question of how to locally define temperature as an intensive quantity, i.e., the locality-of-temperature problem.

C. Locality of temperature

In order to locally assign a temperature to the subsystem $S \subset V$, it was suggested, e.g., in Ref. [9], to extend S by a buffer region and define the temperature of S via the thermal state of the Hamiltonian truncated outside the extended region B; see Fig. 2. The role of the buffer region B is to remove the boundary effects and the correlations with the rest of the system that are intuitively the reason for the locality-of-temperature problem. Nevertheless, it is not obvious how these correlations should be quantified and how large this buffer region needs to be. We will see shortly that Theorem 1 answers these questions.

By $\partial B \subset E$, we denote the set of boundary edges of B, i.e., the edges having overlap with both B and its complement $B^c \coloneqq V\backslash B$. Then, by choosing $H_0 = H - H_{\partial B}$ in Theorem 1, using that $g_0 = g_{\restriction B} \otimes g_{\restriction B^c}$, and tracing over B^c, we obtain the following corollary (see also Fig. 3).

Corollary 1 (Truncation formula).—Let H be a local Hamiltonian, let $B \subset V$ be a subsystem, and denote the corresponding boundary Hamiltonian by $H_{\partial B}$ and the interpolating Hamiltonian by $H(s) \coloneqq H - (1-s)H_{\partial B}$ with its thermal state $g_s \coloneqq g[H(s)]$. Then, for any operator $A = A_B \otimes \mathbb{1}_{B^c}$ supported on B,

$$\mathrm{Tr}[A_B g_{\restriction B}(\beta)] - \mathrm{Tr}[A g(\beta)] = \beta \int_0^1 d\tau \int_0^1 ds\, \mathrm{cov}^\tau_{g_s(\beta)}(H_{\partial B}, A). \tag{5}$$

Now, we choose $S \subset B \subset V$ (see Fig. 2). If, for a given inverse temperature β, correlations over the distance between S and ∂B are negligible, then the corollary clearly implies that

$$\mathrm{Tr}[A g(\beta)] \approx \mathrm{Tr}[A_B g_{\restriction B}(\beta)] \tag{6}$$

for any observable $A_B = A_S \otimes \mathbb{1}_{B\backslash S}$ on S. Also note that such an approximate equality does not hold whenever average correlations over lengths exceeding the distance between S and ∂B are non-negligible.

Hence, we have the following equivalence for the temperature defined via thermal states.

Implication 1 (Locality of temperature).—Temperature is intensive on a given length scale if and only if correlations (measured by the averaged generalized covariance) are negligible compared to $1/\beta$ on that length scale.

In order to fully exploit Corollary 1 it is necessary to bound the generalized covariance, which we will do for high temperatures in the next section.

D. Clustering of correlations at high temperatures

For small temperatures, correlations can be arbitrarily long ranged, as is, e.g., the case for the ferromagnetic Ising model in two or higher dimensions below the Curie temperature. On the other hand, above a universal critical temperature, depending only on a local property of the interaction graph, correlations cluster exponentially, as we will see next. Given the combinatorial nature of parts of the arguments leading to this result, we need additional notation related to edges and vertices of the lattice. Most of the following definitions can be understood intuitively, as is shown in Fig. 2.

We say that two subsystems X, $Y \subset V$ overlap if $X \cap Y \neq \emptyset$, a set $X \subset V$ and a set $F \subset E$ overlap if F contains an edge that overlaps with X, and two sets F, $F' \subset E$ overlap if F overlaps with any of the edges in F'. A subset of edges $F \subset E$ connects X and Y if F contains a sequence of pairwise overlapping edges such that the first overlaps with X and the last overlaps with Y and similarly for the case where X and/or Y are just vertices.

The graph distance on V, and also the induced distance on subsets of V, are denoted by d. The distance $d(X, F)$ of a subset $X \subset V$ and a subset $F \subset E$ is 0 if X and F overlap and otherwise equal to the size of the smallest subset of E that connects X and F. Sometimes, we denote the support of an operator by the operator itself, e.g., for two operators A and A', their distance is $d(A, A') \coloneqq d(\operatorname{supp} A, \operatorname{supp} A')$ and $\partial A \subset E$ are the edges across the boundary of $\operatorname{supp}(A)$.

A subset of edges $F \subset E$ that connects all pairs of its elements λ, $\lambda' \in F$ is called connected. Such a connected set F is also called an (edge) animal. The size $|F|$ of an animal F is given by the number of edges contained in F. The results presented here apply to Hamiltonians with interaction graphs (V, E) whose number a_m of lattice animals of size m containing some fixed edge is exponentially bounded. With

$$a_m \coloneqq \sup_{\lambda \in E} |\{F \subset E \text{ connected} : \lambda \in F, |F| = m\}|, \tag{7}$$

the growth constant α is the smallest constant satisfying

$$a_m \leq \alpha^m. \tag{8}$$

For example, the growth constant of a D-dimensional cubic lattice can be bounded as $\alpha \leq 2De$ (Lemma 2 in Ref. [32]), where e is Euler's number. Moreover, α is finite for any regular lattice [33]. Upper bounds to growth constants for so-called spread-out graphs [32] render our results applicable for the case of bounded-range two-body interactions. By a simple embedding argument, one can also bound the growth constant for the case of local k-body interactions on a regular lattice, which we explain in Sec. IV B in detail.

For any operator A and $p \in [1, \infty]$, we denote by $\|A\|_p$ its Schatten p norm; e.g., $\|A\|_\infty$ is the operator norm and $\|A\|_1$ is the trace norm of A. We call $J \coloneqq \max_{\lambda \in E} \|h_\lambda\|_\infty$ the local interaction strength of a local Hamiltonian, as given in Eq. (4).

We are able to provide a universal inverse critical temperature β^*, which is, in particular, independent of the system size, below which correlations decay exponentially with a thermal correlation length $\xi(\beta)$.

Theorem 2 (Clustering of correlations at high temperatures).—Let $g(\beta)$ be the thermal state at inverse temperature β of a local Hamiltonian with finite interaction (hyper) graph (V, E) having growth constant α and local interaction strength J. Define the quantities

$$\beta^* \coloneqq \ln[(1 + \sqrt{1 + 4/\alpha})/2]/(2J) \tag{9}$$

and

$$\xi(\beta) \coloneqq |\ln[\alpha e^{2|\beta|J}(e^{2|\beta|J} - 1)]|^{-1}. \tag{10}$$

Then, for every $|\beta| < \beta^*$, parameter $\tau \in [0, 1]$, every two operators A and B with $d(A, B) \geq L_0(\beta, a)$ [given in Eq. (50)], and $a \coloneqq \min\{|\partial A|, |\partial B|\}$,

$$|\operatorname{cov}^\tau_{g(\beta)}(A, B)| \leq \frac{4a\|A\|_\infty \|B\|_\infty}{\ln(3)(1 - e^{-1/\xi(\beta)})} e^{-d(A,B)/\xi(\beta)}. \tag{11}$$

The proof is given in Sec. V B.

In the following sections, we outline some of the applications of Theorem 2.

E. Universal locality and stability at high temperatures

If one is interested in the state $g^S(\beta)$ of some subsystem S, then one can truncate the Hamiltonian to S extended by some buffer region and obtain the approximation via the thermal state of the truncated Hamiltonian. The following theorem implies that the approximation error is exponentially small in the width of the buffer region.

For any operator ρ, we denote its reduction to a subsystem $S \subset V$ by $\rho^S \coloneqq \operatorname{Tr}_{S^c}[\rho]$ and note that

$$\|\rho^S\|_1 = \sup\{|\operatorname{Tr}[A\rho]| : \operatorname{supp}(A) = S, \|A\|_\infty = 1\}. \tag{12}$$

Then, as a consequence of Corollary 1 and Theorem 2, we obtain the following corollary.

Corollary 2 (Universal locality at high temperatures).—Let H be a Hamiltonian satisfying the conditions of Theorem 2, let $|\beta| < \beta^*$, and let $S \subset B \subset V$ be subsystems with $d(S, \partial B) \geq L_0(\beta, |\partial S|)$. Then,

$$\|g^S(\beta) - g^S_{\restriction B}(\beta)\|_1 \leq \frac{v|\beta|J}{1 - e^{-1/\xi(\beta)}} e^{-d(S,\partial B)/\xi(\beta)}, \tag{13}$$

where $g^S_{\restriction B}$ denotes the thermal state of B reduced to S and $v \coloneqq 4|\partial S||\partial B|/\ln(3)$.

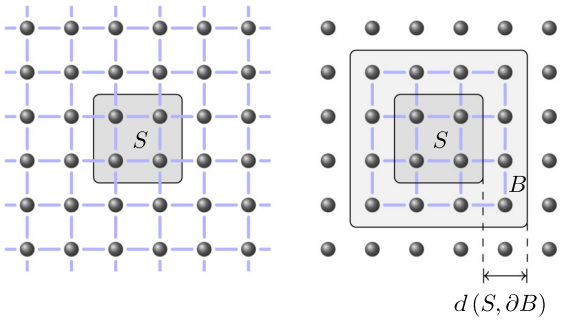

FIG. 3. The truncation from Corollary 2 and Implication 3: For $\beta < \beta^*$ and $d(S, \partial B) \ll \xi(\beta)$, Corollary 2 implies that $g^S(\beta)$, depicted on the left, and $g^S_{\restriction B}(\beta)$, depicted on the right, are approximately equal.

Similarly, as a corollary of Theorems 1 and 2, we obtain the following implication.

Implication 2 (Stability).—Below the critical inverse temperature β^* [from Eq. (9)], thermal states of local Hamiltonians are exponentially stable against distant, locally bounded perturbations.

F. Efficient approximation

Corollary 2 on the universal locality of thermal states also has the following complexity-theoretic consequence.

Implication 3 (Efficient approximation).—For $|\beta| < \beta^*$, local expectation values can be approximated with a computational cost independent of the system size and bounded polynomially in the reciprocal error.

In this sense, the error bound (see Fig. 4) of Corollary 2 is reminiscent of the quasilocality of dynamics, as, e.g., presented in Ref. [34], which is a consequence of

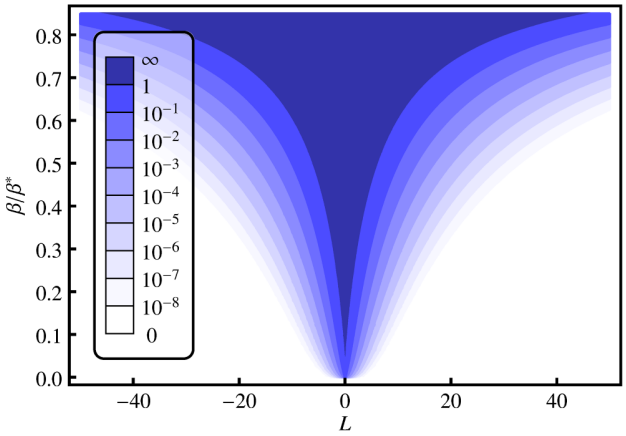

FIG. 4. One can obtain slightly tighter error bounds in Corollaries 2 and 5 by directly using Eq. (47). The plot shows this bound on the approximation error $\|g^S(\beta) - g^S_{\restriction B}(\beta)\|_1$ for the case of S being a single site on a 2D square lattice as a function of the inverse temperature β in units of the critical temperature and the width of the buffer region L. This bound can be seen as an imaginary-time Lieb-Robinson "cone" with diverging width as $\beta \to \beta^*$.

Lieb-Robinson bounds [35,36]. The quasilocality theorem [34] allows for an approximation of time-evolved local observables by truncating the Hamiltonian in the time-evolution operator at a distance $L > 0$ far away from the spacetime cone of the observable's support and has an approximation error that is exponentially small in L.

G. Fermions

In Ref. [37], it was shown for fermionic systems that two-point functions of observables that are odd polynomials in the fermionic operators decay exponentially with a correlation length proportional to the inverse temperature. Here, we obtain an exponential decay of the covariance above the critical temperature for all operators.

Observation 1 (Fermions).—All results also hold for locally interacting fermions on a lattice. See Theorem 4 and Corollaries 4 and 5 in Sec. IV C for the precise statements.

III. RELATIONS TO KNOWN RESULTS

In this section, we discuss the critical temperature from the clustering theorem, the connection of this work to concepts related to thermalization, and approximations of thermal states with so-called matrix product operators; as a last point, we briefly mention similarities with local topological quantum order.

A. Critical temperatures and phase transitions

Our results show that the quantity β^*, as defined in Eq. (9), provides a potentially coarse but universal and completely general upper bound on physical critical temperatures like the Curie temperature. For the ferromagnetic two-dimensional isotropic Ising model without an external field, our bound yields, for example, $1/(\beta^* J) = 2/\ln[(1 + \sqrt{1 + 1/e})/2] \approx 24.58$, whereas the phase transition between the disordered paramagnetic and the ordered ferromagnetic phases is known to really happen at $1/(\beta_c J) = 2/\ln(1 + \sqrt{2}) \approx 2.27$ [16]. Our universal bound is about an order of magnitude higher than the actual value for this example. To put this discrepancy into perspective, it is worth pointing out that it is a very difficult task to estimate physical critical temperatures—numerically or analytically. In fact, analytic expressions for critical temperatures or even just bounds on their values are known only for very few models.

One of the few known general statements is the Mermin-Wagner-Hohenberg theorem [38]. It states that in certain low-dimensional systems with short-range interactions, there cannot be any phase transition involving the spontaneous breaking of a continuous symmetry at any nonzero temperature. However, such systems can still have a low-temperature phase with quasi-long-range order characterized by power-law-like decaying correlations. Consequently, even for systems covered by the Mermin-Wagner-Hohenberg theorem, our Theorem 2 is nontrivial.

For example, it implies an upper bound on the critical temperature of the Kosterlitz-Thouless transition in the two-dimensional XY model [39].

In this work, we have concentrated on the general picture, but it seems likely that refinements of the methods employed and developed here can yield much tighter bounds on critical temperatures if more specific properties of a model are taken into account. At the same time, it remains an open problem to actually find a model with a phase transition with long-range order at the universal highest possible temperature.

B. Foundations of statistical mechanics

The recent years have seen a large number of numerical and experimental (see Ref. [21] for a review) as well as analytical investigations (see, for example, Refs. [17–19,22–24]) of equilibration and thermalization in closed quantum systems. In the focus of these works are the approach to equilibrium or properties of energy eigenstates. The current work complements this body of literature in that it shows fundamental properties of systems in thermal equilibrium. A feature that makes the current work unique is that, contrary to essentially all other works, the results derived here explicitly use the structure of locality interacting systems. (Noteworthy exceptions are Ref. [24] and, albeit in a very special setting, Ref. [18].)

The locality of thermal states is also of interest for recent results [24] on the dynamical thermalization of translation-invariant lattice models: Our Corollary 2 guarantees the existence of a "unique phase" [24] for all temperatures above our critical temperature. Hence, it implies that at sufficiently high temperatures, Theorems 1, 2, and 3 of Ref. [24] are applicable for any translation-invariant Hamiltonian.

There is also an interesting connection of our locality and stability results to the so-called eigenstate-thermalization hypothesis (ETH) [20,21]. The ETH essentially conjectures that the expectation values of certain physically relevant observables (for example, local ones) in energy eigenstates of sufficiently complex Hamiltonians should be very similar to the expectation values in thermal states with the same average energy. Corollary 2 and Implication 2 thus imply that the eigenstates of a Hamiltonian in the center of the spectrum (which correspond to high-temperature thermal states) must, if the Hamiltonian fulfills the ETH, also be locally stable against perturbations of the Hamiltonian. This insight could put constraints on the class of Hamiltonians that fulfills the ETH, provide new insights into the properties of their eigenstates, and open up new ways to test the ETH.

C. MPO approximation of thermal states

Matrix product operators (MPOs) are a certain class of operators that are tractable on classical computers for one-dimensional systems. Therefore, they play an important role in numerical simulations based on so-called tensor networks.

An important ingredient to our proof of Theorem 2 on the clustering of correlations will be a bound on a truncated cluster expansion (Lemma 1). The original result on the cluster expansion (Lemma 2 in the Appendix) is due to Hastings and was first used to approximate thermal states with inverse temperature below $2\beta^*$ by MPOs [40]. This approximation is summarized in the next theorem.

In one spatial dimension, this MPO approximation yields a tensor size bounded polynomially in the system size and the approximation error (see the subsequent corollary). In higher dimensions, however, the MPO approximation yields a tensor size bounded only subexponentially in the system size and is hence not computationally efficient, albeit exponentially cheaper than storing the full density matrix $g(\beta)$. In order to explain the MPO approximation in more detail, we start the discussion with a slightly nonstandard definition of MPOs.

Definition 1 (MPO).—Let $(b[x]^{(j)})_{j=1}^{d^2}$ be a basis for the operators on $\mathcal{H}_x$ and write an arbitrary operator A on $\mathcal{H}$ in the product basis as

$$A = \sum_{k\in[d^2]^V} A_k \bigotimes_{x\in V} b[x]^{(k_x)}, \tag{14}$$

with expansion coefficients $A_k \in \mathbb{C}$ and where $[d^2] := \{1, 2, \ldots, d^2\}$. If the A_k are of the form

$$A_k = \prod_{x\in V} a[x](k), \tag{15}$$

where every $a[x](k)$ only depends on at most r of the $|V|$ indices k_x, then A is called an MPO with tensor size d^{2r}.

Thermal states can be approximated by such MPOs. The following theorem is a consequence of Lemma 2, which we will prove in the Appendix along with Lemma 1.

Theorem 3 (MPO approximation of thermal states [40]).—Let $H = \sum_{\lambda\in E} h_\lambda$ be a local Hamiltonian with finite interaction graph (V, E) having a growth constant α and local interaction strength $J = \max_{\lambda\in E}\|h_\lambda\|_\infty$, and define $b(\beta J) := \alpha e^{|\beta J|}(e^{|\beta J|} - 1)$. Moreover, let β be small enough such that $b(\beta J) < 1$. Then, for each $L \in \mathbb{Z}^+$, there exists a self-adjoint MPO $\rho(\beta, L)$ [given in Eq. (A3)] with tensor size $d^{2N(L)}$, where

$$N(L) := \sup_{x_0\in V} |\{x \in V : d(x, x_0) < L\}| \tag{16}$$

is the number of vertices within a distance less than L. The approximation error is bounded as

$$\|g(\beta) - \rho(\beta, L)\|_1 \leq \exp\left(|E|\frac{b(\beta J)^L}{1 - b(\beta J)}\right) - 1; \tag{17}$$

i.e., for fixed $|\beta J| < b^{-1}(1)$, the trace norm difference scales as $O(|E|e^{-|\ln[b(\beta J)]|L})$ for large enough L.

In particular, the above theorem implies the following corollary.

Corollary 3 (Bound on the tensor size).—Let D be the spatial dimension of the Hamiltonian's interaction graph (V, E), let $n := |E|$ be the system size, and let $\beta < 2\beta^*$ with β^* from Eq. (9). Then, the MPO approximation in Theorem 3 gives rise to a tensor size of the MPO $\rho(\beta, L)$ scaling as

$$\log_d(\text{tensor size}) \leq O[\ln(Cn/\epsilon)^D], \tag{18}$$

with some β-dependent constant C. In particular, for $D = 1$, the bound on the tensor size scales polynomially with n/ϵ.

Let us consider a one-dimensional system and suppose we are explicitly given the MPO tensors $a'[x]$ [see Eq. (15)] of an approximation to a state ρ and, similarly, an observable A of MPO form with MPO tensors $a[x]$. If the tensor sizes of both MPOs scale at most polynomially in the system size, then one can compute the corresponding approximation to the expectation value $\mathrm{Tr}(\rho A)$ with a computational cost scaling polynomially in the system size. Thus, for instance, global product observables can be approximated efficiently, which is not guaranteed by our Implication 3. The problem with the MPO approximation, however, is that Theorem 3 only guarantees the existence of the MPO tensors but it is not obvious how they can be computed (efficiently).

Proof of Corollary 3.—The condition $\beta < 2\beta^*$ is equivalent to $b(\beta J) < 1$. Let us denote the bound to the approximation error in Eq. (17) by ϵ. Note that the upper bound in Eq. (17) satisfies

$$\epsilon := \exp\left(|E|\frac{b(\beta J)^L}{1 - b(\beta J)}\right) - 1 \leq Cn\, b(\beta J)^L \tag{19}$$

for distances L being at least logarithmically large in $n = |E|$ and some β-dependent constant C. Then, the distance L necessary to reach ϵ must asymptotically be at least as large as

$$L \geq \frac{\ln(Cn/\epsilon)}{|\ln[b(\beta J)]|}. \tag{20}$$

Bounding $N(L)$ in terms of the spatial dimension D as $N(L) \leq ML^D$ with some constant M yields a tensor size bounded as

$$\log_d(\text{tensor size}) \leq 2M\left(\frac{\ln(Cn/\epsilon)}{\ln[1/b(\beta J)]}\right)^D. \tag{21}$$

■

D. Local topological quantum order

It is worth mentioning that Corollary 2 and Implication 2 are very reminiscent of the local topological quantum order condition for open quantum systems introduced in Ref. [41] and the results on the local stability of stationary states of local Liouvillians in Ref. [42]. A slightly different family of local topological quantum order conditions for closed quantum systems [41–44] has played a very important role in the theory of locally stable (topological) lattice systems and for rigorous proofs of entropic area laws. Corollary 2 similarly characterizes the regime where local perturbations cannot drive any thermal phase transition.

IV. DETAILS

In this section, we first discuss the generalized covariance and then provide details concerning the applicability of our results to Hamiltonians with k-body interactions. Finally, we justify Observation 1 by stating the fermionic versions of our results.

A. The generalized covariance

The generalized covariance defined in Eq. (2), which depends on a parameter $\tau \in [0, 1]$, provides more information about the correlations between two observables than the standard covariance in a similar way as the class of Rényi entropies characterizes more completely the entanglement properties of a state than simply the von Neumann entropy [45]. While it occurs quite naturally in the perturbation formula of Theorem 1, other possible applications are to be explored. Here, we discuss possible generalizations of the generalized covariance to operators of arbitrary rank, show that for operators A and A' they are always bounded by $\|A\|_\infty\|A'\|_\infty$, and comment on convexity and a symmetrized version of the generalized covariance.

A definition of the generalized covariance for states of arbitrary rank is not relevant for this work because for nonzero temperature, thermal states are full-rank operators. However, the discussion of possible generalizations also hints at the behavior of cov^τ at the end points of the unit interval. On the open interval $\tau \in]0, 1[$, it is natural to simply keep the definition from Eq. (2). There are two natural ways to define ρ^0: Either as $\rho^0 := \mathbb{1}$ or as $\rho^{0+} := \lim_{\tau\to 0}\rho^\tau$, where ρ^{0+} turns out to be the projector onto the image of the operator ρ. For each end point $\tau = 0$ and $\tau = 1$, there are hence two natural ways to define cov^τ, either such that the generalized covariance is continuous or such that $\mathrm{cov}^0_\rho(A, A') = \mathrm{cov}_\rho(A', A)$ and $\mathrm{cov}^1_\rho(A, A') = \mathrm{cov}_\rho(A, A')$, where

$$\mathrm{cov}_\rho(A, A') := \mathrm{Tr}(\rho AA') - \mathrm{Tr}(\rho A)\mathrm{Tr}(\rho A') \tag{22}$$

defines the standard covariance.

Note that for product states and operators with disjoint support, all versions of the generalized covariance vanish. Moreover, for pure states, the continuous version of the generalized covariance vanishes also, meaning that classical correlations are needed to yield a nonzero value.

Next, we show that the generalized covariance is always bounded as

$$|\operatorname{cov}_\rho^\tau(A,A')| \leq \|A\|_\infty \|A'\|_\infty, \tag{23}$$

irrespective of which definitions are chosen for cov^0 and cov^1. We consider a state ρ and define $\bar{A} := A - \operatorname{Tr}(\rho A)$. Then,

$$\operatorname{cov}_\rho^\tau(A,A') = \operatorname{Tr}(\rho^\tau \bar{A} \rho^{1-\tau} A'). \tag{24}$$

Hölder's inequality generalized to several operators and the fact that $\|X\|_p = \||X|^p\|_1^{1/p}$ then imply that

$$|\operatorname{cov}_\rho^\tau(A,A')| \leq \|\rho^\tau\|_{1/\tau} \|\bar{A}\|_\infty \|\rho^\tau\|_{1/(1-\tau)} \|A'\|_\infty \tag{25}$$

$$= \|\bar{A}\|_\infty \|A'\|_\infty, \tag{26}$$

and, by noting that $\|\bar{A}\|_\infty = \|A\|_\infty$, the bound (23) is proven for the continuous version of the generalized covariance. For the noncontinuous versions, the bound follows similarly.

The variance $\operatorname{cov}_\rho^\tau(A,A)$ induced by the continuous version of the covariance is convex in τ, as can be seen by writing out ρ in its eigenbasis. As one can change the sign of $\operatorname{cov}_\rho^\tau(A,A')$ by just changing the sign of A', the generalized covariance is not convex in τ. But, it might be that its magnitude $|\operatorname{cov}_\rho^\tau(A,A')|$ is convex, which is unclear. If this were the case, it would be enough to prove the clustering Theorem 2 only for the end points $\tau \in \{0,1\}$, and hence the proof could be significantly simplified.

Similarly, as there is a symmetrized version of the standard covariance, one can also symmetrize the generalized covariance with respect to the two operators. Because of the cyclicity of the trace, the generalized covariance satisfies the symmetry property

$$\operatorname{cov}_\rho^\tau(A,A') = \operatorname{cov}_\rho^{1-\tau}(A',A). \tag{27}$$

Hence, one can define the symmetrized version of the generalized covariance as follows:

$$\overline{\operatorname{cov}}_\rho^\tau(A,A') := \frac{1}{2}[\operatorname{cov}_\rho^\tau(A,A') + \operatorname{cov}_\rho^\tau(A',A)]. \tag{28}$$

Our results can also be phrased in terms of this symmetrized version, since the averaged generalized covariance in the perturbation formula of Theorem 1 can easily be rewritten in terms of $\overline{\operatorname{cov}}$, and a bound analogous to the clustering of Theorem 2 holds also for the symmetrized quantity.

B. Bound on the growth constant for local k-body interactions

In this section, we show that regular hyperlattices also have a finite growth constant, which renders our results applicable to Hamiltonians with local k-body interactions.

In the case of k-body interactions, the Hamiltonian is again a sum of local terms h_λ whose supports are hyperedges $\lambda = \operatorname{supp}(h_\lambda) \subset V$ with $|\lambda| \leq k$. As before, V denotes the vertex set and E the set of hyperedges.

We assume that the interaction hypergraph (V,E) is a regular hyperlattice, i.e., that it can be embedded into a regular hypercubic lattice of a certain dimension D with hyperedges of hypercubic form. Let us denote by R the edge length of the resulting hypercubes. Note that such an embedding is, in general, not unique and changes both the number of terms in the Hamiltonian and the local interaction strength of H. Moreover, the grouping changes the values of the metric d in our results.

In order to find an exponential upper bound to the number a_m of hyperanimals composed of m hypercubes, let us define a spread-out graph of range R as the graph with the edge set consisting of all pairs $\{x,y\}$ with $0 < \|x-y\|_\infty \leq R$ and $x,y \in \mathbb{Z}^D$ (see Ref. [32]). Notice that as any hypercube is uniquely specified by the coordinates of its "lower left corner," any hyperanimal of size m corresponds to a lattice animal of size $m-1$ and range R in the spread-out graph. It follows from Lemma 2 in Ref. [32] that $a_m \leq (Ke)^m$ with $K = (2R+1)^D - 1$ being the coordination number. Hence, the hyperlattice has a growth constant bounded by $\alpha \leq [(2R+1)^D - 1]e$.

The bound obtained is, for most models, far from optimal, in particular, in situations where the supports of the local Hamiltonian terms are very different from hypercubes. For such cases, one can derive tighter but more specific bounds from known results about lattice animals in a similar way.

C. Fermionic versions of the main results

To make Observation 1 about fermions precise, we introduce the setting of interacting fermions on lattices. For each site $x \in V$, the corresponding fermionic operators, i.e., the creation and annihilation operators $f_x^\dagger$ and f_x, act on the fermionic Fock space and satisfy

$$\{f_x, f_y^\dagger\} = \delta_{x,y}\mathbb{1}, \tag{29}$$

$$\{f_x, f_y\} = 0, \tag{30}$$

where $\{A,B\} := AB + BA$ is the anticommutator. For such systems, all operators can be given in terms of polynomials in the fermionic operators. A monomial of fermionic operators is called even (odd) if it can be written as a product of an even (odd) number of fermionic operators f_x and $f_y^\dagger$. A polynomial of fermionic operators is called even (odd) if it can be written as a linear combination of only

even (odd) monomials, and an operator is called even (odd) if it can be written as an even (odd) polynomial of fermionic operators. According to the fermion-number-parity superselection rule, only operators that are even polynomials in the fermionic operators are physical observables and Hamiltonians.

As with spin-lattice systems, we have again a finite interaction graph (V, E); however, the support of an operator is now to be understood in the picture of the second quantization as follows: The support of any operator A being a polynomial in the fermionic operators is the set of vertices of the fermionic operators that occur in the polynomial. Correspondingly, we denote the algebra of the even operators supported on a region $X \subset V$ by $\mathcal{G}_X$ and denote $\mathcal{G} := \mathcal{G}_V$ for short. The Hamiltonian of a fermionic-lattice system is of the form

$$H = \sum_{\lambda \in E} h_\lambda \tag{31}$$

with $h_\lambda \in \mathcal{G}_\lambda$. For $B \subset V$, the truncated Hamiltonian $H_{\restriction B}$ is similarly the sum only over the edges contained in B. As for spin systems, $H_{\partial B}$ is the sum over the boundary edges of B.

Theorem 1 also holds for such fermionic-lattice systems, and we can prove statements analogous to Corollary 1, Theorem 2, and Corollary 2. Hence, all implications stated in Sec. II also hold. All proofs are presented in Sec. V C.

Corollary 4 (Fermionic truncation formula).—Let $H = \sum_{\lambda \in E} h_\lambda$ be a fermionic local Hamiltonian with local terms $h_\lambda \in \mathcal{G}$, let $B \subset V$ be a subsystem, and let the interpolating Hamiltonian by $H(s) := H - (1-s)H_{\partial B}$ with its thermal state $g_s := g[H(s)]$. Then, for any operator A with support $\mathrm{supp}(A) \subset B$,

$$\mathrm{Tr}\{Ag[H_{\restriction B}](\beta)\} - \mathrm{Tr}[Ag(\beta)] = \beta \int_0^1 d\tau \int_0^1 ds\, \mathrm{cov}^\tau_{g_s(\beta)}(H_{\partial B}, A). \tag{32}$$

Theorem 4 (Clustering of correlations in fermionic systems).—Let $g(\beta)$ be the thermal state at inverse temperature β of a local fermionic Hamiltonian $H = \sum_{\lambda \in E} h_\lambda$ with finite interaction graph (V, E) having growth constant α, local terms $h_\lambda \in \mathcal{G}$, and local interaction strength J. Define the functions β^*, ξ, and L_0 as in Eqs. (9), (10), and (50). Then, for every $|\beta| < \beta^*$, $\tau \in [0,1]$, and every two operators A and B with $d(A,B) \geq L_0(\beta, a)$, where $a := \min\{|\partial A|, |\partial B|\}$,

$$|\mathrm{cov}^\tau_{g(\beta)}(A,B)| \leq \frac{4a\|A\|_\infty \|B\|_\infty}{\ln(3)(1 - e^{-1/\xi(\beta)})} e^{-d(A,B)/\xi(\beta)}. \tag{33}$$

Corollary 5 (Locality of fermionic thermal states).—Let H be a Hamiltonian satisfying the conditions of Theorem 4, let $|\beta| < \beta^*$, and let $S \subset B \subset V$ be subsystems with $d(S, \partial B) \geq L_0(\beta, |\partial S|)$. Then,

$$\|g^S(\beta) - g^S[H_{\restriction B}](\beta)\|_1 \leq \frac{v|\beta|J}{1 - e^{-1/\xi(\beta)}} e^{-d(S,\partial B)/\xi(\beta)}, \tag{34}$$

where $v = 4|\partial S||\partial B|/\ln(3)$.

V. PROOFS

We start this section with the proofs of Theorems 1 and 2. One important stepping stone for the proof of the latter is a tailored version of a bound on a truncated cluster expansion (Lemma 1) from Ref. [40]. Both versions are proven in the Appendix. In the last part of the section, we prove the fermionic versions of our main results, Theorem 4 and Corollaries 4 and 5.

A. Proof of the perturbation formula (Theorem 1)

The two main ingredients in the proof of Theorem 1 are the fundamental theorem of calculus and Duhamel's formula. The generalized covariance appears as a natural measure of correlations.

Proof of Theorem 1.—Using the fundamental theorem of calculus, we obtain

$$\mathrm{Tr}[Ag_0(\beta)] - \mathrm{Tr}[Ag_1(\beta)] = -\mathrm{Tr}\left(A \int_0^1 \frac{d}{ds} \frac{e^{-\beta H(s)}}{Z_s(\beta)} ds\right)$$

with $Z_s := Z[H(s)]$. The derivative can be written as

$$\frac{d}{ds}\frac{e^{-\beta H(s)}}{Z_s(\beta)} = \frac{1}{Z_s(\beta)}\frac{d}{ds} e^{-\beta H(s)} - \frac{g_s(\beta)}{Z_s(\beta)} \mathrm{Tr}\left(\frac{d}{ds} e^{-\beta H(s)}\right).$$

After applying Duhamel's formula to both derivatives, i.e., using that

$$\frac{d}{ds} e^{-\beta H(s)} = -\beta \int_0^1 (e^{-\beta H(s)})^\tau \left(\frac{d}{ds} H(s)\right) (e^{-\beta H(s)})^{1-\tau} d\tau,$$

we obtain

$$\mathrm{Tr}(Ag_0) - \mathrm{Tr}(Ag) = -\beta\, \mathrm{Tr}\left(A \int_0^1 \int_0^1 \{-g_s^\tau (H - H_0) g_s^{1-\tau} + g_s \mathrm{Tr}[g_s^\tau (H - H_0) g_s^{1-\tau}]\} d\tau ds\right). \tag{35}$$

Together with the cyclicity of the trace and the definition of the generalized covariance in Eq. (2), the last equation finishes the proof. ∎

B. Proof of Theorem 2 on clustering of correlations

The proof of Theorem 2 builds on and develops further a cluster expansion of the power series of $e^{-\beta H}$ in terms of summands of the form

$$h(w) := h_{w_1} h_{w_2} \ldots h_{w_{|w|}}, \tag{36}$$

where $w_j \in E$. For the sake of a compact presentation, we refer to edges as letters, to the edge set E as an alphabet, and call sequences of edges words. For any subalphabet $F \subset E$, we denote by $F^* := \bigcup_{l=0}^{\infty} F^l$ the set of words with letters in F and arbitrary length l, where the length $|w|$ of a word $w \in E^*$ is the total number of letters it contains. For two words $w, v \in E^*$, their concatenation is denoted by $w \circ v := (w_1, w_2, \ldots, w_{|w|}, v_1, v_2, \ldots, v_{|v|})$. We call a word $c \in E^*$ connected or a cluster if the set of letters in c is an animal, i.e., connected. So, clusters are connected sequences of edges where the edges can also occur multiple times, while animals are sets of edges without any order or repetition. A word v is called a subsequence of $w \in E^*$ if v can be obtained from w by omitting letters, i.e., if there is an increasing sequence $j_1 < j_2 < \cdots < j_{|v|}$ such that $v_i = w_{j_i}$. This will be denoted by $v \subset w$. A connected subsequence $c \subset w$ is called a maximal cluster of w if c is not a subsequence of any other connected subsequence of w. Importantly, for any word $w \in E^*$, one can permute its letters to a new word w' such that $h(w') = h(w)$, irrespective of the choice of the local terms h_λ and such that $w' = c_1 \circ c_2 \circ \cdots \circ c_k$ is a concatenation of maximal clusters $c_j \subset w$ of w. Note that this decomposition is unique up to the order of the c_j.

In the following, we will consider systems that are either $n = 2$ or $n = 4$ copies of the original system with Hilbert space $\mathcal{H}$. For any operator A on $\mathcal{H}$, we denote by $A^{(j)}$ the operator on $\mathcal{H}^{\otimes n}$ that acts as A on the jth copy, e.g., $A^{(2)} := \mathbb{1} \otimes A$ for $n = 2$. By $\mathcal{S}^{(i,j)}$, we denote the swap operator on $\mathcal{H}^{\otimes n}$ that swaps the ith and jth tensor factors, e.g., $\mathcal{S}^{1,2}|k_1, k_2, k_3, k_4\rangle = |k_2, k_1, k_3, k_4\rangle$ for $n = 4$. For $n = 2$, we write $\mathcal{S}$ instead of $\mathcal{S}^{1,2}$.

We can now state the subsequent lemma, which is a bound on a truncated cluster expansion that is based on a more general, but for our purposes not tight enough bound, used previously in Ref. [40]. (See Lemma 2 in the Appendix.) The lemma will play an important role in the subsequent proof of Theorem 2.

Lemma 1 (Truncated cluster expansion).—Let $\tau \in [0, 1]$ and $H = \sum_{\lambda \in E} h_\lambda$ be a local Hamiltonian on $\mathcal{H}$ with finite interaction graph (V, E) having growth constant α and local interaction strength $J = \max_{\lambda \in E} \|h_\lambda\|_\infty$. We denote by $\tilde{H}$ the Hamiltonian of two weighted copies with local terms $\tilde{h}_\lambda := \tau h_\lambda^{(1)} + (1 - \tau) h_\lambda^{(2)}$. Consider two operators A and B on $\mathcal{H}$, define $b(x) := \alpha e^{|x|}(e^{|x|} - 1)$, and let $|\beta|$ be small enough such that $b(\beta J) < 1$. For some set of edges $F \subset E$, let $\mathcal{C}_{\geq L}(F) \subset E^*$ be the set of words containing at least one cluster c that contains at least one letter of F and has size $|c| \geq L$ (see Fig. 5) and let us denote the corresponding truncated cluster expansion of $e^{-\beta \tilde{H}}$ by

$$\Omega[\tilde{H}](\beta) := \sum_{w \in \mathcal{C}_{\geq L}(F)} \frac{(-\beta)^{|w|}}{|w|!} \tilde{h}(w), \tag{37}$$

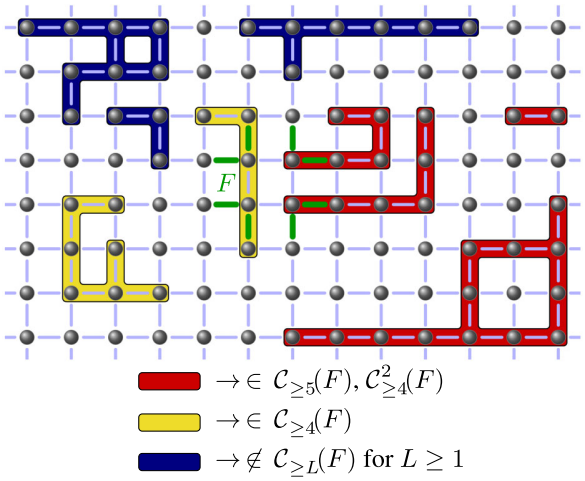

FIG. 5. A 2D square lattice. Three different subalphabets are indicated: Words that contain all letters in those alphabets are members of different sets $\mathcal{C}_{\geq L}(F)$.

with $\tilde{h}(\omega) := \tilde{h}_{\omega 1} \tilde{h}_{\omega 2} \ldots \tilde{h}_{\omega |\omega|}$. Then, for all $\tau \in [0, 1]$,

$$\frac{|\mathrm{Tr}[\mathcal{S} A^{(1)} B^{(2)} \Omega[\tilde{H}](\beta)]|}{\|A\|_\infty \|B\|_\infty Z(\beta)} \leq \exp\left(|F| \frac{b(\beta J)^L}{1 - b(\beta J)} \right) - 1. \tag{38}$$

We provide a detailed proof of this lemma in the Appendix. The terms resulting from the expansion of the exponential series of $e^{-\beta H}$ are classified according to whether they contain a cluster of size at least L that contains a letter from F. One can then show that there is a percolation transition at $\beta^* = b^{-1}(1)/(2J)$ such that for $|\beta| < \beta^*$, the contribution of long clusters is exponentially suppressed.

In the following proof of the exponential clustering, we will use the so-called swap trick: For any two operators A and B, it holds that

$$\mathrm{Tr}(AB) = \mathrm{Tr}[\mathcal{S}(A \otimes B)], \tag{39}$$

which can be checked by a straightforward calculation.

Proof of Theorem 2.—Fix some $\tau \in [0, 1]$. For any operator $A: \mathcal{H} \to \mathcal{H}$, we define $A^{(\pm)} := A \otimes \mathbb{1} \pm \mathbb{1} \otimes A$ and $\tilde{A}^{(+)} := \tau(A^{(1)} + A^{(2)}) + (1 - \tau)(A^{(3)} + A^{(4)})$.

As the first step, we write the covariance as

$$\mathrm{cov}_\rho^\tau(A, B) = \frac{1}{2} 2 \,\mathrm{Tr}[A^{(-)}(\rho^\tau \otimes \rho^\tau) B^{(-)}(\rho^{1-\tau} \otimes \rho^{1-\tau})].$$

Using the swap trick (39) yields (see Fig. 6)

$$\mathrm{cov}_\rho^\tau(A, B) = \frac{1}{2} 2 \,\mathrm{Tr}[\mathcal{S}^{1,3} \mathcal{S}^{2,4} (A^{(-)} \otimes B^{(-)}) \rho_4], \tag{40}$$

where $\rho_4 := \rho^\tau \otimes \rho^\tau \otimes \rho^{1-\tau} \otimes \rho^{1-\tau}$. For the case $\rho = g(\beta)$, the operator ρ_4 turns out to be

$$\rho_4 = \frac{e^{-\beta\tilde{H}^{(+)}}}{Z(\beta)^2}. \tag{41}$$

Writing out ρ_4 as a power series yields

$$\mathrm{cov}^{\tau}_{g(\beta)}(A,B) = \frac{1}{2Z(\beta)^2}\sum_{w\in E^*}\frac{(-\beta)^{|w|}}{|w|!}t(w) \tag{42}$$

with

$$t(w) := \mathrm{Tr}[\mathcal{S}^{1,3}\mathcal{S}^{2,4}(A^{(-)}\otimes B^{(-)})\tilde{h}^{(+)}(w)] \tag{43}$$

and $\tilde{h}^{(+)}(w) := \tilde{h}^{(+)}_{w_1}\tilde{h}^{(+)}_{w_2}\ldots\tilde{h}^{(+)}_{w_{|w|}}$. Next, we argue that $t(w)$ vanishes whenever w does not contain a cluster connecting the supports of A and B. Without loss of generality, we assume that $|\partial A| \le |\partial B|$ and consider $\mathcal{C}_{\ge L}(\partial A)^c = E^*\backslash\mathcal{C}_{\ge L}(\partial A)$, the set of words that do not contain a cluster containing an edge in ∂A of size $L := d(A,B)$ or larger. The set $\mathcal{C}_{\ge L}(\partial A)^c$ hence contains no words with clusters that connect $\mathrm{supp}(A)$ and $\mathrm{supp}(B)$. Any word $w \in \mathcal{C}_{\ge L}(\partial A)^c$ can be replaced by a concatenation of two words w_A and w_B such that $\tilde{h}^{(+)}(w) = \tilde{h}^{(+)}(w_A)\tilde{h}^{(+)}(w_B)$, where w_A contains all maximal clusters of w that overlap with $\mathrm{supp}(A)$ and w_B all other maximal clusters of w. The operators $\tilde{h}^{(+)}(w_A)$ and $\mathbb{1}\otimes\mathbb{1}\otimes B^{(-)} =: \hat{B}$, and $\tilde{h}^{(+)}(w_B)$ and $A^{(-)}\otimes\mathbb{1}\otimes\mathbb{1} =: \hat{A}$, then have disjoint supports, respectively, and the trace in Eq. (43) factorizes into a product of two traces, one over the subsystem $X := \mathrm{supp}(\hat{A})\cup\mathrm{supp}[\tilde{h}^{(+)}(w_A)]$ and the other over the rest of the system. It turns out that both vanish: By using the symmetries $\hat{A} = -\mathcal{S}^{1,2}\hat{A}\mathcal{S}^{1,2}$, $\tilde{h}^{(+)}(w_A) = \mathcal{S}^{1,2}\mathcal{S}^{3,4}\tilde{h}^{(+)}(w_A)\mathcal{S}^{3,4}\mathcal{S}^{1,2}$, $\hat{A}\mathcal{S}^{3,4} = \mathcal{S}^{3,4}\hat{A}$, and that $(\mathcal{S}^{i,j})^2 = \mathbb{1}$, one can show, e.g., that

$$\mathrm{Tr}[\mathcal{S}^{1,3}\mathcal{S}^{2,4}\hat{A}\tilde{h}^{(+)}(w_A)] = -\mathrm{Tr}[\mathcal{S}^{1,3}\mathcal{S}^{2,4}\hat{A}\tilde{h}^{(+)}(w_A)]. \tag{44}$$

This equation implies that for every $w \in \mathcal{C}_{\ge L}(\partial A)^c$,

$$t(w) \propto \mathrm{Tr}[\mathcal{S}^{1,3}\mathcal{S}^{2,4}\hat{A}\tilde{h}^{(+)}(w_A)] = 0. \tag{45}$$

Together with Eq. (42), realizing that $Z(\beta)^2 = Z[H^{(+)}](\beta)$, and using the notation from Eq. (37) with $F = \partial A$ and $L = d(A,B)$, it follows that

$$\mathrm{cov}^{\tau}_{g(\beta)}(A,B) = \mathrm{Tr}\left(\frac{\mathcal{S}^{1,3}\mathcal{S}^{2,4}\hat{A}\,\hat{B}}{2Z(\beta)^2}\Omega[\tilde{H}^{(+)}](\beta)\right). \tag{46}$$

After applying Lemma 1 and using that $\|\hat{A}\|_\infty \le 2\|A\|_\infty$, and similarly for B, we obtain

$$2\,\mathrm{cov}^{\tau}_{\rho}(A,B) = \begin{array}{|c|c|c|c|}\hline A\otimes\mathbb{1}-\mathbb{1}\otimes A & & B\otimes\mathbb{1}-\mathbb{1}\otimes B & \\ \hline g^{\tau} & g^{\tau} & g^{1-\tau} & g^{1-\tau}\\ \hline\end{array}$$

FIG. 6. The "multiple swap trick": Eq. (40) as a tensor network.

$$\frac{|\,\mathrm{cov}^{\tau}_{g(\beta)}(A,B)|}{\|A\|_\infty\|B\|_\infty} \le 2(e^{|\partial A|b(2\beta J)^L/[1-b(2\beta J)]}-1). \tag{47}$$

The fact that the condition $\beta<\beta^*$ is equivalent to $b(2\beta J)<1$ implies that $b(2\beta J)^L$ decays exponentially with L. In order to obtain the desired exponential bound (11), we apply the bound $\forall\, x\in[0,x_0]\colon \exp(x)-1 \le x(e^{x_0}-1)/x_0$ with the choice $x_0=\ln(3)$. In order to have $|\partial A|\frac{b(2\beta J)^L}{1-b(2\beta J)} \le \ln(3)$, we impose

$$L \ge \left|\ln\left(\frac{|\partial A|}{\ln(3)[1-b(2\beta J)]}\right)/\ln(2\beta J)\right| \tag{48}$$

$$= \xi(\beta)|\ln[\ln(3)(1-e^{-1/\xi(\beta)})/|\partial A|]| \tag{49}$$

$$=: L_0(\beta,|\partial A|). \tag{50}$$

This inequality guarantees the exponential bound (11) and finishes the proof. ∎

C. Proofs of the fermionic versions of the main results

In order to also establish our main results for fermionic systems, we go through the proofs for spin systems and discuss the necessary modifications.

Proof of Corollary 4.—In Theorem 1, we choose $H_0 = H - H_{\partial B}$. As the local terms are all in $\mathcal{G}$, we have that the thermal state of H_0 factorizes, i.e., $g_0 = g[H_{\restriction B}]g[H_{\restriction B^c}]$. After tracing over B^c, the statement follows. ∎

Proof of Theorem 4.—We use the same tensor copy trick as in the proof of Theorem 2. Equation (40) still holds in the fermionic setting. Note that the Hilbert space over which the trace is performed in Eq. (40) is not the Fock space of a system of 4 times the number of modes but the tensor product of four identical fermionic Fock spaces with the canonical inner product. This Hilbert space can be interpreted as that of a system of four types of fermionic particles that are each mutually indistinguishable and subject to (up to τ-dependent prefactors) identical Hamiltonians but do not interact with each other and can be distinguished from each other. It is spanned by tensor products of Fock states. The state $g[\tilde{H}^{(+)}](\beta)$ is the thermal state of this system. Equation (42) with $t(w)$ as defined as in Eq. (43) still holds. Note that the swap operators swap tensor factors, not fermionic modes. Thus, they still satisfy the symmetry relations that are

used to prove that only terms corresponding to the words $w \in \mathcal{C}_{\geq L}(\partial A)$ can contribute to the covariance.

It remains to show that Lemma 1 still holds in the fermionic setting. Lemmas 3 and 8 are purely combinatorial. Lemmas 4–7, 9, and 10 only use the local boundedness of the Hamiltonian and that Hamiltonian terms with disjoint support commute. The same holds in the fermionic setting because the Hamiltonian terms must be physical operators, i.e., even polynomials in the fermionic operators. Hence, all lemmas used in the proof of Lemma 1 carry over to the fermionic setting. It is then straightforward to see that the proof itself also goes through without any modifications. ∎

Proof of Corollary 5.—Tracing out B^c in the second trace in Eq. (32) and bounding the integral yields

$$|\mathrm{Tr}[Ag(\beta)] - \mathrm{Tr}\{Ag[H_{\restriction B}](\beta)\}| \leq |\beta| \sup_{s\in[0,1]} \sup_{\tau\in[0,1]} |\operatorname{cov}^{\tau}_{g_s(\beta)}(A, H_{\partial B})|. \quad (51)$$

Taking the supremum over all A with $\|A\|_\infty = 1$ and $\mathrm{supp}(A) \subseteq S$ and using Theorem 4 finish the proof. ∎

VI. CONCLUSIONS

In this work, we clarify the limitations of a universal concept of scale-independent temperature by showing that temperature is intensive on a given length scale if and only if correlations are negligible. The corresponding correlation measure turns out to also quantitatively capture the stability of thermal states against perturbations of the Hamiltonian. Moreover, we find a universal critical temperature above which correlations always decay exponentially with the distance. We compare our results to known results on phase transitions, comment on recent advances concerning thermalization in closed quantum systems (e.g., concerning the eigenstate-thermalization hypothesis), and discuss known matrix product operator approximations of thermal states. More concretely, our results imply that at high enough temperatures, the error made when truncating a Hamiltonian at some distance away from the system of interest is exponentially suppressed with the distance. As a computational consequence, expectation values of local observables can be approximated efficiently.

ACKNOWLEDGMENTS

We thank I. H. Kim, M. Holzäpfel, M. B. Hastings, B. Nachtergaele, M. Friesdorf, and A. Werner for helpful feedback and discussions. This work was supported by the Studienstiftung des Deutschen Volkes, the Alexander von Humboldt Stiftung, the EU (Q-Essence, SIQS, RAQUEL), and the ERC (TAQ).

APPENDIX: CLUSTER EXPANSIONS AND PROOF OF LEMMA 1

The following discussion of cluster expansions is expected to be interesting in its own right, as it contains a rigorous formulation of the ideas outlined in Ref. [40]. We will provide a proof of the original statement used to establish Theorem 3 as well as of the tailored statement in Lemma 1, which is used to prove Theorem 2 on the clustering of correlations.

1. The original cluster expansion from Ref. [40]

The original cluster expansion is similar to Lemma 1 with just one copy of the system instead of two weighted ones.

Lemma 2 (Truncated cluster expansion [40]).—Let $H = \sum_{\lambda\in E} h_\lambda$ be a local Hamiltonian with finite interaction graph (V, E) having growth constant α and local interaction strength $J = \max_{\lambda\in E}\|h_\lambda\|_\infty$, and define $b(x) := \alpha e^{|x|}(e^{|x|} - 1)$. Moreover, let β be small enough such that $b(\beta J) < 1$. For some subset of edges $F \subset E$, let $\mathcal{C}_{\geq L}(F) \subset E^*$ be the set of words containing at least one cluster c that contains at least one letter of F and has size $|c| \geq L$ and denote the corresponding truncated cluster expansion by

$$\Omega[H](\beta) := \sum_{w\in\mathcal{C}_{\geq L}(F)} \frac{(-\beta)^{|w|}}{|w|!} h(w). \quad (A1)$$

Then,

$$\frac{\|\Omega[H](\beta)\|_1}{Z(\beta)} \leq \exp\left(|F|\frac{b(\beta J)^L}{1 - b(\beta J)}\right) - 1. \quad (A2)$$

If one applies this lemma to the setting of Lemma 1, one obtains a bound similar as the one in Eq. (38) but with $Z[\tilde{H}](\beta)$ instead of $Z(\beta)$, where the ratio $Z[\tilde{H}](\beta)/Z[H](\beta)$ can be exponentially large in the system size for $\tau \in]0, 1[$.

Lemma 2 was used in Ref. [40] to establish a mathematically (not algorithmically) constructive version of Theorem 3, on MPO approximations, where the MPO in Eq. (17) is given by

$$\rho(\beta, L) = \frac{1}{Z(\beta)} \sum_{w\in E^*\setminus\mathcal{C}_{\geq L}(E)} \frac{(-\beta)^{|w|}}{|w|!} h(w). \quad (A3)$$

2. Proofs of Lemmas 1 and 2

The purpose of this section is to prove Lemma 1. But, along the way, we also prove Lemma 2. In order to do so, we start with the introduction of some more notation, mainly concerning clusters and lattice animals. For $w \in E^*$ and any subalphabet $G \subset E$, we write $G \subset w$ if every letter in G also occurs in w. By $G^c := E\setminus G$, we denote the complement of $G \subset E$. The extension of G is defined to be $\bar{G} := \{\lambda \in E |\ \exists \lambda' \in G : \lambda' \cap \lambda \neq \emptyset\}$ and, similarly as for subsystems, its boundary is $\partial G := \bar{G}\setminus G$. Throughout the proof, we fix some subset of edges $F \subset E$. We denote by $\mathcal{C}_{\geq L}(F) \subset E^*$ the set of words that contain at least one

cluster c with $c \cap F \neq \emptyset$ and $|c| \geq L$, and we denote by $\mathcal{C}^k_{\geq L}(F)$ the set of words that contain exactly k such clusters. Note that for an animal $G \subset E$, there exists a cluster $c \in E^*$ such that $G = \{\lambda \in c\}$, and if one imposes some order on G, one obtains a cluster. We denote by $\mathcal{A}_{=l}(F)$ and $\mathcal{A}_{\geq L}(F)$ the sets of animals that contain at least one edge of F and are of size exactly l or at least L, respectively. Moreover, we denote by $\mathcal{A}^k_{\geq L}(F)$ the corresponding sets of k-fold animals, i.e.,

$$\mathcal{A}^k_{\geq L} := \{\biguplus_{j=1}^k G_j : G_j \in \mathcal{A}_{\geq L}(F) \text{ nonoverlapping}\}.$$

For a more compact notation, we write the terms in the exponential series as

$$f(w) := \frac{(-\beta)^{|w|}}{|w|!} h(w). \tag{A4}$$

We will frequently use the following fact: For any Hamiltonian with a finite interaction graph (V, E), the partial series over any set of words $\mathcal{W} \subseteq E^*$ converges absolutely, i.e.,

$$\Big\| \sum_{w \in \mathcal{W}} f(w) \Big\|_\infty \leq \sum_{w \in \mathcal{W}} \frac{(|\beta| J)^{|w|}}{|w|!} \tag{A5}$$

$$\leq \sum_{w \in E^*} \frac{(|\beta| J)^{|w|}}{|w|!} \tag{A6}$$

$$= \exp(|\beta| J |E|). \tag{A7}$$

In particular, this bound implies that the order of the terms in the series over any subset of words $\mathcal{W}$ does not matter.

In the following proofs of Lemmas 1 and 2, we use several technical auxiliary lemmas, which we will only state and prove subsequently.

Proof of Lemma 1.—During this proof, we indicate quantities corresponding to $\tilde{H}$ by a tilde accent, e.g., $\tilde{f}(w)$ is defined as in Eq. (A4) but with respect to the local terms $\tilde{h}_\lambda$ of $\tilde{H}$ while $f(w)$ is defined with respect to the local terms h_λ of H.

We start the proof by rearranging the terms in the series over $\mathcal{C}_{\geq L}(F)$ in Eq. (37) according to the number of relevant clusters they contain and use Lemma 3 with b_k being the series over $\mathcal{C}^k_{\geq L}(F)$ to obtain

$$\Omega[\tilde{H}](\beta) = \sum_{k=1}^\infty \sum_{w \in \mathcal{C}^k_{\geq L}(F)} \tilde{f}(w)$$

$$= -\sum_{m=1}^\infty (-1)^m \sum_{k=m}^\infty \binom{k}{m} \sum_{w \in \mathcal{C}^k_{\geq L}(F)} \tilde{f}(w). \tag{A8}$$

Lemmas 5, 6, and 9 are the core of the proof. They define a series of operators $(\tilde{\rho}_m)_{m=1}^\infty$ that have a particularly useful form given in Lemma 10. This form exactly matches the series over k in Eq. (A8), which leads to the following identity:

$$\Omega[\tilde{H}](\beta) = -\sum_{m=1}^\infty (-1)^m \tilde{\rho}_m. \tag{A9}$$

The operators $\tilde{\rho}_m$ are defined in Eq. (A57) as series over m-fold lattice animals G of operators $\rho(G)$ [defined in Eq. (A31)]. This definition implies

$$\mathrm{Tr}(\mathcal{S} A B \tilde{\rho}_m) = \sum_{G \in \mathcal{A}^m_{\geq L}(F)} \mathrm{Tr}[\mathcal{S} A B \tilde{\rho}(G)]. \tag{A10}$$

In the previous steps, the series over words has been rewritten as a series over m-fold animals. Lemma 7 provides a bound on $\tilde{\rho}(G)$ that, together with Eqs. (A9) and (A10), yields

$$\frac{|\mathrm{Tr}\{\mathcal{S} A B \Omega[\tilde{H}](\beta)\}|}{\|A\|_\infty \|B\|_\infty Z(\beta)} \leq \sum_{m=1}^\infty \sum_{G \in \mathcal{A}^m_{\geq L}(F)} y(\beta J)^{|G|}. \tag{A11}$$

Now, a counting argument for lattice animals from Lemma 8 allows us to bound the series over m-fold animals G in terms of a series of animals

$$\frac{|\mathrm{Tr}\{\mathcal{S} A B \Omega[\tilde{H}](\beta)\}|}{\|A\|_\infty \|B\|_\infty Z(\beta)} \leq \sum_{m=1}^\infty \frac{1}{m!} \Big(\sum_{G \in \mathcal{A}_{\geq L}(F)} y(\beta J)^{|G|} \Big)^m.$$

Using that the number a_l [see Eq. (7)] of lattice animals G with $G \cap F \neq \emptyset$ and of size $|G| = l$ is bounded by $|F| a_l$ and that $a_l \leq \alpha^l$ [see Eq. (8)], we obtain

$$|\mathrm{Tr}\{\mathcal{S} A B \Omega[\tilde{H}](\beta)\}| \leq Z(\beta) \sum_{m=1}^\infty \frac{1}{m!} \Big(|F| \sum_{l=L}^\infty b(\beta J)^l \Big)^m$$

with $b(x) := \alpha y(x)$. Performing the partial geometric series over l with argument $b(\beta J) < 1$ and the exponential series over m yields Eq. (A2) and completes the proof. ∎

Similarly, we prove Lemma 2.

Proof of Lemma 2.—By the same argument that led us to Eq. (A9) in the proof of Lemma 1, we obtain

$$\Omega[H](\beta) = -\sum_{m=1}^\infty (-1)^m \rho_m. \tag{A12}$$

Applying the triangle inequality and using the bound on ρ_m from Lemma 9 yields

$$\|\Omega[H](\beta)\|_1 \leq Z(\beta) \sum_{m=1}^\infty \frac{1}{m!} \Big(|F| \sum_{l=L}^\infty b(\beta J)^l \Big)^m. \tag{A13}$$

Performing the partial geometric series over l with argument $b(\beta J) < 1$ and the exponential series over m yields Eq. (A2) and completes the proof. ∎

We now prove various lemmas that are used in the previous proofs of Lemmas 1 and 2.

Lemma 3.—Let $(b_k)_{k=1}^{\infty}$ be a sequence of complex matrices

$$A_K := \sum_{k=1}^{K} b_k \tag{A14}$$

and

$$B_K := -\sum_{m=1}^{K} (-1)^m \sum_{k=m}^{K} \binom{k}{m} b_k. \tag{A15}$$

Then, $A_K = B_K$ for all $K \in \mathbb{N}$. In particular, if both sequences converge, then their limits are the same, i.e., $\lim_{K\to\infty} A_K = \lim_{K\to\infty} B_K$.

Proof.—Applying the binomial theorem to $(1-1)^k = 0$ yields

$$\sum_{l=0}^{k} (-1)^l \binom{k}{l} = 0, \tag{A16}$$

which we will use. We prove the identity by induction. $A_1 = B_1$ is easy to see. Under the assumption that $A_K = B_K$ for some $K \in \mathbb{N}$, we obtain

$$B_{K+1} = B_K - (-1)^{K+1} \binom{K+1}{K+1} b_{K+1} \tag{A17}$$

$$-\sum_{m=1}^{K} (-1)^m \binom{K+1}{m} b_{K+1} \tag{A18}$$

$$= A_K + \left[-(-1)^{K+1} - \sum_{m=1}^{K} (-1)^m \binom{K+1}{m} \right] b_{K+1}$$

$$= A_{K+1}, \tag{A19}$$

where we have used Eq. (A16) in the last step. ∎

The goal of the following lemmas is to show that ρ_m is well-defined and to upper bound it in trace norm. The order of the lemmas is chosen in a way that makes clear that the two quantities ρ_m and $\rho(G)$, which will be defined shortly, are actually well-defined.

We start with a trace norm bound on the perturbed exponential series.

Lemma 4 [Eq. (21) from Ref. [40]].—Let H be a Hamiltonian with finite interaction graph (V, E). For any sequence $(G_j)_{j=1}^{k}$ of subalphabets $G_j \subset E$,

$$\left\| e^{-\beta(H-\sum_{j=1}^{k} H_{G_j})} \right\|_1 \leq Z(\beta) \prod_{j=1}^{k} \left\| e^{|\beta| H_{G_j}} \right\|_{\infty}. \tag{A20}$$

Proof.—The lemma is essentially a consequence of the Golden-Thompson inequality and the fact that the trace norm of a positive operator coincides with its trace. Using first the Golden-Thompson and then Hölder's inequality, we obtain

$$\begin{aligned} \left\| e^{-\beta(H-\sum_{j=1}^{k} H_{G_j})} \right\|_1 &\leq \mathrm{Tr}\left[e^{-\beta(H-\sum_{j=1}^{k-1} H_{G_j})} e^{\beta H_{G_k}} \right] \\ &\leq \mathrm{Tr}\left[e^{-\beta(H-\sum_{j=1}^{k-1} H_{G_j})} \right] \left\| e^{|\beta| H_{G_k}} \right\|_{\infty}. \end{aligned} \tag{A21}$$

Now, iteration completes the proof. ∎

We will use the following lemma to bound the operator norm of certain subseries of $f(w)$.

Lemma 5.—Let (V, E) be a finite graph and $J \geq 0$. For any $G \subset E$,

$$\sum_{w \in G^*: G \subset w} \frac{|\beta J|^{|w|}}{|w|!} = (e^{|\beta J|} - 1)^{|G|}. \tag{A22}$$

Proof.—Ordering the words in the sum in Eq. (A22) with respect to their length yields

$$\sum_{w \in G^*: G \subset w} \frac{|\beta J|^{|w|}}{|w|!} = \sum_{l=|G|}^{\infty} \sum_{w \in G^l: G \subset w} \frac{|\beta J|^{|w|}}{|w|!} \tag{A23}$$

$$= \sum_{l=|G|}^{\infty} \frac{|\beta J|^l}{l!} |\{w \in G^l : G \subset w\}|. \tag{A24}$$

From basic combinatorial considerations, we obtain

$$|\{w \in G^l : G \subset w\}| = \sum_{\substack{j_1, j_2, \ldots, j_n \geq 1, \\ j_1 + j_2 + \ldots + j_n = l}} \binom{l}{j}, \tag{A25}$$

where $\binom{l}{j}$ is a multinomial coefficient. Therefore, the right-hand side of Eq. (A24) only depends on $n := |G|$ and we denote it by

$$\gamma(n) := \sum_{l=n}^{\infty} \gamma(n, l) \tag{A26}$$

with

$$\gamma(n,l) \coloneqq \frac{|\beta J|^l}{l!} \sum_{\substack{j_1,j_2,\ldots,j_n \geq 1,\\ j_1+j_2+\ldots+j_n=l}} \binom{l}{j} = \sum_{\substack{j_1,j_2,\ldots,j_n \geq 1,\\ j_1+j_2+\ldots+j_n=l}} \frac{|\beta J|^{j_1}}{j_1!}\frac{|\beta J|^{j_2}}{j_2!}\cdots\frac{|\beta J|^{j_n}}{j_n!}. \quad \text{(A27)}$$

Then,

$$\begin{aligned}\gamma(n) &= \sum_{l=n}^{\infty} \sum_{\substack{j_1,j_2,\ldots,j_n \geq 1,\\ j_1+j_2+\ldots+j_n=l}} \frac{|\beta J|^{j_1}}{j_1!}\frac{|\beta J|^{j_2}}{j_2!}\cdots\frac{|\beta J|^{j_n}}{j_n!} \\ &= \sum_{l=n}^{\infty} \sum_{j_1=1}^{l-(n-1)} \frac{|\beta J|^{j_1}}{j_1!} \sum_{\substack{j_2,\ldots,j_n \geq 1,\\ j_2+\ldots+j_n=l-j_1}} \frac{|\beta J|^{j_2}}{j_2!}\cdots\frac{|\beta J|^{j_n}}{j_n!} \\ &= \sum_{l=1}^{\infty} \sum_{j_1=1}^{l} \frac{|\beta J|^{j_1}}{j_1!}\gamma(n-1, l+n-1-j_1) \quad \text{(A28)}\end{aligned}$$

and, after realizing that the last series is a Cauchy product,

$$\gamma(n) = \sum_{j_1=1}^{\infty} \frac{|\beta J|^{j_1}}{j_1!} \sum_{l=n-1}^{\infty} \gamma(n-1,l) \quad \text{(A29)}$$

$$= (e^{|\beta J|}-1)\gamma(n-1). \quad \text{(A30)}$$

We note that $\gamma(1) = e^{|\beta J|}-1$, and iteration finishes the proof. ∎

The following lemma provides a factorization of the series $\rho(G)$ in Eq. (A33) over words that have no letters on the boundary of an m-fold animal $G \in \mathcal{A}^m_{=l}(F)$ and contain all letters in G, into $\exp(-\beta H_{(\bar{G})^c})$, whose norm we have bounded in Lemma 4, times a product of operators $\eta(G_j)$. The $\eta(G_j)$ are supported on the single animals G_j composing the m-fold animal G. As we will see, a norm bound for $\eta(G_j)$ follows immediately from the previous lemma, which, in turn, also yields an upper bound on $\rho(G)$. The form of $\rho(G)$ given in Eq. (A33) together with this upper bound plays an important role in the main cluster expansion.

Lemma 6.—Let H be a Hamiltonian with finite interaction graph (V,E). For $G \subset E$, let $G = \biguplus_{j=1}^m G_j$ be the decomposition of G into nonoverlapping animals $G_j \subset E$ and define

$$\rho(G) \coloneqq e^{-\beta H_{(\bar{G})^c}} \prod_{j=1}^m \eta(G_j) \quad \text{(A31)}$$

with

$$\eta(G) \coloneqq \sum_{w \in \bar{G}^* : G \subset w} f(w). \quad \text{(A32)}$$

Then,

$$\rho(G) = \sum_{w \in [(\partial G)^c]^* : G \subset w} f(w). \quad \text{(A33)}$$

Proof.—To simplify the notation, we denote the relevant set of words that contain no letters in ∂G and each letter in G at least once by

$$\mathcal{W}^{\supset G} \coloneqq \{w \in [(\partial G)^c]^* : G \subset w\}. \quad \text{(A34)}$$

The idea is to group these words into subsets $[w] \subset \mathcal{W}^{\supset G}$ that coincide on the connected components of G and on $(\bar{G})^c$ and correspondingly split up the series (A33). We formalize this idea by introducing an equivalence relation on $\mathcal{W}^{\supset G}$. For $v, w \in \mathcal{W}$, we define

$$v \sim w : \Leftrightarrow \begin{cases} v\restriction G^c = w\restriction G^c \\ v\restriction G_j = w\restriction G_j \quad \forall j = 1,2,\ldots,k, \end{cases}$$

where, for any subalphabet $G' \subset E$, the restriction $w\restriction G'$ of a word $w \in E^*$ is obtained from w by omitting all letters that are not in G'. Then, the size of each equivalence class $[w] \in \mathcal{W}^{\supset G}/\sim$ is given by the multinomial coefficient

$$|[w]| = \binom{|w|}{(|w\restriction G^c|, |w\restriction G_1|, \ldots, |w\restriction G_k|)}. \quad \text{(A35)}$$

Note also that $h([w]) \coloneqq h(w\restriction \overline{G}^c)\prod_{j=1}^k h(w\restriction G_j) = h(w)$ is well-defined as a function on the classes. Let us denote the set of words over the alphabet G_j that contain all letters at least once by

$$\mathcal{W}^{=G_j} \coloneqq \{w \in (G_j)^* : G_j \subset w\}. \quad \text{(A36)}$$

Then, the quotient set can be identified with a Cartesian product of these sets

$$\mathcal{W}^{\supset G}/\sim \cong [(\bar{G})^c]^* \times \bigtimes_{j=1}^k \mathcal{W}^{=G_j}. \quad \text{(A37)}$$

For each equivalence class $K \in \mathcal{W}^{\supset G}/\sim$, we pick an arbitrary representative $w_K \in \mathcal{W}^{\supset G}$, use the definition of f in Eq. (A4), and determine that k is the number of connected components of G to obtain

$$\sum_{w\in[(\partial G)^c]^*:G\subset w} f(w) = \sum_{K\in\mathcal{W}^{\supset G}/\sim} |K|\frac{(-\beta)^{|w_K|}}{|w_K|!}h(w_K) \tag{A38}$$

$$= \sum_{v\in[(\tilde{G})^c]^*}\sum_{w_1\in\mathcal{W}^{=G_1}}\sum_{w_2\in\mathcal{W}^{=G_2}}\cdots\sum_{w_k\in\mathcal{W}^{=G_k}}\binom{|v|+\sum_{j=1}^k|w_j|}{(|v|,|w_1|,\ldots,|w_k|)}\frac{(-\beta)^{|v|+\sum_{j=1}^k|w_j|}}{(|v|+\sum_{j=1}^k|w_j|)!}h(v)\prod_{j=1}^k h(w_j)$$

$$= \sum_{v\in[(\tilde{G})^c]^*} f(v)\Big(\sum_{w_1\in\mathcal{W}^{=G_1}} f(w_1)\Big)\Big(\sum_{w_2\in\mathcal{W}^{=G_2}} f(w_2)\Big)\ldots\Big(\sum_{w_k\in\mathcal{W}^{=G_k}} f(w_k)\Big). \tag{A39}$$

Using the definition of η from Eq. (A32) on the last factors yields

$$\sum_{w\in[(\partial G)^c]^*:G\subset w} f(w) = e^{-\beta H_{(\tilde{G})^c}}\prod_{j=1}^k\eta(G_j) = \rho(G). \tag{A40}$$

■

The following lemma is a tighter variant of some of the original arguments leading to Lemma 2 for Hamiltonians consisting of two weighted copies of a local Hamiltonian. Its purpose is to provide a specialized tighter bound on $\rho(G)$, which turns out to be sufficient for our purposes. The central idea of the lemma is to expand $\rho(G)$ in the left-hand side of Eq. (A41) in order to be able to bound the trace using the generalized Hölder's inequality.

Lemma 7.—Let τ, H, $\tilde{H}$, A, and B be as in Lemma 1 and let $G\in\mathcal{A}_{\geq L}^{=m}(F)$ be an m-fold lattice animal with $G=\bigcup_{j=1}^m G_j$ and $G_j\in\mathcal{A}_{\geq L}(F)$. Moreover, let $\tilde{\rho}(G)$ be defined as $\rho(G)$ in Eq. (A31) but with respect to $\tilde{H}$. Then,

$$\frac{|\mathrm{Tr}[\mathcal{S}A^{(1)}B^{(2)}e^{-\beta\tilde{H}_{\tilde{G}^c}}\tilde{\rho}(G)]|}{\|A\|_\infty\|B\|_\infty Z(\beta)} \leq y(\beta J)^{|G|}, \tag{A41}$$

where $y(x) := e^{|x|}(e^{|x|}-1)$.

Proof.—Let us denote $k_\lambda^{(1)} := \tau h_\lambda^{(1)}$ and $k_\lambda^{(2)} := (1-\tau)h_\lambda^{(2)}$. For $w\in E^*$ and $v\in\{1,2\}^{|w|}$, we define $\tilde{h}(w,v) := k_{w_1}^{v_1}k_{w_2}^{v_2}\cdots k_{w_{|w|}}^{v_{|w|}}$. Then, by expanding the product $\tilde{h}(w)$, it can be written as

$$\tilde{h}(w) = \sum_{v\in\{1,2\}^{|w|}}\tilde{h}(w,v). \tag{A42}$$

Importantly, we can reorder the terms in $\tilde{h}(w,v)$ so that

$$\tilde{h}(w,v) = \tilde{h}^{(1)}(w,v)\tilde{h}^{(2)}(w,v), \tag{A43}$$

where $\tilde{h}^{(1)}(w,v) = h^{(i)}(w,v)\otimes\mathbb{1}$ and $\tilde{h}^{(2)}(w,v) = \mathbb{1}\otimes h^{(ii)}(w,v)$. Factorizing the operators and using the swap trick (39), we obtain

$$\mathrm{Tr}[\mathcal{S}A^{(1)}B^{(2)}e^{-\beta\tilde{H}_{\tilde{G}^c}}\tilde{h}(w,v)] = \mathrm{Tr}\{\mathcal{S}[Ae^{-\beta\tau H_{\tilde{G}^c}}h^{(i)}(w,v)]\otimes[Be^{-\beta(1-\tau)H_{\tilde{G}^c}}h^{(ii)}(w,v)]\} \tag{A44}$$

$$= \mathrm{Tr}\{[Ae^{-\beta\tau H_{\tilde{G}^c}}h^{(i)}(w,v)][Be^{-\beta(1-\tau)H_{\tilde{G}^c}}h^{(ii)}(w,v)]\}. \tag{A45}$$

Bounding the trace by the trace norm and applying Hölder's inequality generalized to several operators yields

$$|\mathrm{Tr}[\mathcal{S}A^{(1)}B^{(2)}e^{-\beta\tilde{H}_{\tilde{G}^c}}\tilde{h}(w,v)]| \leq \|A\|_\infty\|B\|_\infty\|e^{-\beta\tau H_{\tilde{G}^c}}\|_{1/\tau}\|e^{-\beta(1-\tau)H_{\tilde{G}^c}}\|_{1/(1-\tau)}\|h^{(i)}(w,v)\|_\infty\|h^{(ii)}(w,v)\|_\infty$$
$$\leq \|A\|_\infty\|B\|_\infty\|e^{-\beta H_{\tilde{G}^c}}\|_1 J^{|w|}\tau^{n^{(1)}(v)}(1-\tau)^{n^{(2)}(v)}, \tag{A46}$$

where in the second step, we have used that $\|X\|_p = \||X|^p\|_1^{1/p}$ and that with $n^{(j)}(v) := |\{v_k : v_k = j\}|$ for $j\in\{1,2\}$, the bounds $\|h^{(i)}(w,v)\|_\infty \leq (\tau J)^{n^{(1)}(v)}$ and $\|h^{(ii)}(w,v)\|_\infty \leq [(1-\tau)J]^{n^{(2)}(v)}$ hold. Now, we apply Lemma 4 and use that $\|e^{|\beta|H_{G_j}}\|_\infty \leq e^{|\beta|J|G_j|}$ to arrive at

$$|\mathrm{Tr}[\mathcal{S}A^{(1)}B^{(2)}e^{-\beta\tilde{H}_{\tilde{G}^c}}\tilde{h}(w,v)]| \leq \|A\|_\infty\|B\|_\infty Z(\beta)e^{|\beta|J|G|}J^{|w|}\tau^{n^{(1)}(v)}(1-\tau)^{n^{(2)}(v)}. \tag{A47}$$

From the definition of η in Eq. (A32), it follows that

$$\prod_{j=1}^{m}\tilde{\eta}(G_j)=\sum_{\left\{\substack{w^{(j)}\in G_j^*:\\ G_j\subset w^{(j)}}\right\}_{j=1}^{m}}\prod_{i=1}^{m}\frac{(-\beta)^{|w^{(i)}|}}{|w^{(i)}|!}\tilde{h}(w^{(i)}) \tag{A48}$$

$$=\sum_{\left\{\substack{w^{(j)}\in G_j^*:\\ G_j\subset w^{(j)}}\right\}_{j=1}^{m}}\frac{(-\beta)^{|w|}}{\prod_{i=1}^{m}|w^{(i)}|!}\sum_{v\in\{1,2\}^{|w|}}\tilde{h}(w,v), \tag{A49}$$

where $w\coloneqq w^{(1)}\circ w^{(2)}\circ\ldots\circ w^{(m)}$ and hence $\tilde{h}(w)=\prod_{i=1}^{m}h(w^{(i)})$. Together with the bound (A47) we obtain

$$\begin{aligned}&\frac{|\mathrm{Tr}[\mathcal{S}A^{(1)}B^{(2)}e^{-\beta\tilde{H}_{\bar{G}^c}}\prod_{j=1}^{m}\tilde{\eta}(G_j)]|}{\|A\|_\infty\|B\|_\infty Z(\beta)}\\ &\le e^{|\beta|J|G|}\sum_{\left\{\substack{w^{(j)}\in G_j^*:\\ G_j\subset w^{(j)}}\right\}_{j=1}^{m}}\frac{|\beta|^{|w|}}{\prod_{i=1}^{m}|w^{(i)}|!}J^{|w|}\\ &\quad\times\sum_{v\in\{1,2\}^{|w|}}\tau^{n^{(1)}(v)}(1-\tau)^{n^{(2)}(v)}.\end{aligned} \tag{A50}$$

Using the definition (A31) of $\rho(G)$ and the multinomial formula yields

$$\frac{|\mathrm{Tr}[\mathcal{S}A^{(1)}B^{(2)}e^{-\beta\tilde{H}_{\bar{G}^c}}\tilde{\rho}(G)]|}{\|A\|_\infty\|B\|_\infty Z(\beta)} \tag{A51}$$

$$=e^{|\beta|J|G|}\sum_{\left\{\substack{w^{(j)}\in G_j^*:\\ G_j\subset w^{(j)}}\right\}_{j=1}^{m}}\prod_{i=1}^{m}\frac{(|\beta|J)^{|w^{(i)}|}}{|w^{(i)}|!} \tag{A52}$$

$$=e^{|\beta|J|G|}\prod_{i=1}^{m}\left(\sum_{\substack{w^{(i)}\in G_i^*:\\ G_i\subset w^{(i)}}}\frac{(|\beta|J)^{|w^{(i)}|}}{|w^{(i)}|!}\right) \tag{A53}$$

$$\le e^{|\beta|J|G|}\prod_{i=1}^{m}(e^{|\beta|J}-1)^{|G_i|}, \tag{A54}$$

where in the second-to-last step, we have factorized the series and in the last step, we have used Lemma 5. ∎

We will need the following combinatorial lemma.

Lemma 8.—Let (V,E) be a finite (hyper)graph and $y\in[0,1[$. Then, for any $F\subset E$,

$$\sum_{G\in\mathcal{A}_{\ge L}^{m}(F)}y^{|G|}\le\frac{1}{m!}\left(\sum_{G\in\mathcal{A}_{\ge L}(F)}y^{|G|}\right)^{m}. \tag{A55}$$

Proof.—Remember that $\mathcal{A}_{\ge L}^{m}(F)$ is the set of m-fold (edge) animals of size at least L that contain a letter from F. For every $G\in\mathcal{A}_{\ge L}^{m}(F)$, one finds m pairs (G_1,G_2) with $G_1\in\mathcal{A}_{\ge L}^{m-1}(F)$ and $G_2\in\mathcal{A}_{\ge L}^{m}(F)$ such that $G=G_1\uplus G_2$; hence,

$$\begin{aligned}m\sum_{G\in\mathcal{A}_{\ge L}^{m}(F)}y^{|G|}&\le\sum_{G_1\in\mathcal{A}_{\ge L}^{m-1}(F)}\sum_{G_2\in\mathcal{A}_{\ge L}(F)}y^{|G_1|+|G_2|}\\ &=\left(\sum_{G\in\mathcal{A}_{\ge L}^{m-1}(F)}y^{|G|}\right)\left(\sum_{G\in\mathcal{A}_{\ge L}(F)}y^{|G|}\right).\end{aligned}$$

By iterating this inequality, we obtain

$$\sum_{G\in\mathcal{A}_{\ge L}^{m}(F)}y^{|G|}\le\frac{1}{m!}\left(\sum_{G\in\mathcal{A}_{\ge L}(F)}y^{|G|}\right)^{m}. \tag{A56}$$

∎

In the following lemma, we define a family of operators ρ_m and bound their trace norms. The bounds, in particular, guarantee that the ρ_m are well-defined. In addition, they are useful for the proof of Lemma 2, albeit they are not explicitly needed for the proof of Lemma 1.

Lemma 9.—Let $\rho(G)$ be defined as in Lemma 6 with respect to a Hamiltonian H having a finite interaction (hyper)graph (V,E) with growth constant α and let

$$\rho_m\coloneqq\sum_{G\in\mathcal{A}_{\ge L}^{m}(F)}\rho(G) \tag{A57}$$

for some $F\subset E$. Then,

$$\|\rho_m\|_1\le\frac{Z(\beta)}{m!}\left(|F|\sum_{l=L}^{\infty}b(\beta J)^l\right)^{m}, \tag{A58}$$

where $b(x)\coloneqq\alpha e^{|x|}(e^{|x|}-1)$.

Proof.—Consider a k-fold animal $G\in\mathcal{A}_{\ge L}^{m}(F)$ and decompose it into its k nonoverlapping animals $G_j\in\mathcal{A}_{\ge L}(F)$ as $G=\biguplus_{j=1}^{k}G_j\subset E$. Then, Eq. (A31) and Hölder's inequality imply

$$\|\rho(G)\|_1\le\|e^{-\beta H_{(\bar{G})^c}}\|_1\prod_{j=1}^{k}\|\eta(G_j)\|_\infty, \tag{A59}$$

and it follows from Lemmas 4 and 5 in conjunction with the definition of η in Eq. (A32) that

$$\|\rho(G)\|_1\le Z(\beta)y(\beta J)^{|G|}. \tag{A60}$$

Hence, by the definition from Eq. (A57) and Lemma 8, we obtain

$$\|\rho_m\|_1 \leq Z(\beta) \sum_{G \in \mathcal{A}^m_{\geq L}(F)} y(\beta J)^{|G|} \tag{A61}$$

$$\leq \frac{Z(\beta)}{m!} \left(\sum_{G \in \mathcal{A}_{\geq L}(F)} y(\beta J)^{|G|} \right)^m. \tag{A62}$$

By decomposing the set of animals of size at least L into a union of sets of animals of fixed size l, i.e., $\mathcal{A}_{\geq L}(F) = \biguplus_{l=L}^{\infty} \mathcal{A}_{=l}(F)$, we can write

$$\|\rho_m\|_1 \leq \frac{Z(\beta)}{m!} \left(\sum_{l=L}^{\infty} |\mathcal{A}_{=l}(F)| y(\beta J)^l \right)^m. \tag{A63}$$

The bound (8) on the number of lattice animals, the fact that the number $|F|$ of edges in F upper bounds the number of possibilities of translating an animal G such that $G \subset F$, and $b = \alpha y$ finish the proof. ∎

While the last lemma provides a bound on ρ_m and, in particular, implies that ρ_m is well-defined, the next lemma provides a useful form of ρ_m.

Lemma 10.—Let ρ_m be defined as in Eq. (A57). Then,

$$\rho_m = \sum_{k=m}^{\infty} \binom{k}{m} \sum_{w \in \mathcal{C}^k_{\geq L}(F)} f(w). \tag{A64}$$

Proof.—For $G \in \mathcal{A}^m_{\geq L}(F)$, let

$$\mathcal{W}(G) := \{w \in [(\partial G)^c]^* : G \subset w\}. \tag{A65}$$

According to Eqs. (A33) and (A57),

$$\rho_m = \sum_{G \in \mathcal{A}^m_{\geq L}(F)} \sum_{w \in [(\partial G)^c]^* : G \subset w} f(w). \tag{A66}$$

As

$$\bigcup_{G \in \mathcal{A}^m_{\geq L}(F)} \mathcal{W}(G) = \biguplus_{k=m}^{\infty} \mathcal{C}^k_{\geq L}(F), \tag{A67}$$

the sums in Eqs. (A57) and (A64) contain the same terms. It remains to show that the multiplicities are correct, i.e., are given by the binomial factor. Every word in $\mathcal{W}(G)$ contains at least m maximal clusters of size at least L, each of which contains a letter in F. The key is to decompose this set as

$$\mathcal{W}(G) = \biguplus_{k=m}^{\infty} \mathcal{W}^k(G) \tag{A68}$$

with

$$\mathcal{W}^k(G) := \{w \in \mathcal{W}(G) : \exists \text{ exactly } k \text{ maximal clusters } c \subset w : c \in \mathcal{C}_{\geq L}(F)\},$$

i.e., into sets of words having exactly $k \geq m$ such clusters. Then, the observation that for every $w \in \mathcal{W}^k(G)$ there are exactly $\binom{k}{m}$ many m-fold animals $G' \in \mathcal{A}^m_{\geq L}(F)$ with $w \subset G'$ completes the proof. ∎

[1] H. Pothier, S. Guéron, N. O. Birge, D. Esteve, and M. H. Devoret, *Energy Distribution Function of Quasiparticles in Mesoscopic Wires*, Phys. Rev. Lett. **79**, 3490 (1997).

[2] P. Peng, Z. Su, Z. Liu, Q. Yu, Z. Cheng, and J. Bao, *Nanowire Thermometers*, Nanoscale **5**, 9532 (2013).

[3] Y. Gao and Y. Bando, *Nanotechnology: Carbon Nanothermometer Containing Gallium*, Nature (London) **415**, 599 (2002).

[4] N. Linden, S. Popescu, and P. Skrzypczyk, *How Small Can Thermal Machines Be? The Smallest Possible Refrigerator*, Phys. Rev. Lett. **105**, 130401 (2010).

[5] A. Mari and J. Eisert, *Cooling by Heating*, Phys. Rev. Lett. **108**, 120602 (2012).

[6] M. Hartmann, G. Mahler, and O. Hess, *Existence of Temperature on the Nanoscale*, Phys. Rev. Lett. **93**, 080402 (2004).

[7] M. Hartmann and G. Mahler, *Measurable Consequences of the Local Breakdown of the Concept of Temperature*, Europhys. Lett. **70**, 579 (2005).

[8] M. Hartmann, *Minimal Length Scales for the Existence of Local Temperature*, Contemp. Phys. **47**, 89 (2006).

[9] A. Ferraro, A. Garcia-Saez, and A. Acin, *Intensive Temperature and Quantum Correlations for Refined Quantum Measurements*, Europhys. Lett. **98**, 10009 (2012).

[10] M. Vojta, *Quantum Phase Transitions*, Rep. Prog. Phys. **66**, 2069 (2003).

[11] D. Ruelle, *Statistical Mechanics: Rigorous Results* (Benjamin, New York, 1969).

[12] J. Ginibre, *Reduced Density Matrices of Quantum Gases. II. Cluster Property*, J. Math. Phys. (N.Y.) **6**, 252 (1965).

[13] W. Greenberg, *Critical Temperature Bounds of Quantum Lattice Gases*, Commun. Math. Phys. **13**, 335 (1969).

[14] O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics* (Springer, New York, 1981).

[15] Y. M. Park and H. J. Yoo, *Uniqueness and Clustering Properties of Gibbs States for Classical and Quantum Unbounded Spin Lattices*, J. Stat. Phys. **80**, 223 (1995).

[16] S. M. Bhattacharjee and A. Khare, *Fifty Years of the Exact Solution of the Two-Dimensional Ising Model by Onsager*, Curr. Sci. **69**, 816 (1995).

[17] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghi, *Canonical Typicality*, Phys. Rev. Lett. **96**, 050403 (2006).

[18] M. Cramer, C. M. Dawson, J. Eisert, and T. J. Osborne, *Exact Relaxation in a Class of Nonequilibrium Quantum Lattice Systems*, Phys. Rev. Lett. **100**, 030602 (2008).

[19] N. Linden, S. Popescu, A. J. Short, and A. Winter, *Quantum Mechanical Evolution towards Thermal Equilibrium*, Phys. Rev. E **79**, 061103 (2009).

[20] M. Rigol, V. Dunjko, and M. Olshanii, *Thermalization and Its Mechanism for Generic Isolated Quantum Systems*, Nature (London) **452**, 854 (2008).
[21] V. I. Yukalov, *Equilibration and Thermalization in Finite Quantum Systems*, Laser Phys. Lett. **8**, 485 (2011).
[22] P. Reimann and M. Kastner, *Equilibration of Isolated Macroscopic Quantum Systems*, New J. Phys. **14**, 043020 (2012).
[23] A. Riera, C. Gogolin, and J. Eisert, *Thermalization in Nature and on a Quantum Computer*, Phys. Rev. Lett. **108**, 080402 (2012).
[24] M. P. Mueller, E. Adlam, L. Masanes, and N. Wiebe, *Thermalization and Canonical Typicality in Translation-Invariant Quantum Lattice Systems*, arXiv:1312.7420.
[25] L. Masanes, A. J. Roncaglia, and A. Acín, *The Complexity of Energy Eigenstates as a Mechanism for Equilibration*, Phys. Rev. E **87**, 032137 (2013).
[26] M. Troyer, F. Alet, S. Trebst, and S. Wessel, *Non-local Updates for Quantum Monte Carlo Simulations*, AIP Conf. Proc. **690**, 156 (2003).
[27] M. B. Hastings and T. Koma, *Spectral Gap and Exponential Decay of Correlations*, Commun. Math. Phys. **265**, 781 (2006).
[28] B. Nachtergaele and R. Sims, *Lieb-Robinson Bounds and the Exponential Clustering Theorem*, Commun. Math. Phys. **265**, 119 (2006).
[29] B. Nachtergaele, Y. Ogata, and R. Sims, *Propagation of Correlations in Quantum Lattice Systems*, J. Stat. Phys. **124**, 1 (2006).
[30] Z. Landau, U. Vazirani, and T. Vidick, *A Polynomial-Time Algorithm for the Ground State of 1D Gapped Local Hamiltonians*, arXiv:1307.5143.
[31] Everything also works with more technical assumptions for certain cases where the local Hilbert spaces are not finite dimensional. But, in this case, the local terms of the Hamiltonian need to be bounded in operator norm by a constant. Hence, our results are not applicable to bosonic systems.
[32] Y. M. Miranda and G. Slade, *The Growth Constants of Lattice Trees and Lattice Animals in High Dimensions*, Electron. Commun. Probab. **16**, 129 (2011).
[33] M. Penrose, *Self-Avoiding Walks and Trees in Spread-Out Lattices*, J. Stat. Phys. **77**, 3 (1994).
[34] T. Barthel and M. Kliesch, *Quasi-locality and Efficient Simulation of Markovian Quantum Dynamics*, Phys. Rev. Lett. **108**, 230504 (2012).
[35] D. Poulin, *Lieb-Robinson Bound and Locality for General Markovian Quantum Dynamics*, Phys. Rev. Lett. **104**, 190401 (2010).
[36] B. Nachtergaele, A. Vershynina, and V. A. Zagrebnov, *Lieb-Robinson Bounds and Existence of the Thermodynamic Limit for a Class of Irreversible Quantum Dynamics*, AMS Contemporary Mathematics **552**, 161 (2011).
[37] M. B. Hastings, *Decay of Correlations in Fermi Systems at Non-zero Temperature*, Phys. Rev. Lett. **93**, 126402 (2004).
[38] N. D. Mermin and H. Wagner, *Absence of Ferromagnetism or Antiferromagnetism in One-or Two-Dimensional Isotropic Heisenberg Models*, Phys. Rev. Lett. **17**, 1133 (1966).
[39] J. Fröhlich and T. Spencer, *The Kosterlitz-Thouless Transition in Two-Dimensional Abelian Spin Systems and the Coulomb Gas*, Commun. Math. Phys. **81**, 527 (1981).
[40] M. B. Hastings, *Solving Gapped Hamiltonians Locally*, Phys. Rev. B **73**, 085115 (2006).
[41] T. S. Cubitt, A. Lucia, S. Michalakis, and D. Perez-Garcia, *Stability of Local Quantum Dissipative Systems*, arXiv:1303.4744.
[42] M. J. Kastoryano and J. Eisert, *Rapid Mixing Implies Exponential Decay of Correlations*, J. Math. Phys. (N.Y.) **54**, 102201 (2013).
[43] S. Michalakis and J. Pytel, *Stability of Frustration-Free Hamiltonians*, Commun. Math. Phys. **322**, 277 (2013).
[44] S. Bravyi, M. Hastings, and S. Michalakis, *Topological Quantum Order: Stability under Local Perturbations*, J. Math. Phys. (N.Y.) **51**, 093512 (2010).
[45] H. Li and F. D. M. Haldane, *Entanglement Spectrum as a Generalization of Entanglement Entropy: Identification of Topological Order in Non-Abelian Fractional Quantum Hall Effect States*, Phys. Rev. Lett. **101**, 010504 (2008).

2.4 On state space structures

We have seen (in Section 2.2.1 on pages 40f) that the set of states that can be effectively reached with local dynamics in polynomial time is exponentially small compared to the full set of states. This means that natural dynamics only exploits a tiny subset of state space. For ground states (see pages 22f) and thermal states [WVHC08] of local Hamiltonians, similar statements can be made in the form of so-called *area laws* [ECP10]. In the last section, we saw that at high temperatures correlations in thermal states cluster exponentially and that thermal states can be approximated by matrix product operators (see Theorems 2 and 3 on pages 50 and 52). This all justifies the picture that most of many-body physics takes place in a small subset of state space. Hence, if one wants to numerically simulate a system with restricted resources available, it seems to be a very good idea to also restrict the simulation to a small subset of state space, which contains the interesting physics. This is the basic idea of so-called *tensor network* methods. As the development and investigation of such methods have grown into a large field, there are several review articles available [ECP10, Eis13, Oru13, PGVWC07, Sch11, Sch13, VMC08] (see also pages 22ff). A *tensor network* is a partially contracted family of tensors (see, e.g., Figure 1 in the following Publication on page 101), where the dimensions of the single contracted indices are the *bond dimension*s. It can indeed be understood as a network, where the tensors correspond to the vertices, the indices to lines on the vertices, and contractions to connections of the corresponding lines (see pages 101f for figures and a more detail explanation). The basic idea of tensor network methods is to represent states and operators in terms of such networks, and to keep the tensors as low-dimensional as possible. This leads to several ansatz classes of states that serve as variational sets in the numerical simulations. As correlations are captured by the contraction lines, the topology of the network determines what kind of correlations it can capture. Hence, it is crucial to find the right network to solve a particular problem.

As an example, we introduce and compare two important classes of tensor network states in the following section. Then, in Section 2.4.2, we identify and discuss a fundamental problem arising when generalizing tensor network based methods from pure states to mixed state scenarios.

2.4.1 Real-space renormalization yields finite correlations

The most widely investigated example for tensor network states is given by *matrix product states (MPS)* that are one-dimensional tensor networks parameterizing state vectors of spin chains. In a way, they are pure state with only local correlations. Similarly, ground states of gapped local Hamiltonians have local corrections. More precisely, they obey an area law of the entanglement Rényi entropies [ECP10]. Hence, such ground states can be approximated by MPS faithfully [VC06]. If a Hamiltonian is not gapped, integer-spin chains can feature power

law violations of the area law [MS14]. However, in many situations its ground state still satisfies an area law only with logarithmic correlations [ECP10], meaning that it can have entanglement scaling logarithmically in the length of the chain. Being important in the study of critical phenomena [FIK08, Kor04, LRV04, VLRK03], this insight has lead to the idea of approximating ground states with hierarchical tensor network states, which are called *multiscale entanglement renormalization ansatz (MERA)* states [Vid07, Vid08] and which also feature a logarithmic scaling of entanglement. Correspondingly, MERA states have unbounded correlations and can be used as a variational class to simulate critical ground states efficiently [Vid07]. This method can even be used to study fermions [PBE10], which seems to be intractable using quantum Monte Carlo simulations due to the sign problem [TW05].

Of course, one would also like to simulate systems of dimensions larger than one. The two-dimensional analogue of MPS, called *projected entangled pair states (PEPS)*, are not efficiently contractible [SWVC07] and one needs to resolve to approximation schemes [LCB14]. MERA states, in contrast, are efficiently contractible in any dimension: In order to compute an expectation value of a local observable only tensors inside the *causal cone* need to be taken into account and the cone has a system size independent width [Vid07].

Quite surprising, the following holds (see Publication [BKE10] in Appendix B): In dimensions larger than one, MERA states are just PEPS with system-size independent tensor sizes. This indeed means that the hierarchical tensor network, such as generated in real-space renormalization schemes, can be mapped to a planer tensor network with bounded local dimensions. Hence, MERA states are an effectively contractible subclass of PEPS. As a corollary, we obtain that higher dimensional MERA states also satisfy the area law for the entanglement entropy (for an alternative argument leading to this statement, see the Preprint [Vid06] of Ref. [Vid08]).

2.4.2 A hard and an undecidable problem for translation invariant 1D systems

In any simulation one only needs to keep track of the relevant degrees of freedom. Consequently, numerical algorithms include a truncation step, where in the case of tensor network based algorithms, the bond dimensions are reduced. For the case of unitary dynamics in real and imaginary time, this truncation works particularly well in one dimension, making DMRG techniques [Whi92] a great success [ECP10, Eis13, Oru13, PGVWC07, Sch11, Sch13, VMC08]. In the case of Liouvillian dynamics (see Section 2.2), however, numerical simulations seem to be significantly less successful, even though some progress has been made [CB13, VGRC04, ZV04]. Similarly, as the simulation of imaginary time unitary dynamics can be used to approximate the ground state of a system by an MPS, one would like to approximate stationary states of local Liouvillians by MPOs via simulating the time evolution. One obstacle seems to be that known truncation schemes do not preserve positivity in that case and generate negative eigen-

values in the state approximation. The situation seems to be even more serious: Given an MPO approximation of a state it is not known how one can check whether or not the MPO is a state itself.

In the following Publication [KGE14a] we show that this is no coincidence by pointing out a fundamental obstacle: Testing whether or not an MPO is a positive semi-definite operator is NP-hard in the system size and even undecidable if no bound on the system size is assumed. This means that positivity is a genuinely global property without any local witness. Our results also hold if one allows for some fixed violation of the positivity by some threshold and we also formulate them in terms of a weak membership formulation of the problem (see pages 71f). We also discuss connections between different notions of one-dimensional tensor network states (see Table 1 on page 70 for an overview) and hidden Markov models (see Ref. [Vid11] for a survey) from classical probability theory as well as different types of positive matrix factorizations. Our undecidability result adds to a short list of undecidable problems in quantum mechanics [CPGW14, EMG12, WCPG11].

Matrix-Product Operators and States: NP-Hardness and Undecidability

M. Kliesch,[1] D. Gross,[2] and J. Eisert[1]
[1]QMIO Group, Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany
[2]Physikalisches Institut and Freiburg Center for Data Analysis and Modeling, Universität Freiburg, 79104 Freiburg, Germany
(Received 6 May 2014; revised manuscript received 12 July 2014; published 16 October 2014)

Tensor network states constitute an important variational set of quantum states for numerical studies of strongly correlated systems in condensed-matter physics, as well as in mathematical physics. This is specifically true for finitely correlated states or matrix-product operators, designed to capture mixed states of one-dimensional quantum systems. It is a well-known open problem to find an efficient algorithm that decides whether a given matrix-product operator actually represents a physical state that in particular has no negative eigenvalues. We address and answer this question by showing that the problem is provably undecidable in the thermodynamic limit and that the bounded version of the problem is NP-hard (nondeterministic-polynomial-time hard) in the system size. Furthermore, we discuss numerous connections between tensor network methods and (seemingly) different concepts treated before in the literature, such as hidden Markov models and tensor trains.

DOI: 10.1103/PhysRevLett.113.160503 PACS numbers: 03.67.−a, 02.60.Pn, 03.65.Ud, 89.70.Eg

Computational quantum many-body physics is marred by the fact that standard computational descriptions of states require exponentially many parameters. Fortunately, for many physically relevant problems, one does not need to consider all those parameters to capture natural properties extremely accurately. One of the pillars on which computational many-body approaches rest is the framework of tensor network methods. Here, the relevant degrees of freedom are parametrized by very few numbers, which are organized in terms of tensor networks that are contracted in order to compute expectation values [1–8]. Notably, the density-matrix renormalization group approach, the most successful method to numerically determine ground state properties of strongly correlated one-dimensional models, can be cast into such a form [1,2]. In this language, the problem of minimizing the energy can be phrased as a variational principle over matrix-product (or purely generated C^*-finitely correlated) states [9]. The natural analogue that also encompasses mixed quantum states is matrix-product operators. Again, they feature strongly in numerical algorithms [10,11], for example when investigating stationary states of local Liouvillians modeling open quantum systems [12,13] or Gibbs states [14,15].

However, general matrix-product operators are not guaranteed to represent physical states, which is the source of considerable conceptual and computational difficulties. It would thus be highly desirable to design an efficient algorithm capable of checking whether a given matrix-product representation defines a positive operator. To decide if such an efficient "local test for positivity" exists is a fundamental problem in the field, implicit already in its early formulations (see the Appendix of Ref. [9]).

Here, we address and answer this question: determining whether a matrix-product operator defines a physical state in the thermodynamic limit is a provably undecidable problem. We also show that the bounded version of the problem is nondeterministic-polynomial-time (NP) hard in the number of tensors, burying hopes that one could find an efficient algorithm testing for positivity exactly. This is proven for quantum spin chains with local dimension $d = 2$ by a polynomial reduction from the Post correspondence problem and a bounded variant thereof.

To give a practical example, one can approximate stationary states of local Liouvillians by iteratively applying the Liouvillian to a state described as a matrix-product operator and subsequently truncating the tensors. To avoid inconsistent results, one has to check whether the truncation step has caused the state to become "too unphysical" in that it has created eigenvalues that are more negative than some chosen tolerance threshold. We prove such a check to be unfeasible. The practical implications of our work are as follows. On the one hand, they motivate the quest for finding specific feasible instances that might exist. This quest is reminiscent of the task of finding, e.g., efficient contractions of two-dimensional planar tensor networks, even though this task has been identified to be #P-complete [16]. On the other hand, it shows that one should direct one's efforts towards finding approximate solutions.

The insight presented here adds a natural many-body problem to the list of quantum mechanical questions that have recently been identified not only as computationally hard, but as outright undecidable [17–20]. Along the way of introducing these novel results, we discuss a number of connections between concepts that have arisen in the literature, but whose relation has received surprisingly little attention (see Table I).

Tensor networks: In quantum many-body theory, tensor network methods are widely used in order to avoid

TABLE I. Concepts of tensor networks discussed here.

States	Classical	Quantum
Pure	Deterministic finite [55]	Matrix-product states [3], purely generated C^*-finitely correlated states [9], tensor trains [26]
Mixed, inherently positive	Hidden Markov models [21], probabilistic finite automaton [46]	C^*-finitely correlated states [9], local purification [29], quantum Markov chains [27]
Mixed, not inherently positive	Quasirealizations [21]	Finitely correlated states [9], matrix-product density operators [10,11]

intractability problems. The idea is to resort to variational classes of states, where the attention is restricted to low-dimensional manifolds of states that seem to capture well the relevant physics of the model under study. It is less widely appreciated in the physics community that similar structures are ubiquitous in classical probability theory: the hidden Markov model (HMM) is a generalization of a Markov chain, where the observable process $(Y_t)_{t\in\mathbb{N}}$ does not need to be Markovian but there is a stochastic process $(X_t)_{t\in\mathbb{N}}$ carrying additional information that renders the combined process (X_t, Y_t) Markovian. We only consider the case where X_t and Y_t have finitely many outcomes and call the number of outcomes of X_t the bond dimension D. With transition probabilities $M^{(\alpha)}_{i,j} = \Pr[(X_{t+1}, Y_{t+1}) = (j,\alpha)|X_t = i]$, boundary condition $p_j = \Pr[X_1 = j]$, and $\underline{1} \coloneqq (1,1,\ldots,1)^T \in \mathbb{R}^D$, the probabilities of outcome sequences of the process Y_t are given by the matrix product

$$\Pr[Y_1 = \alpha_1, \ldots, Y_n = \alpha_n] = pM^{(\alpha_1)} \cdots M^{(\alpha_n)}\underline{1}\,. \qquad (1)$$

In order for a HMM to describe a stationary process, p is usually taken to be a stationary distribution, i.e., $\sum_{\alpha=1}^{d} M^{(\alpha)} p = p$. The description complexity of the HMM is independent of n, or, if we allow the Markov kernels M to vary as a function of t, linear. Non-negativity of the probabilities in Eq. (1) is guaranteed because they arise as the contraction over elementwise non-negative tensors.

From Eq. (1), it follows that the matrices $F^{(k,n)}$ defined by

$$F^{(k,n)}_{(\alpha_1,\ldots,\alpha_k),(\alpha_{k+1},\ldots,\alpha_{k+n})} \coloneqq \Pr[Y_1 = \alpha_1, \ldots, Y_{k+n} = \alpha_{k+n}]$$

have rank at most D, which upper bounds the so-called Hankel rank [21].

A natural question is whether the rank condition alone characterizes those distributions that allow for a HMM with bond dimension D. It has been known since the 1960 s that this is not the case: there are distributions where $\operatorname{rank}(F^{(k,n)}) \leq D$ for all k, n, yet no HMM with finite bond dimension exists [22,23]. However, a relatively straightforward argument (based on sequential "rank-revealing decompositions," e.g., singular value decompositions) shows that every distribution with rank bounded by D allows for a representation as in Eq. (1) where the tensors M are not necessarily positive. This fact seems to have been discovered independently in different contexts, e.g., Refs. [7,9,24–26]. The resulting form is known as a quasirealization, offering the same concise description of the distribution as a HMM. These are, however, more difficult to work with computationally, as any variation of the local tensors can destroy global positivity. An important question thus is as follows: are the conditions on the tensor M that guarantee global positivity computationally efficiently verifiable? As we will see, the answer is no.

The above constructions generalize to the quantum setting: a C^*-finitely correlated state [9] (also known as quantum Markov chain [27,28]) ρ is obtained by replacing the elements of Eq. (1) by their quantum counterparts. We substitute p by a $D \times D$ density matrix σ, the stochastic map M by a completely positive map Φ that maps states on $\mathbb{C}^D$ to those on $\mathbb{C}^D \otimes \mathbb{C}^d$, and $\underline{1}$ by the partial trace (cf. Fig. 1). This immediately yields a local purification [29]: one can write Φ in Kraus representation by choosing operators $K_i : \mathbb{C}^D \to \mathbb{C}^D \otimes \mathbb{C}^d$ satisfying $\sum_{i=1}^{E} K_i K_i^\dagger = \mathbb{1}$ and $\Phi(\cdot) = \sum_{i=1}^{E} K_i^\dagger \cdot K_i$. Here, E can be assumed to be smaller than or equal to dD. Then the n-fold application of K to a purification $|\sqrt{\sigma}\rangle$ of σ yields the state vector $|\psi\rangle = K \circ K \circ \ldots \circ K |\sqrt{\sigma}\rangle$ in $\mathbb{C}^D \otimes (\mathbb{C}^d \otimes \mathbb{C}^E)^{\otimes n} \otimes \mathbb{C}^D$ that is a local purification of ρ, see Fig. 1. For quantum states one can, once more, define "quasirepresentations." Here Φ can be a general linear map and σ is some operator (no positivity constraints). This results in what is known as a matrix-product density operator (MPDO) or finitely correlated state (FCS) (not C^*-FCS). A discussion of different notions of positivity is provided in the Supplemental Material [30]. More concretely, with $[d] \coloneqq \{1, 2, \ldots, d\}$,

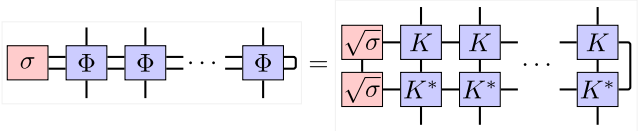

FIG. 1 (color online). A C^*-FCS as a tensor network. The channel Φ can be written in terms of Kraus operators, $\Phi(\rho) = \sum_i K_i^\dagger \rho K_i$. The vertically contracted indices between K and K^* correspond to the sum over i. The tensor network to the right is referred to as a local purification.

a MPDO is a density matrix written in the following form.

Definition 1 [matrix-product operator (MPO)].—An instance of MPO tensors is given by $M = (M_{i,j}^{(\alpha,\beta)})_{\alpha,\beta\in[d],i,j\in[D]} \in \mathbb{C}^{d\times d\times D\times D}$ and boundary vectors $|L\rangle, |R\rangle \in \mathbb{C}^D$. The dimension d is called the physical dimension and D the (MPO)-bond dimension. The generated translation invariant MPO for system size n is

$$\rho(L,M,R,n) := \sum_{j\in[D]^{n+1}} L_{j_1} M_{j_1,j_2} \otimes \ldots \otimes M_{j_n,j_{n+1}} R_{j_{n+1}}.$$

Main results: In order to precisely define the problems that are shown to be computationally unfeasible, we employ the standard language of theoretical computer science: the task of identifying objects with a certain property (e.g., those MPOs that are positive) is a decision problem. A specific case (e.g., given by a concrete tensor and boundary vectors) is an instance. A decision problem is NP-hard if it is at least as hard (in a precise sense) as all other problems from the complexity class NP. It is deemed highly unlikely that any NP-hard problem can be solved efficiently on either a classical or a quantum computer. A problem is (Turing) undecidable if no computer, even if endowed with unbounded resources, is capable of correctly solving all instances. In the statements of the various problems below, MPO tensors are specified by rational numbers. These have finite descriptions and can thus serve as inputs to computer programs. Allowing for more general numbers (e.g., complex rationals) would make the problem only harder.

In the precise statement of the problem, we allow for a threshold λ, which bounds the "degree of negativity" that is deemed acceptable. Moreover, we call positive semidefinite operators more concisely just positive.

Problem 2.—[bounded MPO threshold problem (BTP)] Instance: MPO tensors $M \in \mathbb{Q}^{d\times d\times D\times D}$, $|L\rangle, |R\rangle \in \mathbb{Q}^D$, threshold $\lambda \in \mathbb{Q}$, and system size n. Question: is the MPO $\rho(L,M,R,n) + \lambda\mathbb{1}$ positive?

Problem 3.—[MPO threshold problem (TP)] The TP is defined in the same way as the the BTP except that there is no restriction on the system size and the question is as follows: is there an $n \in \mathbb{Z}^+$ such that $\rho(L,M,R,n) + \lambda\mathbb{1}$ is not positive?

We obtain the following results, where the latter one adapts ideas from Ref. [46] to the quantum setting.

Theorem 4.—(NP-hardness of the bounded MPO threshold problem) For any $\lambda \in \mathbb{Q}$ and physical dimension $d \geq 2$, the BTP is NP-hard.

Theorem 5.—(undecidability of the MPO threshold problem) For each threshold $\lambda \in \mathbb{Q}$ the TP is undecidable. In particular, this holds for the case where the physical dimension is $d = 2$, the bond dimension is $D = 42$, and the matrices $M_{i,j}$ are diagonal for all $i, j = 1, \ldots, D$.

Outlook: An important question is whether there are physically relevant instances for which positivity is efficiently decidable and how this can be exploited best in numerical algorithms. Sometimes one can, e.g., efficiently detect negativity locally by calculating expectation values with respect to matrix-product states (MPS) of small bond dimension, see Fig. 2.

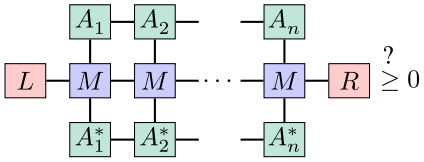

FIG. 2 (color online). Contracting MPOs with MPS can detect negativity for some instances. Hence, this provides a hierarchy of efficient tests labeled by the MPS-bond dimension, a strategy practically accessible by density-matrix renormalization group (DMRG) approaches.

In Ref. [29] local purifications of positive MPOs in terms of matrix-product states are investigated and it is shown that the arising MPS-bond dimension can in general not be bounded independently of the system size. This already suggests that such purifying MPS would require high bond dimensions when used instead of MPOs in numerical simulations. However, two constructive purification methods are suggested that are efficient when the rank of the MPO is polynomially bounded but in general necessarily inefficient [29]. From our Theorem 4 it also follows that this is no coincidence. To be more precise, a local purification method is an algorithm that receives a MPO instance $M, |L\rangle, |R\rangle$ and a system size n with $\rho(L,M,R,n) \geq 0$ as input and outputs a local purification of $\rho(L,M,R,n)$.

Corollary 6.—Local purification methods are inefficient in the system size.

In the BTP one is asked to exactly delineate the MPOs with smallest eigenvalues above $-\lambda$ from those with smallest eigenvalues below $-\lambda$. In practice, it would be acceptable if an algorithm reliably recognizes whether a state ρ is either sufficiently positive, i.e., $\rho \geq -\lambda$, or violates a threshold by at least $\epsilon \geq 0$, i.e., $\rho \not\geq -(\lambda+\epsilon)$. Such an approximate version is allowed to give unspecified results on the narrow band between the two cases. In order to make this precise, we state the BTP as a weak membership problem: for $\epsilon > 0$ one is only required to decide whether a MPO instance $(L,\ M,\ R)$ with $\mathrm{Tr}(\rho(L,M,R,n)) = 1$ satisfies either $\rho(L,M,R,n) \geq -\lambda$ or $\rho(L,M,R,n) \not\geq -(\lambda+\epsilon)$. The MPO provided in the proof of Theorem 4 has a trace that is exponentially bounded from above. Hence, as a corollary, one obtains that the BTP remains NP-hard as a weak membership problem if ϵ is exponentially small in n. This statement remains true for algebraic and not necessarily rational inputs. Weak membership formulations seem to be natural for a variety of problems in quantum information. For instance, NP-hardness of testing separability of quantum states as a weak membership problem was established first [47] for an exponentially small "error" ϵ and, much later [48], for a polynomially

small ϵ, in fact, using key methods of the previous approach [47]. Hence, our work is an invitation to explore whether the BTP as a weak membership problem is also NP-hard for only polynomially bounded ϵ or, instead, to actually find an algorithm that efficiently solves that problem.

Details: For any finite set Σ (alphabet) we denote by Σ^n the set of sequences w (words) of $n = |w|$ elements (letters) from Σ and by $\Sigma^* := \bigcup_{n\in\mathbb{N}}\Sigma^n$ the set of words. For $w \in [d]^*$ we denote by $|w\rangle = \bigotimes_{j=1}^{|w|} |w_j\rangle$ the tensor product of the corresponding canonical basis states $|1\rangle, \ldots, |d\rangle$. Given two monoids W and W' we call a map $U: W \to W'$ a morphism if U maps the identity element of W to the identity element of W' and $U(w_1 w_2) = U(w_1)U(w_2)$ for all $w_1, w_2 \in W$. The monoids we encounter here are either given by words over an alphabet with concatenation or by matrices with matrix multiplication. Next, we introduce the famous Post correspondence problem [49] and a bounded variant thereof.

Problem 7.—[bounded Post correspondence problem (BPCP)] Instance: pairs of words $(u_\alpha, v_\alpha)_{\alpha\in[d]}$ over a finite alphabet Σ and length n in unary notation [50]. Question: does there exist a nonempty word $w \in [d]^n$ of length n such that $u_{w_1} u_{w_2} \cdots u_{w_n} = v_{w_1} v_{w_2} \cdots v_{w_n}$?

Problem 8.—[Post correspondence problem (PCP)] The PCP is defined in the same way as the BPCP except that there is no restriction on the word length.

The two sets of words $(u_\alpha, v_\alpha)_{\alpha\in[d]}$, referred to as dominos, define two morphisms $U, V: [d]^* \to \Sigma^*$ given by $U(w) = u_{w_1} u_{w_2} \cdots u_{w_{|w|}}$ and similar for V.

Theorem 9.—(PCP is undecidable, see Ref. [51]) For every $d \geq 7$ the PCP with $\Sigma = \{0, 1\}$ is undecidable.

In fact, by noting a simpler proof [52] of a variant of this theorem with larger d, one can make the following computation theoretic statement.

Observation 10.—(BPCP is NP complete) There is a polynomial p such that for any nondeterministic Turing machine $\mathcal{M}$ and input x there is a reduction to an instance U, V of the BPCP such that $\mathcal{M}$ accepts x in n steps iff there is a solution of U, V of length $p(n)$.

In the usual textbook proof of the undecidability of the PCP (see, e.g., Ref. [52]) the halting problem is reduced to the PCP. The idea of the proof is to encode the computation history of a given Turing machine into the two morphisms of the PCP in two different ways such that there is a solution iff the Turing machine halts. Specifically, there is a polynomial p such that exactly when a Turing machine accepts an input after n steps then there is a solution of the corresponding PCP instance of length $p(n)$. The encoding works in the same way when the Turing machine is replaced by a nondeterministic Turing machine because having a transition relation instead of a transition function allows us to define the PCP instances in the same way (possibly with more "dominos" (u_α, v_α)). This shows that if one could solve the BPCP in polynomial time, then one could also solve NP problems in polynomial time. As solutions to the BPCP can be verified by a Turing machine in polynomial time the BPCP is in NP.

For the proofs of Theorems 4 and 5 the two following polynomial reductions are needed (see the Supplemental Material [30] for proofs building on Refs. [46,53,54]).

Lemma 11.—Let $(u_j, v_j)_{j\in[d]}$ be any instance of the PCP and $\lambda \in \mathbb{Q}$ be a threshold. Then there exist boundary vectors $|L\rangle, |R\rangle$, and matrices $A^{(1)}, \ldots, A^{(d)} \in \mathbb{N}^{6\times 6}$ that define a morphism $A: [d]^* \to \mathbb{N}^{6\times 6}$ such that for all $w \in [d]^*$

$$\begin{aligned} \langle L|A^{(w)}|R\rangle &= -(\lambda+1) \quad \text{if} \quad U(w) = V(w), \\ \langle L|A^{(w)}|R\rangle &\geq -\lambda \qquad\quad \text{if} \quad U(w) \neq V(w). \end{aligned} \tag{2}$$

Lemma 12.—(see Ref. [54]) Let $d, D \geq 2$, $A^{(1)}, \ldots, A^{(d)} \in \mathbb{Q}^{D\times D}$ be matrices that define a morphism $A: [d]^* \to \mathbb{Q}^{D\times D}$, and $|L\rangle, |R\rangle \in \mathbb{Q}^D$ be boundary vectors. Then there exist two matrices $B^{(1)}$ and $B^{(2)}$ that define a morphism $B: [2]^* \to \mathbb{Q}^{Dd\times Dd}$ together with an injective morphism $X: [d]^* \to [2]^*$ satisfying $|X(w)| = d|w|$ such that $\langle \tilde{L}|B^{(X(w))}|\tilde{R}\rangle = \langle L|A^{(w)}|R\rangle$ for all $w \in [d]^*$, where $|\tilde{X}\rangle := (|X\rangle, 0, \ldots, 0)^T$.

Proof of Theorems 4 and 5.—We prove the theorem by using the encoding $A: [d]^* \to \mathbb{Q}^{D\times D}$ of the PCP with d dominos into the matrices from Lemma 11. Using Lemma 12 we reduce the physical dimension d to 2 at the expense of having a larger bond dimension dD and an increase of the system size n to dn. This results in an encoding $C: [2]^* \to \mathbb{Q}^{dD\times dD}$ with boundary vectors $|\tilde{L}\rangle$ and $|\tilde{R}\rangle$. Now we define a MPO tensor M by $M^{(\alpha,\beta)} = \sum_{\gamma=1}^{d} \delta_{\alpha,\gamma}\delta_{\beta,\gamma} C^\gamma$. Then $(|L\rangle, M, |R\rangle)$ is an encoding of the PCP to the TP. All successively used encodings are polynomial reductions. In particular, an instance of the BPCP with word length n can be written as an instance of the BTP with system size dn. Hence, Theorem 9 and Observation 10 finish the proof.

Conclusions: In this work, we have shown that a problem naturally occurring in the context of tensor network states is NP-hard and in the thermodynamic limit even undecidable. The findings point to the challenge for reliable numerical methods for, e.g., finding Gibbs and stationary states of quantum many-body systems: truncations in the bond dimension—a common step in existing numerical algorithms—can introduce inconsistencies that cannot be found computationally. This insight provides an interesting twist to numerical methods to capture mixed quantum many-body systems as well as to notions of Hamiltonian complexity [56,57]. Future research should follow a dual aim: first, identify instances and approximations where (near) positivity can be guaranteed; second, search for further problems in the context of tensor network states that are not decidable algorithmically.

The work of M. K. and J. E. is supported by the Studienstiftung des Deutschen Volkes, the FQXi, the EU

(RAQUEL, SIQS, AQuS), and the ERC (TAQ). D. G. acknowledges support from the Excellence Initiative of the German Federal and State Governments (Grant No. ZUK 43), from the U.S. Army Research Office under Contracts No. W911NF-14-1-0098 and No. W911NF-14-1-0133, and from the Freiburg Research Innovation Fund.

[1] S. Rommer and S. Ostlund, Phys. Rev. B **55**, 2164 (1997).
[2] U. Schollwoeck, Ann. Phys. (N.Y.) **326**, 96 (2011).
[3] F. Verstraete, J. I. Cirac, and V. Murg, Adv. Phys. **57**, 143 (2008).
[4] N. Schuch, arXiv:1306.5551.
[5] R. Orus, Ann. Phys. (N.Y.) **349**, 117 (2014).
[6] J. Eisert, Model. Simul. **3**, 520 (2013).
[7] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Quantum Inf. Comput. **7**, 401 (2007).
[8] J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys. **82**, 277 (2010).
[9] M. Fannes, B. Nachtergaele, and R. F. Werner, Commun. Math. Phys. **144**, 443 (1992).
[10] F. Verstraete, J. J. Garcia-Ripoll, and J. I. Cirac, Phys. Rev. Lett. **93**, 207204 (2004).
[11] M. Zwolak and G. Vidal, Phys. Rev. Lett. **93**, 207205 (2004).
[12] M. C. Banuls, "Tensor Network Techniques for the Study of Dissipative Dynamics," 2013 (unpublished).
[13] Z. Cai and T. Barthel, Phys. Rev. Lett. **111**, 150403 (2013).
[14] T. Barthel, U. Schollwöck, and S. R. White, Phys. Rev. B **79**, 245101 (2009).
[15] E. M. Stoudenmire and S. R. White, New J. Phys. **12**, 055026 (2010).
[16] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **98**, 140506 (2007).
[17] J. Eisert, M. P. Muller, and C. Gogolin, Phys. Rev. Lett. **108**, 260501 (2012).
[18] M. M. Wolf, T. S. Cubitt, and D. Perez-Garcia, arXiv:1111.5425.
[19] J. Morton and J. Biamonte, Phys. Rev. A **86**, 030301(R) (2012).
[20] T. S. Cubitt, D. Perez-Garcia, and M. M. Wolf, in Proceedings of the XVII Conference on Quantum Information Processing, Barcelona, 2014 (unpublished).
[21] M. Vidyasagar, Math. Control Signals Syst. **23**, 1 (2011).
[22] M. Fox and H. Rubin, Ann. Math. Stat. **39**, 938 (1968).
[23] S. W. Dharmadhikari and M. G. Nadkarni, Ann. Math. Stat. **41**, 207 (1970).
[24] R. V. Erickson, Ann. Math. Stat. **41**, 843 (1970).
[25] G. Vidal, Phys. Rev. Lett. **93**, 040502 (2004).
[26] I. V. Oseledets, SIAM J. Sci. Comput. **33**, 2295 (2011); I. V. Oseledets and E. E. Tyrtyshnikov, SIAM J. Sci. Comput. **31**, 3744 (2009).
[27] M. Guta and J. Kiukas, arXiv:1402.3535.
[28] O. E. Barndorff-Nielsen, U. Franz, R. Gohm, B. Kümmerer, and S. Thorbjørnsen, *Quantum Independent Increment Processes II* (Springer, Heidelberg, 2006).
[29] G. De las Cuevas, N. Schuch, D. Pérez-García, and J. I. Cirac, New J. Phys. **15**, 123021 (2013).
[30] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.113.160503, which includes Refs. [31–45], for a discussion of different notions of positivity and proofs of Lemmas 11 and 12.
[31] S. A. Vavasis, SIAM J. Optim. **20**, 1364 (2010).
[32] D. D. Lee and H. S. Seung, Nature (London) **401**, 788 (1999).
[33] D. D. Lee and H. S. Seung, Adv. Neural Inf. Process. Syst. **13**, 556 (2000).
[34] S. Fiorini, S. Massar, S. Pokutta, H. R. Tiwary, and R. de Wolf, Proc. ACM STOC **44**, 95 (2012).
[35] M. Yannakakis, Proc. ACM STOC **20**, 223 (1988).
[36] J. Gouveia, P. A. Parrilo, and R. Thomas, Math. Oper. Res. **38**, 248 (2013).
[37] R. de Wolf, SIAM J. Comput. **32**, 681 (2003).
[38] C. Stark, arXiv:1210.1105.
[39] C. D. Aliprantis and R. Tourky, *Cones and Duality* (American Mathematical Society, Providence, 2007).
[40] S. Fiorini, S. Massar, M. K. Patra, and H. Raj Tiwary, arXiv:1310.4125.
[41] S. W. Dharmadhikari, Ann. Math. Stat. **34**, 705 (1963).
[42] J. Barrett, Phys. Rev. A **75**, 032304 (2007).
[43] H. Barnum, J. Barrett, M. Leifer, and A. Wilce, arXiv:quant-ph/0611295.
[44] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam, 1982).
[45] D. Gross, M. Muller, R. Colbeck, and O. C. O. Dahlsten, Phys. Rev. Lett. **104**, 080402 (2010).
[46] V. D. Blondel and V. Catarini, Theory Comput. Syst. **36**, 231 (2003).
[47] L. Gurvits, Proc. ACM STOC **35**, 10 (2003).
[48] S. Gharibian, Quantum Inf. Comput. **10**, 343 (2010).
[49] E. Post, Bull. Am. Math. Soc. **52**, 264 (1946).
[50] "Unary notation" effectively means that the problem size of the input n equals n.
[51] Y. Matiyasevicha and G. Sénizergues, Theor. Comput. Sci. **330**, 145 (2005).
[52] M. Sipser,"*Introduction to the Theory of Computation*", 2nd ed. (Thomson Course Technology, Boston, 2005).
[53] M. Hirvensalo, Lect. Notes Comput. Sci. **4362**, 309 (2007).
[54] V. D. Blondel and J. N. Tsitsiklis, Inf. Proc. Lett. **63**, 283 (1997).
[55] J. E. Hopcroft, R. Motwani, and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation* (Addison Wesley, Reading, MA, 2007).
[56] T. J. Osborne, arXiv:1106.5875.
[57] D. Aharonov, I. Arad, Z. Landau, and U. Vazirani, arXiv:1011.3445.

Supplemental Material: Matrix product operators and states – NP-hardness and undecidability

M. Kliesch,[1] D. Gross,[2] and J. Eisert[1]

[1]*QMIO group, Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*
[2]*Physikalisches Institut, Universität Freiburg, 79104 Freiburg, Germany*

APPENDIX

In the first parte of this appendix we discuss different kinds of postive matrix factorizations, positive ranks, and cones. In the last part we prove Lemmas 10 and 11 using ideas from Refs. [5, 7].

Matrix factorizations

In this section we first comment on the gap between the usual *rank* and the *non-negative rank* of a matrix. Then we briefly discuss another notion of positive rank, called *semidefinite rank*.

Non-negative rank. The problem of finding HMM realizations is closely connected with the theory of *non-negative matrix factorizations*. Any entry-wise non-negative matrix F, e.g., given by $F_{\alpha,\beta} = \Pr[Y_1 = \alpha, Y_2 = \beta]$, can be written as

$$F = \sum_{j=1}^{D} |L_i\rangle\langle R_i| \,, \tag{6}$$

i.e., as a sum of rank-1 matrices. This is a quasi-realization of the process (Y_1, Y_2) with two time steps. The minimal D is the usual rank of F. Similarly, the *positive rank* of F is the minimum number of entry-wise non-negative rank-1 matrices that sum up to F. Any such decomposition into rank-1 matrices with the proper normalization is a HMM for (Y_1, Y_2).

We thus find that the auxiliary dimension D of a quasi-realization of two discrete random variables is nothing but the rank of the joint probability distribution seen as a matrix; while the smallest D in a HMM is the non-negative rank. Old results [2, 3] showing that minimal bond dimensions for realizations and quasi-realizations are distinct can thus be reinterpreted in terms of gaps between rank and non-negative rank. Actually finding the non-negative rank is known to be NP-hard [8].

Non-negative matrix factorizations have been studied extensively, just to name two famous examples, in Ref. [9] for influential applications in machine learning theory, and in Ref. [10] for algorithms.

Positive semidefinite rank. If A_α, B_β are non-negative matrices, then $\mathrm{Tr}(A_\alpha B_\beta)$ is clearly a non-negative number. One defines a *positive semidefinite (PSD) factorization* of a matrix F to be a choice of positive semidefinite matrices $A_\alpha \geq 0, B_\beta \geq 0$ such that

$$F_{\alpha,\beta} = \mathrm{Tr}(A_\alpha B_\beta). \tag{7}$$

The *PSD rank* of F is the smallest dimension D such that there is a PSD factorization of F in terms of $D \times D$-matrices. These notions have recently attracted considerable attention [11–13] as novel lower-bounds on the PSD rank (partly derived in the study of quantum communication complexity [14]) have been used to disprove the existence of efficient semidefinite programs that would solve certain NP-hard combinatorial optimization problems [11]. The problem of identifying semidefinite factorizations for empirically observed statistics of quantum experiments has been treated in Ref. [15] and prior works by the same author referenced there.

Bond dimensions of MPOs and local purifications

As in the previous case, one can easily check that a PSD factorization for F immediately gives a realization of the *classical* bivariate distribution defined by F in terms of a *quantum* HMM, i.e., a local purification. Of course, classical distributions can be embedded into quantum states. This was used in Ref. [4] to leverage the fact [11] that there are families of Fs whose rank is bounded, but whose PSD rank diverges to show that there are bi-partite quantum states with bounded bond dimensions as MPOs, but unbounded ones for any local purification.

Cones and generalizations

One can abstract even further. Element-wise positivity and semidefiniteness are two notions of positivity in vector spaces. There is a systematic theory of general ordered vector spaces [16]. Notions of order in real vector spaces stands in one-one correspondence to (Archimedean) closed convex cones. The cone represents those elements of the vector space that are deemed "non-negative". For two vectors $v, w \in V$, one says $v \leq w$ iff $v - w$ is an element of that cone. One can now define generalized HMMs by the "physical" and the "auxiliary" vector spaces carrying various orders and by requiring that the Fs are positive maps in the sense that they preserve these orders. We can then distinguish various cases. If both physical and classical spaces carry the usual element-wise order, we recover the traditional notion of HMM. If both carry the semidefinite order, we obtain quantum Markov processes. If the physical one is element-wise, but the auxiliary one semidefinite, one arrives at the model treated in Ref. [4] and also in this work.

Such *conal matrix factorizations* have recently been discussed in the optimization literature [11, 13, 17]. Notably, researchers working on stochastic process both in a classical [18] and quantum [1, Appendix] setting have appreciated these ideas already quite early. They are also inherent to the *generalized probabilistic theories* approach to quantum foundations [19–22].

Proof of Lemmas 10 and 11

Proof of Lemma 10. For an alphabet Σ of size $b = |\Sigma|$ the *numeric representation* of Σ^* is the map $\sigma : \Sigma^* \to \mathbb{N}$ given by

$$\sigma(w) \coloneqq \sum_{j=1}^{|w|} \sigma(w_j) b^{|w|-j} \tag{8}$$

where $\sigma(\emptyset) = 0$ and $\sigma\restriction_\Sigma : \Sigma \to [b]$ is a bijection (enumeration of Σ). Note that for any two words $u, v \in \Sigma^*$

$$\sigma(uv) = b^{|v|}\sigma(u) + \sigma(v) \tag{9}$$

and $\sigma(u) = \sigma(v)$ iff $u = v$. Note that for all w

$$b^{|w|-1} \leq \sigma(w) \leq b^{|w|} - 1. \tag{10}$$

Next, we define $A : \Sigma^* \times \Sigma^* \to \mathbb{N}^{6\times 6}$ as in Refs. [5, 6] by

$$A(u,v) = \begin{pmatrix} b^{2|u|} & 0 & 0 & 0 & 0 & 0 \\ 0 & b^{|u|+|v|} & 0 & 0 & 0 & 0 \\ 0 & 0 & b^{2|v|} & 0 & 0 & 0 \\ \sigma(u)b^{|u|} & \sigma(v)b^{|u|} & 0 & b^{|u|} & 0 & 0 \\ 0 & \sigma(u)b^{|v|} & \sigma(v)b^{|v|} & 0 & b^{|v|} & 0 \\ \sigma(u)^2 & 2\sigma(u)\,\sigma(v) & \sigma(v)^2 & 2\sigma(u) & 2\sigma(v) & 1 \end{pmatrix}.$$

Using the property (9) of the numeric representation of Σ^* it follows that

$$A(u_1, v_1)\, A(u_2, v_2) = A(u_1 u_2, v_1 v_2) \tag{11}$$

for all words $u_1, u_2, v_2, v_2 \in \Sigma^*$. Now let $U, V : [d] \to \Sigma$ be an instance of the PCP and define $A^{(\alpha)} \coloneqq A(U(\alpha), V(\alpha))$ for all $\alpha \in [d]^*$ and the boundary vectors as $|L\rangle = |6\rangle$ and $|R\rangle = |1\rangle - |2\rangle + |3\rangle - (\lambda+1)\,|6\rangle$. Thanks to Eq. (11), A viewed as a map on $[d]$ extends to a morphism on $[d]^*$, as required. For any $w \in [d]^*$ we have $U(w) = V(w)$ iff $A^{(w)}_{6,4} = A^{(w)}_{6,5}$. Moreover,

$$\langle L|\, A^{(w)}\, |R\rangle = \big(\sigma(U(w)) - \sigma(V(w))\big)^2 - (\lambda+1)\,, \tag{12}$$

and Eq. (4) follows. □

Proof of Lemma 11. We using a technique from Ref. [7]. Define $B^{(1)} \coloneqq \operatorname{diag}\left(A^{(\alpha)}\right)_{\alpha\in[d]}$ to be the block diagonal matrix having $A^{(\alpha)}$ in increasing order as the diagonal blocks and

$$B^{(2)} \coloneqq \begin{pmatrix} 0 & \mathbb{1}_{D(d-1)} \\ \mathbb{1}_D & 0 \end{pmatrix}. \tag{13}$$

Now we construct a bijective morphism $X : [d]^* \to [2]^*$ satisfying

$$\big\langle \tilde{L}\big|B^{(X(w))}\big|\tilde{R}\big\rangle = \langle L|\, A^{(w)}\, |R\rangle\,. \tag{14}$$

First, note that $B^{(2)} = S \otimes \mathbb{1}_D$, where S is the permutation matrix representing the cyclic permutation $(1, 2, \ldots, d) \mapsto (2, 3, \ldots, d, 1)$. This means that $B^{(2)}$ acts as the cyclic left shift on the d blocks of D neighbouring components of column vectors. Hence

$$C^{(\alpha)} \coloneqq B^{(2)^{\alpha-1}}\, B^{(1)}\, B^{(2)^{d-(\alpha-1)}} \tag{15}$$

$$= \operatorname{diag}(A^{(\alpha)}, \ldots, A^{(d)}, A^{(1)}, \ldots, A^{(\alpha-1)}) \tag{16}$$

for all $\alpha \in [d]$. Eq. (16) implies that for any $w \in [d]^*$ the upper left $D \times D$ block of $C^{(w)}$ is $A^{(w)}$. Next, define $\big|\tilde{L}\big\rangle, \big|\tilde{R}\big\rangle \in \mathbb{Z}^{dD}$ to be the vectors that have $|L\rangle$ and $|R\rangle$ as their first D components respectively and all other components equal to zero. Then we obtain

$$\big\langle \tilde{L}\big|C^{(w)}\big|\tilde{R}\big\rangle = \langle L|\, A^{(w)}\, |R\rangle \tag{17}$$

for all $w \in [d]^*$.

For $s \in \mathbb{N}$ and $\alpha \in [d]$ we denote by α^s the word that consists of s times the letter α. Now we define a morphism $X : [d]^* \to [2]^*$ by

$$X(\alpha) \coloneqq 2^{\alpha-1}1\,2^{d-\alpha} \tag{18}$$

for $\alpha \in [d]$. Note that for any $w \in [d]^*$ we have $|X(w)| = d|w|$ and that one can reconstruct w from just knowing $X(w)$, i.e., X is injective. Using Eq. (16) we obtain

$$C^{(w)} = B^{(X(w))}. \tag{19}$$

Together with Eq. (17) this finishes the proof. □

[1] M. Fannes, B. Nachtergaele, and R. F. Werner, Comm. Math. Phys. **144**, 429 (1992).
[2] M. Fox and H. Rubin, Ann. Math. Stat. **39**, 938 (1968).
[3] S. W. Dharmadhikari and M. G. Nadkarni, Ann. Math. Stat. **41**, 207 (1970).
[4] G. De las Cuevas, N. Schuch, D. Pérez-García, and J. I. Cirac, New J. Phys. **15**, 123021 (2013).
[5] V. D. Blondel and V. Catarini, Theory Comput. Syst. **36**, 231 (2003).
[6] M. Hirvensalo, LNCS **4362**, 309 (2007).
[7] V. D. Blondel and J. N. Tsitsiklis, Inform. Process. Lett. **63**, 283 (1997).
[8] S. A. Vavasis, SIAM J. Optim. **20**, 1364 (2009).
[9] D. D. Lee and H. S. Seung, Nature **401**, 788 (1999).
[10] D. D. Lee and H. S. Seung, Adv. Neural Inform. Process. Systems **13**, 556 (2000).
[11] S. Fiorini, S. Massar, S. Pokutta, H. R. Tiwary, and R. de Wolf, Proc. ACM STOC **44**, 95 (2012); arXiv:1111.0837.
[12] M. Yannakakis, Proc. ACM STOC **20**, 223 (1988).

[13] J. Gouveia, P. A. Parrilo, and R. Thomas, Math. Oper. Res. **38**, 248 (2013); arXiv:1111.3164.
[14] R. de Wolf, SIAM J. Comput. **32**, 681 (2003).
[15] C. Stark, arXiv:1210:1105.
[16] C. D. Aliprantis and R. Tourky, *Cones and Duality* (American Mathematical Society, Providence, 2007).
[17] S. Fiorini, S. Massar, M. K. Patra, and H. Raj Tiwary, arXiv:1310.4125.
[18] S. W. Dharmadhikari, Ann. Math. Statist. **34**, 705 (1963).
[19] J. Barrett, Phys. Rev. A **75**, 032304 (2007).
[20] H. Barnum, J. Barrett, M. Leifer, and A. Wilce, quant-ph/0611295.
[21] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam, 1982).
[22] D. Gross, M. P. Muller, R. Colbeck, and O. C. O. Dahlsten, Phys. Rev. Lett. **104**, 080402 (2010).

3 Quantum simulations and the verification problem

In the last chapter, we discussed several results relevant in the broader context of simulations of complex quantum systems. More precisely, we have discussed simulations on both, on classical computers and on quantum computers. But in fact, there is also a notion of so-called *quantum simulations* that is weaker than a scalable universal quantum computer but can solve at least some task more efficiently than classical computers. Since such simulations started recently to receive a lot of attention, I would like to start the conclusion of this thesis with the discussion of a fundamental problem that arises once one starts solving problems other than those in NP, i.e., problems other than those solutions of which can be verified efficiently on a classical computer. In order to explain this, it is worth having a look at the intractability of decision problems in NP first. For instance, deciding whether or not a natural number can be non-trivially factorized is believed to be intractable on classical computers. If someone had a machine factoring numbers efficiently, he clearly could solve this decision problem. But even more, he could easily convince everybody of his machine's power, since using a classical computer one can efficiently check whether or not the factorizations are indeed correct. Coming back to quantum simulations, in general there is no such efficient check that decides whether a quantum simulation actually does what it is intended to do. This means that someone operating a quantum simulation with supra-classical powers can not automatically convince a sceptic second party that he actually does have those powers. We call this the *verification problem*. A detailed discussion of this problem has strangely been absent in the quantum simulations literature so far.

While building a fully fledged universal quantum computer is the ultimate goal, realizing some sort of computation that is intractable on classical computers would constitute a great breakthrough [CZ12, Pre12]. The hope is to achieve this breakthrough using quantum resources, i.e., in a quantum simulation. In fact, there are already a number of works tackling this more modest problem with photonic quantum systems [AGW12], ultracold gases in optical lattices [BDN12], trapped ions [BR12], superconducting circuits [HTK12]. Also annealing processes with quantum signatures for finding certain ground states have been realized experimentally [Boi+13, Boi+14]. As their functioning cannot be verified easily, they were criticized for apparent "classical signatures" [SS13, Wan+13]. Another famous example for quantum simulations is given by the Boson-Sampling problem [AA11], which was suggested solely to

show that quantum systems can outperform classical ones in performing some particular task. Therefore, Boson-Sampling is ideal for discussing the verification problem.

3.1 Indication that classical efficient Boson-Sampling verification is impossible

The Boson-Sampling problem is to approximately sample from the output distribution of a linear optical network with n bosons in m modes. A single sample is the vector of photon numbers obtained when the output state of the network is measured in the Fock basis (see Figure 1 in Publication [GKAE13] in Appendix B). If the passive Gaussian unitary describing the network is drawn uniformly at random then this problem is proven [AA11] to be hard with high probability for m scaling at least as n^5 and under reasonable conjectures. Such Boson-Sampling simulations have even been implemented experimentally with up to three bosons in six optical modes [Bro+13, Spr+13, Til+13, Cre+13] and the theoretical predictions of the theory of passive linear optics have been confirmed. Of course, one would like to conclude that quantum systems provably outperform classical computers (in performing some task). For this conclusion, it remains to show that Boson-Sampling simulations can be scaled efficiently in the number of bosons and the inverse error [AA11].

But more fundamentally, the verification problem needs to be resolved, at least in some way. Of course, one can always gain some evidence that a quantum simulation works correctly by testing it in a classically tractable regime [Bra+14, Tro+12]. Then it is a matter of taste whether one trusts the simulation also in the classically intractable regime. However, it remains open if and how this problem can be overcome in a more satisfactory manner. For Boson-Sampling, the situation is particularly dramatic as its sole purpose is to provide an example for a classical intractable quantum simulation. Hence, it is desirable that it performs particularly well in this task, i.e., that one can convince even a sceptic person of its power. In this sense, for Boson-Sampling a solution of the verification problem is particularly important. A good intuition and also rigorous evidence that Boson-Sampling cannot be verified effectively with only classical resources is given in Publication [GKAE13], which we will discuss in the next section. As we will discuss in Section 3.2, first steps were already made towards verification of Boson-Sampling using very restricted quantum resources. Whether or not full verification of Boson-Sampling in an efficient manner is possible remains, however, an open problem.

For verification one ideally would like to test whether a potential Boson-Sampling device either samples from a distribution that is ϵ-close to the ideal Boson-Sampling distribution or not. But since boundary cases can always be arbitrary hard to distinguish, the best one can hope for is to verify that the distributions have a distance either less than ϵ or larger than $\epsilon + \Delta$, where $\Delta > 0$ can be reduced efficiently. But actually a test that verifies the device if

it samples from the ideal Boson-Sampling distribution and rejects if the device samples from a distribution that is more than ϵ far away from it could be sufficient for verification (see also Publication [AGKE14] for detailed discussion of notions of verification). At this point it is worth noting that the sample space is larger than exponentially large in the number of bosons n and with similar discretization arguments as in Publication [Kli+11b] on pages 40f, it follows that the space of distributions is larger than double exponentially large in n.

However, as a first step, we will discuss Boson-Sampling in a restricted setting, where sample complexity results [Bat+13] can be applied. More specifically, we ask the following: How many samples are needed to distinguish the Boson-Sampling distribution from the uniform one? One important ingredient to the hardness of Boson-Sampling is that not any probability of a single outcome can be calculated efficiently (by a conjecture the hardness of Boson-Sampling relies on) from the Gaussian unitary describing the linear optical network. Hence, it is reasonable to first find an answer to that problem, where one is not allowed to use the unitary. This essentially restricts the verification algorithm to be a so-called *symmetric algorithm* [Bat+13] and one can show that exponentially many samples are needed in that case (see Publication [GKAE13] in Appendix B). The intuitive argument is that the Boson-Sampling distribution is typically very flat and, therefore, the size of the sample space makes it hard to detect its structure from few samples. Importantly, the structure is encoded in polynomially many samples which is implied by Theorem 3 on page 124 of Appendix B. The main technical contribution of this theorem is to show that the Boson-Sampling distribution is flat, i.e., to lower bound the min-entropy of typical Boson-Sampling distributions.

Inspired by this work [GKAE13], algorithms were found that distinguish Boson-Sampling from uniform sampling using only polynomially many samples, where the key is to actually use the information about the Gaussian unitary in a smart way [AA11]. On the other hand, an observation from F. G. S. L. Brandão is the following [AA11]: For every instance of Boson-Sampling with high min-entropy and every circuit length $T \in O(\mathsf{poly}(n))$ there is a classically efficiently samplable distribution indistinguishable from the distribution of the Boson-Sampling instance by all circuits of length T. Hence, efficient classical Boson-Sampling verification, in the sense explained above, is not possible.

3.2 Reliable quantum verification for photonic quantum technologies

The discussion from the last section shows that the lack of practical verification tools is a significant roadblock for the development of reliable quantum simulations. In particular, there seem to be fundamental obstacles for purely classical verification of photonic quantum simulations. But actually, there is no need to restrict to classical verification techniques. Very simple

quantum operations can be performed reliably, such as single mode measurements. These operations can be used to verify that the right quantum state is prepared in a photonic preparation scheme. In Publication [AGKE14] (see Appendix B) different notions of verification are discussed and a verification protocol using single mode measurements is provided. The protocol verifies m-mode Gaussian states and a class of linear optical states, which are given by a passive Gaussian unitary acting on an m-mode Fock basis state with n photons. In both cases, one can even allow for post-selected target states. The protocol is efficient in m and the inverse post-selection success probability but inefficient in n. The only restriction is that quadrature measurements of the prepared state have bounded variances.

The verification protocol rests on Gaussian extremality based fidelity estimation and can replace non-efficient tomography and compressed sensing techniques [Gro+10] for the important case where the target state is known. The protocol can verify states with negative Wigner functions, which are not covered by the known classical simulation schemes [ME12b, VWFE13]. With this verification, Boson-Sampling experiments with a constant number of photons [Bro+13, Spr+13, Til+13, Cre+13], all the states preparable with Knill-Laflamme-Milburn-like schemes [KLM01, Kok+07] with inverse polynomially bounded post-selection probability, state preparations for measurement based quantum computing with photonic states [Men+06, Yok+13], and realization of multipartite entanglement [CMP14, Hua+11, Yao+12] can be verified efficiently. This suggests that, using simple quantum operations, one can indeed efficiently verify quantum simulations that are not tractable classically efficiently.

4 Conclusions

In this thesis, we have established various results on reliable classical as well as quantum simulations of complex quantum systems, in particular, of local Liouvillian time evolution, systems in thermal equilibrium, and photonic simulations. For each of these systems, we now briefly summarize the results subsequently outline a few open problems.

We have seen that time evolution of open quantum lattice systems can be simulated efficiently on quantum computers under reasonable assumptions. This result can be interpreted as a Church-Turing type statement and also reveals limits of efficient state generation. Moreover, such time evolution can be locally simulated efficiently in the system size even on classical computers. Together this also provides a mathematical toolbox for classical simulations of Markovian systems in, e.g., tensor network based methods. We also have identified an obstacle for numerical truncations of mixed states that prevents the reliable simulation of such systems on a larger scale: For the numerically convenient parametrization of operators in matrix product form deciding positivity is NP-hard in the system size. If no bound on the system size is imposed this even becomes an undecidable problem in the translation invariant setting. In order to better understand and overcome this positivity problem, we suggest the following:

- Strengthen the hardness result on the bounded threshold problem as a weak membership problem to positivity checks with polynomial accuracy (see pages 71f), if possible.
- Find optimal tests checking positivity of MPOs.
- Identify tractable and physically relevant instances.
- Finally, find algorithms that efficiently and reliably simulate open quantum systems (for physically relevant instances).
- If a system is subjected to local noise, is it still hard to classically simulate its continuous time evolution? What are conditions for hardness and/or classical tractability? Here a first analysis suggests that the time scale on which the noise is correlated plays a crucial role.

Thermal states are ubiquitous in nature and appear as stationary states in many situations. Here, we have extended the concept of intensive temperature to interacting quantum systems. This is done by relating the length scale on which temperature can be defined (in a canonical

way) to a length scale on which correlations, with respect to a certain correlation measure, are negligible. For temperatures above a universal threshold, we show that these correlations decay exponentially. As a consequence, thermal states are stable against distant Hamiltonian perturbations in this regime of high temperature. Moreover, above the temperature threshold, local expectation values can be approximated efficiently in the system size. The temperature threshold also upper bounds physical critical temperatures of phase transitions involving long-range order. These results are very general but only provide coarse bounds for specific models. Here, one should be able to use more specific properties of a model to improve the results. Moreover, connecting our methods with the recent results [RGE12] on thermalization in closed quantum systems could lead to great progress towards a general microscopic understanding of thermalization. More specifically, we leave open the following problems:

- Understand better the generalized covariance (for specific models). Is its magnitude convex in τ? How does the averaged generalized covariance scale in the limit of large β? One would expect that, in many situations, the averaged generalized covariance also decays exponentially below the critical temperature. Can this intuition be made rigorous? This would automatically extent our results on universal locality.

- Finding smaller values for the critical temperature $1/\beta^*$ (see page 50) would, e.g., constitute tighter bounds on phase transitions temperatures.

- What Hamiltonians have long-range order at the highest possible temperature?

- It is rigorously known that thermal states can arise dynamically in systems, where the coupling to the environment is bounded by a small, system size independent constant [RGE12]. Can our methods to exploit the locality structure be used to improve upon this? With that, one might be able to prove stronger and more general version of thermalization.

  The τ-averaged generalized covariance (called "Matsubara time-ordered average" in its "truncated version" by Kitaev) is closely related to local currents [Kit06]; and physical intuition tells us that systems dynamically thermalize due to transport [CE10, Gog14]. Searching for local transport conditions, maybe in terms of Kitaev's local currents from Ref. [Kit06], and rigorously relating them to thermalization and also to the stability of thermal states seems to be an interesting research program on its own.

Inspired by these results about simulations of complex quantum systems, we have asked the question of what a reliable simulation actually is in the first place and encounter the verification problem of quantum simulations: Once a quantum simulation of some system outperforms classical computers, it becomes unclear how one can verify that the simulation indeed simulates the system. For the case of Boson-Sampling, sampling complexity lower bounds on the verification in a restricted setting are discussed. These bounds suggest that Boson-Sampling

cannot be verified efficiently with classical resources only. Driven by this insight, a reliable verification scheme for photonic state preparations was developed, which uses single-mode measurements. This scheme is efficient in the number of modes and allows for efficient verification of, e.g., Boson-Sampling experiments with a constant number of photons and state preparations necessary for measurement based quantum computing. The verification scheme turns out be be compatible with post-selection on measurements of ancillary modes and turns out to be efficient in the post selection success probability. Nevertheless, in the big goal of achieving reliable quantum simulations, this work constitutes only a first step. There are various problems to be overcome and here we only mention the most fundamental ones:

- What is a reasonable notion of "quantum simulations"? In the literature, this term refers to a lot of different ideas.
- Are there interesting, classically intractable problems, other than decision problems, solutions of which can be verified efficiently classically? For instance, is there a version of Boson-Sampling that can be verified classically efficiently?

In this thesis, we have presented various rigorous results on simulations of quantum systems, where locality and complexity play an intertwined role and by this contribute to a better understanding of complex quantum systems.

Bibliography

[AA11] S. Aaronson and A. Arkhipov, "*The computational complexity of linear optics*", in: STOC'11: Proc. 43rd Ann. ACM Symp. Theor. Comput. Pp. 333–342 (2011), URL: http://arxiv.org/abs/1011.3245.

[AALV09] D. Aharonov, I. Arad, Z. Landau, and U. Vazirani, "*The detectability lemma and quantum gap amplification*", in: STOC'09: Proc. 41st Ann. ACM Symp. Theor. Comput. Pp. 417–426 (2009), URL: http://arxiv.org/abs/0811.3412.

[AAVL11] D. Aharonov, I. Arad, U. Vazirani, and Z. Landau, "*The detectability lemma and its applications to quantum Hamiltonian complexity*", New J. Phys. **13**, 113043 (2011), DOI: 10.1088/1367-2630/13/11/113043.

[AGIK09] D. Aharonov, D. Gottesman, S. Irani, and J. Kempe, "*The Power of Quantum Systems on a Line*", Commun. Math. Phys. **287**, 41–65 (2009), DOI: 10.1007/s00220-008-0710-3.

[AGKE14] L. Aolita, C. Gogolin, M. Kliesch, and J. Eisert, "*Quantum certification of photonic quantum simulations*", submitted to Nat. Comm. (2014), arXiv:quant-ph/1407.4817.

[AGW12] A. Aspuru-Guzik and P. Walther, "*Photonic quantum simulators*", Nat. Phys. **8**, 285–291 (2012), DOI: 10.1038/nphys2253.

[AM76] N. W. Ashcroft and D. N. Mermin, "*Solid state physics*", Thomson Learning, 1976, URL: http://www.worldcat.org/isbn/0030839939.

[AOLL12] C. Ates, B. Olmos, W. Li, and I. Lesanovsky, "*Dissipative Binding of Lattice Bosons through Distance-Selective Pair Loss*", Phys. Rev. Lett. **109**, 233003 (2012), DOI: 10.1103/PhysRevLett.109.233003.

[Bar+11] J. T. Barreiro, M. Muller, P. Schindler, D. Nigg, T. Monz, M. Chwalla, M. Hennrich, C. F. Roos, P. Zoller, and R. Blatt, "*An open-system quantum simulator with trapped ions*", Nature **470**, 486–491 (2011), DOI: 10.1038/nature09801.

[Bar+12] C.-E. Bardyn, M. A. Baranov, E. Rico, A. İmamoğlu, P. Zoller, and S. Diehl, "*Majorana Modes in Driven-Dissipative Atomic Superfluids with a Zero Chern Number*", Phys. Rev. Lett. **109**, 130402 (2012), DOI: 10.1103/PhysRevLett.109.130402.

[Bat+13] T. Batu, L. Fortnow, R. Rubinfeld, W. D. Smith, and P. White, "*Testing closeness of discrete distributions*", J. ACM **60**, 4 (2013), DOI: 10.1145/2432622.2432626.

[BDN12] I. Bloch, J. Dalibard, and S. Nascimbene, "*Quantum simulations with ultracold quantum gases*", Nat. Phys. **8**, 267–276 (2012), DOI: 10.1038/nphys2259.

[BH13] F. G. S. L. Brandao and M. Horodecki, "*An area law for entanglement from exponential decay of correlations*", Nat. Phys. **9**, 721–726 (2013), DOI: 10.1038/nphys2747.

[BK12] T. Barthel and M. Kliesch, "*Quasilocality and Efficient Simulation of Markovian Quantum Dynamics*", Phys. Rev. Lett. **108**, 230504 (2012), DOI: 10.1103/PhysRevLett.108.230504.

[BKE10] T. Barthel, M. Kliesch, and J. Eisert, "*Real-Space Renormalization Yields Finite Correlations*", Phys. Rev. Lett. **105**, 010502 (2010), DOI: 10.1103/PhysRevLett.105.010502.

[Boi+13] S. Boixo, T. Albash, F. M. Spedalieri, N. Chancellor, and D. A. Lidar, "*Experimental signature of programmable quantum annealing*", Nat. Commun. **4** (2013), DOI: 10.1038/ncomms3067.

[Boi+14] S. Boixo, T. F. Ronnow, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer, "*Evidence for quantum annealing with more than one hundred qubits*", Nat. Phys. **10**, 218–224 (2014), DOI: 10.1038/nphys2900.

[BR12] R. Blatt and C. F. Roos, "*Quantum simulations with trapped ions*", Nat. Phys. **8**, 277–284 (2012), DOI: 10.1038/nphys2252.

[BR97] O. Bratteli and D. W. Robinson, "*Operator Algebras and Quantum Statistical Mechanics: Equilibrium States. Models in Quantum Statistical Mechanics*", 2nd ed., Springer, 1997.

[Bra+14] S. Braun, M. Friesdorf, S. S. Hodgman, M. Schreiber, J. P. Ronzheimer, A. Riera, M. del Rey, I. Bloch, J. Eisert, and U. Schneider, "*Emergence of coherence and the dynamics of quantum phase transitions*", (2014), arXiv:cond-mat.quant-gas/1403.7199.

[Bro+13] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White, "*Photonic Boson Sampling in a Tunable Circuit*", Science **339**, 794–798 (2013), DOI: 10.1126/science.1231440.

[Bro11] D. Browne, "*Quantum simulation hits the open road*", Physics **4**, 72 (2011), DOI: 10.1103/Physics.4.72.

[CB13] Z. Cai and T. Barthel, "*Algebraic versus Exponential Decoherence in Dissipative Many-Particle Systems*", Phys. Rev. Lett. **111**, 150403 (2013), DOI: 10.1103/PhysRevLett.111.150403.

[CDEO08] M. Cramer, C. M. Dawson, J. Eisert, and T. J. Osborne, "*Exact Relaxation in a Class of Nonequilibrium Quantum Lattice Systems*", Phys. Rev. Lett. **100**, 030602 (2008), DOI: 10.1103/PhysRevLett.100.030602.

[CE10] M. Cramer and J. Eisert, "*A quantum central limit theorem for non-equilibrium systems: exact local relaxation of correlated states*", New J. Phys. **12**, 055020, 055020 (2010), DOI: 10.1088/1367-2630/12/5/055020.

[CLMPG13] T. S. Cubitt, A. Lucia, S. Michalakis, and D. Perez-Garcia, "*Stability of local quantum dissipative systems*", ArXiv e-prints (2013), arXiv:quant-ph/1303.4744.

[CMP14] M. Chen, N. C. Menicucci, and O. Pfister, "*Experimental Realization of Multipartite Entanglement of 60 Modes of a Quantum Optical Frequency Comb*", Phys. Rev. Lett. **112**, 120505 (2014), DOI: 10.1103/PhysRevLett.112.120505.

[Com] "*Complexity zoo*", URL: https://complexityzoo.uwaterloo.ca/Complexity_Zoo.

[CPGW14] T. S. Cubitt, D. Perez-Garcia, and M. M. Wolf, "*Undecidability of the spectral gap problem*", talk given at XVII Conference on Quantum Information Processing, Barcelona, Feb. 6 (2014), URL: http://benasque.org/2014QIP/cgi-bin/program.pl.

[CZ12] J. I. Cirac and P. Zoller, "*Goals and opportunities in quantum simulation*", Nat. Phys. **8**, 264–266 (2012), DOI: 10.1038/nphys2275.

[Dav79] E. B. Davies, "*Generators of dynamical semigroups*", J. Funct. Anal. **34**, 421–432 (1979), DOI: 10.1016/0022-1236(79)90085-5.

[Die+08] S. Diehl, A. Micheli, A. Kantian, B. Kraus, H. P. Buchler, and P. Zoller, "*Quantum states and phases in driven open quantum systems with cold atoms*", Nat. Phys. **4**, 878–883 (2008), DOI: 10.1038/nphys1073.

[DKSV04] A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal, "*Time-dependent density-matrix renormalization-group using adaptive effective Hilbert spaces*", J. Stat. Mech. Theor. Exp. **2004**, P04005 (2004), DOI: 10.1088/1742-5468/2004/04/P04005.

[DYDZ10] S. Diehl, W. Yi, A. J. Daley, and P. Zoller, "*Dissipation-Induced d-Wave Pairing of Fermionic Atoms in an Optical Lattice*", Phys. Rev. Lett. **105**, 227001 (2010), DOI: 10.1103/PhysRevLett.105.227001.

[ECP10] J. Eisert, M. Cramer, and M. B. Plenio, "*Colloquium: Area laws for the entanglement entropy*", Rev. Mod. Phys. **82**, 277–306 (2010), DOI: 10.1103/RevModPhys.82.277.

[Eis13] J. Eisert, "*Entanglement and tensor network states*", Modeling and Simulation **3**, 520 (2013), arXiv:quant-ph/1308.3318.

[EMG12] J. Eisert, M. P. Müller, and C. Gogolin, "*Quantum measurement occurrence is undecidable*", Phys. Rev. Lett. **108**, 260501 (2012), DOI: 10.1103/PhysRevLett.108.260501.

[EP10] J. Eisert and T. Prosen, "*Noise-driven quantum criticality*", (2010), arXiv:quant-ph/1012.5013.

[FGSA12] A. Ferraro, A. García-Saez, and A. Acín, "*Intensive temperature and quantum correlations for refined quantum measurements*", Europhys. Lett. **98**, 10009 (2012), DOI: 10.1209/0295-5075/98/10009.

[FIK08] F. Franchini, A. R. Its, and V. E. Korepin, "*Renyi entropy of the XY spin chain*", J. Phys. A **41**, 025302 (2008), DOI: 10.1088/1751-8113/41/2/025302.

[GB02] Y. Gao and Y. Bando, "*Nanotechnology: Carbon nanothermometer containing gallium*", Nature **415**, 599–599 (2002), DOI: 10.1038/415599a.

[GI09] D. Gottesman and S. Irani, "*The Quantum and Classical Complexity of Translationally Invariant Tiling and Hamiltonian Problems*", Proc. 50th Ann. FOCS – IEEE, 95–104 (2009), DOI: 10.1109/FOCS.2009.22.

[Gin65] J. Ginibre, "*Reduced Density Matrices of Quantum Gases. II. Cluster Property*", J. Math. Phys. **6**, 252–262 (1965), DOI: 10.1063/1.1704276.

[GKAE13] C. Gogolin, M. Kliesch, L. Aolita, and J. Eisert, "*Boson-Sampling in the light of sample complexity*", (2013), arXiv:quant-ph/1306.3995.

[Gla+12] A. W. Glaetzle, R. Nath, B. Zhao, G. Pupillo, and P. Zoller, "*Driven-dissipative dynamics of a strongly interacting Rydberg gas*", Phys. Rev. A **86**, 043403 (2012), DOI: 10.1103/PhysRevA.86.043403.

[GLTZ06] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, "*Canonical Typicality*", Phys. Rev. Lett. **96**, 050403 (2006), DOI: 10.1103/PhysRevLett.96.050403.

[Gog14] C. Gogolin, private communication on Jul. 23, 2014.

[Gre69] W. Greenberg, "*Critical temperature bounds of quantum lattice gases*", Commun. Math. Phys. **13**, 335–344 (1969), DOI: 10.1007/BF01645417.

[Gro+10] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, "*Quantum State Tomography via Compressed Sensing*", Phys. Rev. Lett. **105**, 150401 (2010), DOI: 10.1103/PhysRevLett.105.150401.

[Hae+11] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pi žorn, H. Verschelde, and F. Verstraete, "*Time-Dependent Variational Principle for Quantum Lattices*", Phys. Rev. Lett. **107**, 070601 (2011), DOI: 10.1103/PhysRevLett.107.070601.

[Has06] M. B. Hastings, "*Solving gapped Hamiltonians locally*", Phys. Rev. B **73**, 085115 (2006), DOI: 10.1103/PhysRevB.73.085115.

[HDR90] J. Huyghebaert and H. De Raedt, "*Product formula methods for time-dependent Schrodinger problems*", J. Phys. A **23**, 5777 (1990), DOI: 10.1088/0305-4470/23/24/019.

[HK06] M. B. Hastings and T. Koma, "*Spectral Gap and Exponential Decay of Correlations*", Commun. Math. Phys. **265**, 781–804 (2006), DOI: 10.1007/s00220-006-0030-4.

[HM05] M. Hartmann and G. Mahler, "*Measurable consequences of the local breakdown of the concept of temperature*", Europhys. Lett. **70**, 579 (2005), DOI: 10.1209/epl/i2004-10518-5.

[HMH04] M. Hartmann, G. Mahler, and O. Hess, "*Existence of Temperature on the Nanoscale*", Phys. Rev. Lett. **93**, 080402 (2004), DOI: 10.1103/PhysRevLett.93.080402.

[HTK12] A. A. Houck, H. E. Tureci, and J. Koch, "*On-chip quantum simulation with superconducting circuits*", Nat. Phys. **8**, 292–299 (2012), DOI: 10.1038/nphys2251.

[Hua+11] Y.-F. Huang, B.-H. Liu, L. Peng, Y.-H. Li, L. Li, C.-F. Li, and G.-C. Guo, "*Experimental generation of an eight-photon Greenberger-Horne-Zeilinger state*", Nat. Commun. **2**, 546 (2011), DOI: 10.1038/ncomms1556.

[HW05] M. B. Hastings and X.-G. Wen, "*Quasiadiabatic continuation of quantum states: The stability of topological ground-state degeneracy and emergent gauge invariance*", Phys. Rev. B **72**, 045141 (2005), DOI: 10.1103/PhysRevB.72.045141.

[KE13] M. J. Kastoryano and J. Eisert, "*Rapid mixing implies exponential decay of correlations*", J. Math. Phys. **54**, 102201 (2013), DOI: 10.1063/1.4822481.

[KGE13] M. Kliesch, C. Gogolin, and J. Eisert, "*Lieb-Robinson bounds and the simulation of time evolution of local observables in lattice systems*", (2013), arXiv:quant-ph/1306.0716.

[KGE14a] M. Kliesch, D. Gross, and J. Eisert, "*Matrix-Product Operators and States: NP-Hardness and Undecidability*", Phys. Rev. Lett. **113**, 160503 (2014), DOI: 10.1103/PhysRevLett.113.160503.

[KGE14b] M. Kliesch, C. Gogolin, and J. Eisert, "*Lieb-Robinson Bounds and the Simulation of Time-Evolution of Local Observables in Lattice Systems*", in: Many-Electron Approaches in Physics, Chemistry and Mathematics, ed. by V. Bach and L. Delle Site, Mathematical Physics Studies, Springer International Publishing, 2014, pp. 301–318, DOI: 10.1007/978-3-319-06379-9_17, arXiv:1306.0716.

[Kit06] A. Kitaev, "*Anyons in an exactly solved model and beyond*", Annals of Physics **321** January Special Issue, 2 –111 (2006), DOI: 10.1016/j.aop.2005.10.005.

[Kli+11a] M. Kliesch, T. Barthel, C. Gogolin, M. Kastoryano, and J. Eisert, "*Dissipative Quantum Church-Turing Theorem*", Phys. Rev. Lett. **107**, 120501 (2011), DOI: 10.1103/PhysRevLett.107.120501.

[Kli+11b] M. Kliesch, T. Barthel, C. Gogolin, M. Kastoryano, and J. Eisert, *Supplementary Material (Dissipative Quantum Church-Turing Theorem)*, 2011, DOI: 10.1103/PhysRevLett.107.120501.

[Kli+14] M. Kliesch, C. Gogolin, M. J. Kastoryano, A. Riera, and J. Eisert, "*Locality of Temperature*", Phys. Rev. X **4**, 031019 (2014), DOI: 10.1103/PhysRevX.4.031019.

[KLM01] E. Knill, R. Laflamme, and G. J. Milburn, "*A scheme for efficient quantum computation with linear optics*", Nature **409**, 46–52 (2001), DOI: 10.1038/35051009.

[KLM07] P. Kaye, R. Laflamme, and M. Mosca, "*An Introduction to Quantum Computing*", Oxford University Press, 2007.

[Kok+07] P. Kok, W. Munro, K. Nemoto, T. Ralph, J. Dowling, and G. Milburn, "*Linear optical quantum computing with photonic qubits*", Rev. Mod. Phys. **79**, 135–174 (2007), DOI: 10.1103/RevModPhys.79.135.

[Kor04] V. E. Korepin, "*Universality of Entropy Scaling in One Dimensional Gapless Models*", Phys. Rev. Lett. **92**, 096402 (2004), DOI: 10.1103/PhysRevLett.92.096402.

[KP13] R. Koenig and F. Pastawski, "*Generating topological order: no speedup by dissipation*", (2013), arXiv:quant-ph/1310.1037.

[Kra+11] H. Krauter, C. A. Muschik, K. Jensen, W. Wasilewski, J. M. Petersen, J. I. Cirac, and E. S. Polzik, "*Entanglement Generated by Dissipation and Steady State Entanglement of Two Macroscopic Objects*", Phys. Rev. Lett. **107**, 080503 (2011), DOI: 10.1103/PhysRevLett.107.080503.

[KT13] M. J. Kastoryano and K. Temme, "*Quantum logarithmic Sobolev inequalities and rapid mixing*", J. Math. Phys. **54**, 052202 (2013), DOI: http://dx.doi.org/10.1063/1.4804995.

[KWE13] M. J. Kastoryano, M. M. Wolf, and J. Eisert, "*Precisely Timing Dissipative Quantum Information Processing*", Phys. Rev. Lett. **110**, 110501 (2013), DOI: 10.1103/PhysRevLett.110.110501.

[LCB14] M. Lubasch, J. I. Cirac, and M.-C. Bañuls, "*Algorithms for finite Projected Entangled Pair States*", ArXiv e-prints (2014), arXiv:quant-ph/1405.3259.

[Lin76] G. Lindblad, "*On the generators of quantum dynamical semigroups*", Commun. Math. Phys. **48**, 119–130 (1976), DOI: 10.1007/BF01608499.

[LPS10] N. Linden, S. Popescu, and P. Skrzypczyk, "*How Small Can Thermal Machines Be? The Smallest Possible Refrigerator*", Phys. Rev. Lett. **105**, 130401 (2010), DOI: 10.1103/PhysRevLett.105.130401.

[LPSW09] N. Linden, S. Popescu, A. J. Short, and A. Winter, "*Quantum mechanical evolution towards thermal equilibrium*", Phys. Rev. E **79**, 061103 (2009), DOI: 10.1103/PhysRevE.79.061103.

[LR72] E. H. Lieb and D. W. Robinson, "*The finite group velocity of quantum spin systems*", Commun. Math. Phys. **28**, 251–257 (1972), URL: http://projecteuclid.org/euclid.cmp/1103858407.

[LRV04] J. I. Latorre, E. Rico, and G. Vidal, "*Ground state entanglement in quantum spin chains*", Quant. Inf. Comput. **4**, 48–92 (2004), eprint: quant-ph/0304098.

[LVV13] Z. Landau, U. Vazirani, and T. Vidick, "*A polynomial-time algorithm for the ground state of 1D gapped local Hamiltonians*", (2013), arXiv:quant-ph/1307.5143.

[MAMW13] M. P. Mueller, E. Adlam, L. Masanes, and N. Wiebe, "*Thermalization and canonical typicality in translation-invariant quantum lattice systems*", (2013), arXiv:quant-ph/1312.7420.

[ME12a] A. Mari and J. Eisert, "*Cooling by Heating: Very Hot Thermal Light Can Significantly Cool Quantum Systems*", Phys. Rev. Lett. **108**, 120602 (2012), DOI: 10.1103/PhysRevLett.108.120602.

[ME12b] A. Mari and J. Eisert, "*Positive Wigner Functions Render Classical Simulation of Quantum Computation Efficient*", Phys. Rev. Lett. **109**, 230503 (2012), DOI: 10.1103/PhysRevLett.109.230503.

[Men+06] N. C. Menicucci, P. van Loock, M. Gu, C. Weedbrook, T. C. Ralph, and M. A. Nielsen, "*Universal Quantum Computation with Continuous-Variable Cluster States*", Phys. Rev. Lett. **97**, 110501 (2006), DOI: 10.1103/PhysRevLett.97.110501.

[MS14] R. Movassagh and P. W. Shor, "*Power law violation of the area law in critical spin chains*", (2014), arXiv:quant-ph/1408.1657.

[MSVC14] A. Molnár, N. Schuch, F. Verstraete, and J. I. Cirac, "*Approximating Gibbs states of local Hamiltonians efficiently with PEPS*", (2014), arXiv:quant-ph/1406.2973.

[MZ13] S. Michalakis and J. P. Zwolak, "*Stability of Frustration-Free Hamiltonians*", Commun. Math. Phys. **322**, 277–302 (2013), DOI: 10.1007/s00220-013-1762-6.

[Nag10] D. Nagaj, "*Fast universal quantum computation with railroad-switch local Hamiltonians*", J. Math. Phys. **51**, 062201 (2010), DOI: 10.1063/1.3384661.

[NC00] M. A. Nielsen and I. L. Chuang, "*Quantum computation and quantum information*", Cambridge University Press, 2000.

[NS06] B. Nachtergaele and R. Sims, "*Lieb-Robinson Bounds and the Exponential Clustering Theorem*", Commun. Math. Phys. **265**, 119–130 (2006), DOI: 10.1007/s00220-006-1556-1.

[NVZ11] B. Nachtergaele, A. Vershynina, and V. A. Zagrebnov, "*Lieb-Robinson bounds and existence of the thermodynamic limit for a class of irreversible quantum dynamics*", Am. Math. Soc. Contemp. Math. **552**, 161–175 (2011), DOI: 10.1090/conm/552.

[Oru13] R. Orus, "*A Practical Introduction to Tensor Networks: Matrix Product States and Projected Entangled Pair States*", (2013), arXiv:cond-mat.str-el/1306.2164.

[Osb06] T. J. Osborne, "*Efficient Approximation of the Dynamics of One-Dimensional Quantum Spin Systems*", Phys. Rev. Lett. **97**, 157202 (2006), DOI: 10.1103/PhysRevLett.97.157202.

[Osb12] T. J. Osborne, "*Hamiltonian complexity*", Rep. Prog. Phys. **75**, 022001 (2012), DOI: 10.1088/0034-4885/75/2/022001.

[PBE10] C. Pineda, T. Barthel, and J. Eisert, "*Unitary circuits for strongly correlated fermions*", Phys. Rev. A **81**, 050303 (2010), DOI: 10.1103/PhysRevA.81.050303.

[PGVWC07] D. Pérez-García, F. Verstraete, M. M. Wolf, and J. I. Cirac, "*Matrix product state representations*", Quant. Inf. Comput. **7**, 401–430 (2007), URL: http://arxiv.org/abs/quant-ph/0608197.

[Pou10] D. Poulin, "*Lieb-Robinson Bound and Locality for General Markovian Quantum Dynamics*", Phys. Rev. Lett. **104**, 190401 (2010), DOI: 10.1103/PhysRevLett.104.190401.

[PQSV11] D. Poulin, A. Qarry, R. Somma, and F. Verstraete, "*Quantum Simulation of Time-Dependent Hamiltonians and the Convenient Illusion of Hilbert Space*", Phys. Rev. Lett. **106**, 170501 (2011), DOI: 10.1103/PhysRevLett.106.170501.

[Pre12] J. Preskill, "*Quantum supremacy now?*", blog entry July 22, 2012 in Quantum Frontiers, URL: http://quantumfrontiers.com/2012/07/22/supremacy-now/.

[PY95] Y. Park and H. Yoo, "*Uniqueness and clustering properties of Gibbs states for classical and quantum unbounded spin systems*", J. Stat. Phys. **80**, 223–271 (1995), DOI: 10.1007/BF02178359.

[RDO08] M. Rigol, V. Dunjko, and M. Olshanii, "*Thermalization and its mechanism for generic isolated quantum systems*", Nature **452**, 854–858 (2008), DOI: 10.1038/nature06838.

[RGE12] A. Riera, C. Gogolin, and J. Eisert, "*Thermalization in Nature and on a Quantum Computer*", Phys. Rev. Lett. **108**, 080402 (2012), DOI: 10.1103/PhysRevLett.108.080402.

[RK12] P. Reimann and M. Kastner, "*Equilibration of isolated macroscopic quantum systems*", New J. Phys. **14**, 043020 (2012), DOI: 10.1088/1367-2630/14/4/043020.

[Rue64] D. Ruelle, "*Cluster Property of the Correlation Functions of Classical Gases*", Rev. Mod. Phys. **36**, 580–584 (1964), DOI: 10.1103/RevModPhys.36.580.

[Rue69] D. Ruelle, "*Statistical Mechanics: Rigorous Results*", W. A. Benjamin, 1969.

[Sch+13] P. Schindler, M. Muller, D. Nigg, J. T. Barreiro, E. A. Martinez, M. Hennrich, T. Monz, S. Diehl, P. Zoller, and R. Blatt, "*Quantum simulation of dynamical maps with trapped ions*", Nat. Phys. **9**, 361–367 (2013), DOI: 10.1038/nphys2630.

[Sch11] U. Schollwöck, "*The density-matrix renormalization group in the age of matrix product states*", Ann. Phys. **326** January 2011 Special Issue, 96 –192 (2011), DOI: http://dx.doi.org/10.1016/j.aop.2010.09.012.

[Sch13] N. Schuch, "*Condensed Matter Applications of Entanglement Theory*", (2013), arXiv:quant-ph/1306.5551.

[Sho97] P. Shor, "*Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*", SIAM J. Comput. **26**, 1484–1509 (1997), DOI: 10.1137/S0097539795293172.

[Spo78] H. Spohn, "*Entropy production for quantum dynamical semigroups*", J. Math. Phys. **19**, 1227–1230 (1978), DOI: 10.1063/1.523789.

[Spr+13] J. B. Spring, B. J. Metcalf, P. C. Humphreys, W. S. Kolthammer, X.-M. Jin, M. Barbieri, A. Datta, N. Thomas-Peter, N. K. Langford, D. Kundys, J. C. Gates, B. J. Smith, P. G. R. Smith, and I. A. Walmsley, "*Boson Sampling on a Photonic Chip*", Science **339**, 798–801 (2013), DOI: 10.1126/science.1231692.

[SS13] J. A. Smolin and G. Smith, "*Classical signature of quantum annealing*", (2013), arXiv:quant-ph/1305.4904.

[Suz76] M. Suzuki, "*Generalized Trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems*", Commun. Math. Phys. **51**, 183–190 (1976), URL: http://projecteuclid.org/euclid.cmp/1103900351.

[SWVC07] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, "*Computational Complexity of Projected Entangled Pair States*", Phys. Rev. Lett. **98**, 140506 (2007), DOI: 10.1103/PhysRevLett.98.140506.

[TATW03] M. Troyer, F. Alet, S. Trebst, and S. Wessel, "*Non-local Updates for Quantum Monte Carlo Simulations*", AIP Conf. Proc. **690**, 156–169 (2003), DOI: 10.1063/1.1632126.

[Til+13] M. Tillmann, B. Dakic, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, "*Experimental boson sampling*", Nat. Photon. **7**, 540–544 (2013), DOI: 10.1038/nphoton.2013.102.

[Tro+12] S. Trotzky, Y.-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwock, J. Eisert, and I. Bloch, "*Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional Bose gas*", Nat. Phys. **8**, 325–330 (2012), DOI: 10.1038/nphys2232.

[Tro59] H. F. Trotter, "*On the product of semi-groups of operators*", Proc. Amer. Math. Soc. **10**, 545–551 (1959), DOI: 10.1090/S0002-9939-1959-0108732-6.

[TW05] M. Troyer and U.-J. Wiese, "*Computational Complexity and Fundamental Limitations to Fermionic Quantum Monte Carlo Simulations*", Phys. Rev. Lett. **94**, 170201 (2005), DOI: 10.1103/PhysRevLett.94.170201.

[VC06] F. Verstraete and J. I. Cirac, "*Matrix product states represent ground states faithfully*", Phys. Rev. B **73**, 094423 (2006), DOI: 10.1103/PhysRevB.73.094423.

[VGRC04] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, "*Matrix Product Density Operators: Simulation of Finite-Temperature and Dissipative Systems*", Phys. Rev. Lett. **93**, 207204 (2004), DOI: 10.1103/PhysRevLett.93.207204.

[Vid04] G. Vidal, "*Efficient Simulation of One-Dimensional Quantum Many-Body Systems*", Phys. Rev. Lett. **93**, 040502 (2004), DOI: 10.1103/PhysRevLett.93.040502.

[Vid06] G. Vidal, "*A class of quantum many-body states that can be efficiently simulated*", (2006), arXiv:quant-ph/0610099v1.

[Vid07] G. Vidal, "*Entanglement Renormalization*", Phys. Rev. Lett. **99**, 220405 (2007), DOI: 10.1103/PhysRevLett.99.220405.

[Vid08] G. Vidal, "*Class of Quantum Many-Body States That Can Be Efficiently Simulated*", Phys. Rev. Lett. **101**, 110501 (2008), DOI: 10.1103/PhysRevLett.101.110501.

[Vid11] M. Vidyasagar, "*The complete realization problem for hidden Markov models: a survey and some new results*", Math. Control Signals Syst. **23**, 1–65 (2011), DOI: 10.1007/s00498-011-0066-7.

[VLRK03] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev, "*Entanglement in Quantum Critical Phenomena*", Phys. Rev. Lett. **90**, 227902 (2003), DOI: 10.1103/PhysRevLett.90.227902.

[VMC08] F. Verstraete, V. Murg, and J. Cirac, "*Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems*", Adv. Phys. **57**, 143–224 (2008), DOI: 10.1080/14789940801912366.

[Voj03] M. Vojta, "*Quantum phase transitions*", Rep. Prog. Phys. **66**, 2069 (2003), URL: http://stacks.iop.org/0034-4885/66/i=12/a=R01.

[VWFE13] V. Veitch, N. Wiebe, C. Ferrie, and J. Emerson, "*Efficient simulation scheme for a class of quantum optics experiments with non-negative Wigner representation*", New J. Phys. **15**, 013037 (2013), DOI: 10.1088/1367-2630/15/1/013037.

[VWIC09] F. Verstraete, M. M. Wolf, and J. Ignacio Cirac, "*Quantum computation and quantum-state engineering driven by dissipation*", Nat. Phys. **5**, 633–636 (2009), URL: http://dx.doi.org/10.1038/nphys1342.

[WC08] M. M. Wolf and J. I. Cirac, "*Dividing Quantum Channels*", Commun. Math. Phys. **279**, 147–168 (2008), DOI: 10.1007/s00220-008-0411-y.

[WCPG11] M. M. Wolf, T. S. Cubitt, and D. Perez-Garcia, "*Are problems in Quantum Information Theory (un)decidable?*", (2011), arXiv:quant-ph/1111.5425.

[Whi92] S. R. White, "*Density matrix formulation for quantum renormalization groups*", Phys. Rev. Lett. **69**, 2863–2866 (1992), DOI: 10.1103/PhysRevLett.69.2863.

[WVHC08] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, "*Area Laws in Quantum Systems: Mutual Information and Correlations*", Phys. Rev. Lett. **100**, 070502 (2008), DOI: 10.1103/PhysRevLett.100.070502.

[Yao+12] X.-C. Yao, T.-X. Wang, P. Xu, H. Lu, G.-S. Pan, X.-H. Bao, C.-Z. Peng, C.-Y. Lu, Y.-A. Chen, and J.-W. Pan, "*Observation of eight-photon entanglement*", Nat. Photon. **6**, 225–228 (2012), DOI: 10.1038/nphoton.2011.354.

[Yok+13] S. Yokoyama, R. Ukai, S. C. Armstrong, C. Sornphiphatphong, T. Kaji, S. Suzuki, J.-i. Yoshikawa, H. Yonezawa, N. C. Menicucci, and A. Furusawa, "*Ultra-large-scale continuous-variable cluster states multiplexed in the time domain*", Nat. Photon. **7**, 982–986 (2013), DOI: 10.1038/nphoton.2013.287.

[ZV04] M. Zwolak and G. Vidal, "*Mixed-State Dynamics in One-Dimensional Quantum Lattice Systems: A Time-Dependent Superoperator Renormalization Algorithm*", Phys. Rev. Lett. **93**, 207205 (2004), DOI: 10.1103/PhysRevLett.93.207205.

[Cre+13] A. Crespi, R. Osellame, R. Ramponi, D. J. Brod, E. F. Galvão, N. Spagnolo, C. Vitelli, E. Maiorino, P. Mataloni, and F. Sciarrino, "*Integrated multimode interferometers with arbitrary designs for photonic boson sampling*", Nat. Photon. **7**, 545–549 (2013), DOI: 10.1038/nphoton.2013.112.

[Wan+13] L. Wang, T. F. Rønnow, S. Boixo, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer, "*Comment on: "Classical signature of quantum annealing"*", (2013), arXiv:quant-ph/1305.5837.

A Appendix: Terminology

In this appendix we briefly explain some basics concepts from computer science and quantum information theory.

A.1 Concepts from computational complexity theory

We will use several terms from computational complexity theory, which is a branch of the theory of computation. The main task is, for a given model of computation, to classify computational problems according to their difficulty. We need the following models of computation:

- A *Turing machine* is the theoretic model of a classical computer without randomness.
- A *probabilistic Turing machine* is equivalent to a Turing machine that can use random numbers as an additional resource.
- The *unitary* or *quantum circuit model* is the standard model of a quantum computation. The inputs are qubits, the actual computation is a sequence of unitaries, called *quantum gates*, each of which acts on at most two qubits, and the output is measured in a canonical basis called *computational basis*.

A computational problem has a natural problem size, e.g., for the problem of factoring an integer the problem size is the number of digits. *Decision problems* are problems with finite input and binary solutions, i.e., yes/no answers. By restricting the resources in terms of the problem size one obtains a large number of complexity classes [Com]. For instance:

- P denotes that class of decision problems that can be solved by a Turing machine in *polynomial time*, i.e., in a time scaling at most polynomially in the problem size.
- NP denotes that class of decision problems where solutions can be verified by a Turing machine in polynomial time. Equivalently, on can define NP as the class of problems that can be solved in polynomial time in another theoretical model of computation called *non-deterministic Turing machine*. Hence, the name, NP stands for *non-deterministic polynomial time*.

- BPP, standing for *bounded-error probabilistic polynomial time*, is the class of decision problems that can be solved by a probabilistic Turing machine in polynomial time with error probability upper bounded by $1/3$.

- BQP, standing for *bounded error quantum polynomial time*, is the class of decision problems that can be solve by a quantum computer in polynomial time, with error probability upper bounded by $1/3$.

- QMA, standing for *Quantum Merlin Arthur*, is a class of decision problems that can be verified by a quantum computer in polynomial time with probability at least $2/3$. Note that QMA is related to BQP in a similar way as NP is related to P. Therefore, it is often called the *quantum analogue of* NP.

It is not hard to show that $\mathsf{P} \subset \mathsf{BPP} \subset \mathsf{BQP} \subset \mathsf{QMA}$ and $\mathsf{P} \subset \mathsf{NP} \subset \mathsf{QMA}$ and it is believed that the inclusions, except maybe the first one, are strict. $\mathsf{P} \stackrel{?}{=} \mathsf{NP}$ is called the P *versus* NP *problem* and is one of the seven *Millennium Prize Problems* of the Clay Mathematics Institute.

We call a decision problem *undecidable* if there cannot exist an algorithm that decides every instance of the problem correctly. For a given model of computation we say that a problem can be solved *efficiently* or is *tractable* if it can be solved in polynomial time. We say that it is *not tractable* if this is not the case and *intractable* if it can be solved but not in polynomial time, so that intractable problems are always decidable but not tractable problems might also be undecidable. Moreover, a problem L is called *hard* for NP (QMA), or NP-hard (QMA-hard), if any other problem in NP (QMA) can be reduced to L with overhead bounded polynomially. A hard problem that is itself in NP (QMA) is called *complete* for NP (QMA) or just NP-complete (QMA-complete). Finally, we briefly name a few famous examples:

- The so-called *local Hamiltonian problem*, which is the decision version of finding the ground state energy of a 2-local Hamiltonian, is QMA-complete. More precisely, the local Hamiltonian problem is to decide, whether the ground state energy of a Hamiltonian with two-body interactions is either below a value a or above another value $b > a$, for $1/(b-a)$ scaling at most polynomially in the number of subsystems.

- The decision version of integer factoring is clearly in NP, believed to not be in P, and also believed not to be NP-complete.

- An NP-complete problem is the famous *travelling salesman problem*, which is the decision version of finding a shortest route, starting from one city, passing any city of a list of cities exactly once, and returning to the origin city in the end.

A.2 Concepts from quantum (information) theory

The following mathematical definition is important to classify valid quantum operations, both in the Schrödinger and the Heisenberg picture. A linear map T between operator spaces is called *completely positive* if for all auxiliary spaces X with identity map id_X, the map $T \otimes \mathrm{id}_X$ takes positive semi-definite operators to positive semi-definite ones (see, e.g., Ref. [Lin76] for details).

A *spin system* is a quantum mechanical system described by a finite dimensional Hilbert space. In the special case where the dimension is two one also calls it a *spin-$\frac{1}{2}$ system* or a *qubit*.

A *quantum lattice system* is a multipartite quantum system with an associated *graph*, where the the subsystems are associated with the vertices and the interactions with the edges. The graph is also called *interaction graph*. In the case where the single subsystems are spins/ fermions/ bosons the system is also called a spin/ fermionic/ bosonic lattice system. See, e.g., Sec. 2.3 and Sec. 5 of Publication [KGE14b] on pages 17f and 26 for more details on spin and fermionic lattice systems, respectively.

If one draws a pure state of a spin lattice system uniformly at random, then the state reduced to a subsystem has an entropy scaling as the number of sites in that subsystem and one says that the entanglement scales as the volume in that case. In many physical situations, however, one encounters pure states where the entropy only scales at most as the boundary of the considered subsystem. One then says that the state fulfills an *area law* for the entanglement entropy. This can be generalized to mixed states and/or different correlation measures such as the Rényi entropy or mutual information. See Publication [KGE14b] on pages 23f for more details.

A class of pure states for one-dimensional spin systems that fulfills such an area-law for all entanglement Rényi entropies are *matrix product states (MPS)*. Here the expansion coefficient in the product basis are given in terms of matrices (see page 101 for a precise definition). They are convenient for numerical simulations as they approximate well important classes of physically relevant states. See Publication [KGE14b] pages 22ff for more details.

Similarly, *matrix product operators (MPOs)* parametrize operators and are used in numerical simulations to describe local observables, update steps in (imaginary) time evolutions, and density matrices. See Publication [KGE14a] pages 71f for more details.

B Appendix: Other publications generated during this thesis

While firs-author publications are presented in the main text, the author's other publications [AGKE14, BK12, BKE10, GKAE13] are included in this appendix for the sake of completeness.

Real-Space Renormalization Yields Finite Correlations

Thomas Barthel,[1] Martin Kliesch,[1] and Jens Eisert[1,2]
[1]*Institute for Physics and Astronomy, Potsdam University, 14476 Potsdam, Germany*
[2]*Institute for Advanced Study Berlin, 14193 Berlin, Germany*
(Received 17 March 2010; revised manuscript received 30 April 2010; published 2 July 2010)

Real-space renormalization approaches for quantum lattice systems generate certain hierarchical classes of states that are subsumed by the multiscale entanglement renormalization Ansatz (MERA). It is shown that, with the exception of one spatial dimension, MERA states are actually states with finite correlations, i.e., projected entangled pair states (PEPS) with a bond dimension independent of the system size. Hence, real-space renormalization generates states which can be encoded with local effective degrees of freedom, and MERA states form an efficiently contractible class of PEPS that obey the area law for the entanglement entropy. It is further pointed out that there exist other efficiently contractible schemes violating the area law.

DOI: 10.1103/PhysRevLett.105.010502 PACS numbers: 03.67.Mn, 02.70.−c, 03.65.Ud, 64.60.ae

Renormalization group (RG) methods aim at solving many-body problems by treating energy scales in an iterative fashion, progressing from high to low energies [1]. One of its earliest formulations is the real-space RG which works by repeated steps of thinning out local degrees of freedom and rescaling of the system as in Kadanoff's block spin transformation [2]. In real-space RG approaches to quantum lattice models [3], in each RG step τ, the system is partitioned into small blocks. From those blocks high-energy states are eliminated and the Hamiltonian $\hat{H}_{\tau+1}$ for the renormalized system is obtained by applying the corresponding projection operators, exactly $\hat{H}_{\tau+1} = \hat{P}_{\tau+1}\hat{H}_\tau\hat{P}^\dagger_{\tau+1}$ or in some appropriate approximation, followed by a coarse graining of the lattice. This is iterated, e.g., until a step $\tau = T$ is reached where the renormalized system consists of a single small block for which the ground state $|\mathrm{gs}_T\rangle$ can be obtained exactly. Applying the RG transformations in reverse order yields an approximation $\hat{P}^\dagger_1\hat{P}^\dagger_2\ldots\hat{P}^\dagger_T|\mathrm{gs}_T\rangle$ to the ground state of the original model. Those states, generated by the real-space RG, fall into the class of so-called tree tensor networks (TTN) [4]. A recent more elaborate real-space RG scheme, the multiscale entanglement renormalization Ansatz (MERA) [5,6], a genuine simulation technique for strongly correlated systems, allows in each RG step for local unitary operations to be applied before the elimination of block basis states. The technique generates hence a more general class of states, referred to as MERA states; see Fig. 1.

Whereas the degrees of freedom of MERA and TTN states are organized in a hierarchical structure encoding correlations on different length scales, there exists a different class of so-called finitely correlated states where the degrees of freedom are organized in a strictly local manner. For $D = 1$ dimensional systems they are often referred to as matrix product states [7], and for $D \geq 1$ as tensor product Ansätze or projected entangled pair states (PEPS) [8]; Fig. 2. PEPS are the basis of powerful numerical techniques, such as the very successful density-matrix renormalization group [9].

In this Letter, we establish the surprising fact that, for $D > 1$, real-space RG, despite of the inherently hierarchical nature of the procedure, generates states that capture correlations by local degrees of freedom. More specifically, it is shown that MERA states form a subclass of PEPS, unifying both approaches. This also explains the failure of real-space RG for some situations for which merely anecdotal evidence had previously accumulated.

PEPS, TTN, and MERA are all tensor network states (TNS). In terms of an orthonormal product basis $|\boldsymbol{\sigma}\rangle = \bigotimes_{i=1}^N |\sigma_i\rangle$ for a lattice of N sites, TNS are of the form $\sum_{\boldsymbol{\sigma}}\psi_{\boldsymbol{\sigma}}|\boldsymbol{\sigma}\rangle$ where the expansion coefficients $\psi_{\boldsymbol{\sigma}}$ are encoded as a partially contracted set of tensors; Fig. 2. Recently, this notion has been generalized to the fermionic case [10]. For a PEPS, to each site i, a tensor A_i is assigned which has one physical index σ_i and further auxiliary indices—one for each nearest neighbor—which need to be contracted to obtain $\psi_{\boldsymbol{\sigma}}$; Fig. 2. For TTN and MERA,

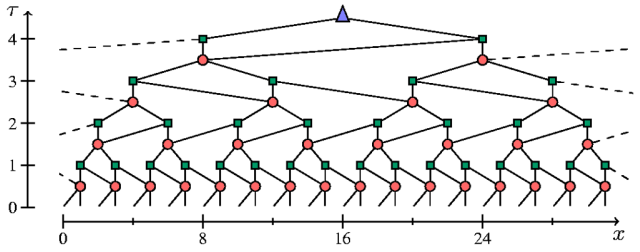

FIG. 1 (color online). A 1D MERA with linear branching ratio $b = 2$. Circles, squares, and the triangle denote tensors, the lines denote contractions of those tensors. The squares are isometries that map two local subsystems $\mathcal{H}_i^\tau$ and $\mathcal{H}_{i+1}^\tau$ into one $\mathcal{H}_{i/2}^{\tau+1}$ as in Kadanoff's block spin transformation. The circles denote unitary operators, disentanglers, reducing the entanglement between $\mathcal{H}_i^\tau \otimes \mathcal{H}_{i+1}^\tau$ and the rest of the system before the action of the isometry. Tensor positions are chosen according to Eq. (6) such that stacking of tensors is avoided.

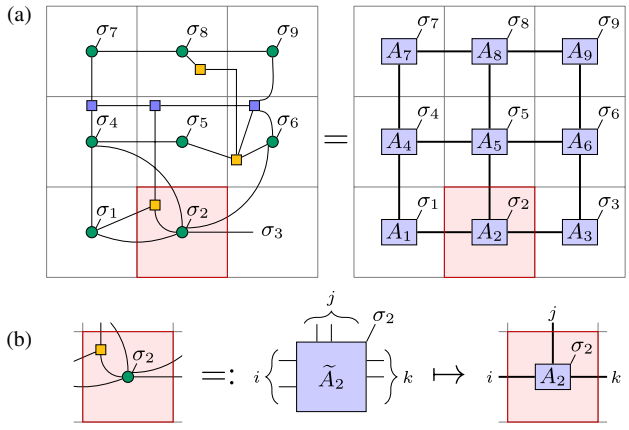

FIG. 2 (color online). (a) Procedure for mapping a TNS (left) to a 2D PEPS (right), by assigning tensors to lattice sites and contraction lines to paths on the lattice. (b) The elements of the PEPS tensors are determined by the elements of the tensors composing the TNS.

the tensors are arranged in a hierarchical pattern with the physical indices in the lowest layer; Fig. 1. The number of degrees of freedom of a TNS can be tuned by changing the number χ of values each auxiliary index runs over. Increasing χ for a fixed structure of the TNS enlarges the variational space, allowing for a more precise approximation to the exact ground state in a variational method, but increases computation costs. Hence, χ is called the refinement parameter of the TNS. The computational costs for efficient simulation techniques scale polynomially in χ.

For $D = 1$, TTN and MERA states can in general not be encoded efficiently as PEPS. There are MERA states with an entanglement entropy that scales logarithmically in the subsystem size [11], as occurring in critical models [12], whereas the entanglement entropy of 1D PEPS saturates for large subsystem sizes. In this respect, MERA states are more useful than PEPS for this case. For $D > 1$, however, our aforementioned result on real-space RG means in the tensor network language that MERA states with a refinement parameter χ can be mapped efficiently to PEPS such that the resulting PEPS refinement parameter χ_{PEPS} is some system-size independent function of χ. This also implies that $D > 1$ MERA states always obey the entanglement area law just as PEPS [12,13]. This behavior is shared by ground states of noncritical systems and critical bosons. Ground states of critical fermions, however, can violate the area law [12,14]. Consequently χ needs to be scaled polynomially in the system size in order to describe such critical fermionic systems accurately. Otherwise, the real-space RG schemes addressed here [3,5] are necessarily imprecise in that case. The remaining advantage of $D > 1$ MERA is that local observables and correlation functions can be evaluated efficiently, whereas, for PEPS, approximations are necessary. In this sense, MERA states simply form an efficiently contractible subclass of PEPS. This raises the question of whether any efficiently contractible tensor network automatically yields an area law which is, however, not the case. To show this, we construct an example of efficiently contractible TNS based on unitary quantum cellular automata (QCA). For a specific choice of the tensors, one obtains instances that violate the area law for generic bipartitions of the system.

General procedure for mapping TNS to PEPS.—All TNS can be mapped to PEPS, although not necessarily in an efficient manner. To map a TNS to a PEPS one can (a) assign each tensor of the TNS to a specific site of the physical lattice [15]

$$\mathcal{V}_{\mathrm{phys}} := \{0, \ldots, L-1\}^D \subset \mathbb{Z}^D, \tag{1}$$

and (b) for each contraction line that connects the tensors, decide on a specific path for that line on the edges $\mathcal{E}_{\mathrm{phys}}$ of the physical lattice,

$$\mathcal{E}_{\mathrm{phys}} := \{(\boldsymbol{r}, \boldsymbol{r}') \in \mathcal{V}_{\mathrm{phys}} \times \mathcal{V}_{\mathrm{phys}} | |\boldsymbol{r} - \boldsymbol{r}'|_1 = 1\}, \tag{2}$$

see Fig. 2. The tensors composing the PEPS are then obtained by introducing for each edge of the lattice an auxiliary vector space that is the tensor product of the vector spaces of all TNS contraction lines that traverse that edge. The elements of the PEPS tensor for site i are determined by the elements of all the TNS tensors that were assigned to site i. See Fig. 2(b).

Given a family of TNS for different linear system sizes L, a mapping of the TNS to PEPS is called efficient if there exists an upper bound χ_{PEPS} on the resulting PEPS refinement parameter that is independent of L. Applying the described mapping procedure for a 1D MERA state inevitably results in an inefficient mapping, i.e., in a PEPS refinement parameter χ_{PEPS} that diverges with the system size. This is not just a feature of the specific procedure. In [11], a family of 1D TTN states is constructed for which any mapping to PEPS necessarily requires χ_{PEPS} to diverge with the system size.

Qualitative argument.—The following argument motivates why an efficient mapping of MERA to PEPS should be possible for $D > 1$. Let us assign to each contraction line of the MERA state a finite cross section, e.g., equal to a^{D-1} with the lattice spacing a. Then one can ask what D-dimensional volume $V(\tau)$ the contraction lines of a certain layer τ connecting to layers with $\tau' \leq \tau$ cover. Those contraction lines of layer τ have length $\ell(\tau) \propto ab^{\tau}$, where b is the linear branching ratio of the MERA. The number of lattice cells in layer τ is $b^{(T-\tau)D}$; Fig. 1. Hence, the volume covered by the contraction lines of layer τ is $V(\tau) \propto a^{D-1}\ell(\tau)b^{(T-\tau)D} \propto b^{DT-(D-1)\tau}$. The density of the MERA contraction lines, or more precisely, a resulting estimate for the average number of contraction line paths traversing a unit cell of the physical lattice ($\tau = 0$) is hence

$$\log_{\chi}(\chi_{\mathrm{PEPS}}) \propto b^{-TD} \sum_{\tau=0}^{T} V(\tau) \propto \sum_{\tau=0}^{T} b^{-(D-1)\tau}$$

$$\Rightarrow \log_{\chi}(\chi_{\mathrm{PEPS}}) \propto \begin{cases} T & \text{for } D = 1 \\ \frac{1}{1-b^{-(D-1)}} & \text{for } D > 1, T \to \infty. \end{cases} \tag{3}$$

Note that for an edge traversed by n paths, one obtains an upper bound $\chi_{\mathrm{PEPS}} = \chi^n$ to the PEPS refinement parameter, i.e., $n = \log_\chi(\chi_{\mathrm{PEPS}})$. As $T = \log_b L$, 1D MERA with a fixed refinement parameter χ have according to Eq. (3) the potential to encode states with a logarithmic scaling of the entanglement entropy [11], as occurring in critical 1D systems. For $D > 1$, however, Eq. (3) suggests that there is enough space on the physical lattice to assign the MERA contraction lines to paths on the lattice such that the mapping to PEPS is efficient. That this is indeed possible is proven constructively in the following.

Preconditions for MERA states.—In order to show that the mapping presented in the following is efficient, it is necessary to exploit the defining properties of MERA states that correspond directly to features of the real-space RG and can be summarized as follows. (i) The MERA state is a TNS for a D-dimensional square lattice $(\mathcal{V}_{\mathrm{phys}}, \mathcal{E}_{\mathrm{phys}})$ consisting of L^D unit cells with

$$L = b^T. \tag{4}$$

(ii) The MERA consists of T layers of tensors labeled by $\tau = 1, \ldots, T$. (iii) There is an upper bound χ on the dimension of the vector spaces associated to the tensor indices, and an upper bound C_o on the order of each tensor. (iv) With each layer, we associate a coarse-grained square lattice $\mathcal{L}_\tau$ of $(L/b^\tau)^D$ cells of the physical lattice

$$\mathcal{L}_\tau := \{0, \ldots, L/b^\tau - 1\}^D \subset \mathbb{Z}^D, \tag{5}$$

and $\mathcal{L}_0 := \mathcal{V}_{\mathrm{phys}}$. Every cell of lattice $\mathcal{L}_\tau$ contains corresponding b^D cells of lattice $\mathcal{L}_{\tau-1}$. (v) There exists an assignment of the tensors of layer τ to cells of the lattice $\mathcal{L}_\tau$ such that the number of tensors inside a single cell is bounded from above by a constant C_t, and the distance of contracted tensors is bounded from above by C_r, where the distance of a tensor of layer τ to a tensor of layer $\tau' \leq \tau$ is defined as the L_1 distance of their corresponding cells in $\mathcal{L}_\tau$ [16]. (vi) For $|\tau - \tau'| > C_T$, there are no contractions between tensors of layer τ with tensors of layer τ'.

The upper bounds χ, C_o, C_t, C_r, and C_T are required to be independent of the system size L. [17] The stated conditions guarantee that the MERA features a so-called causal cone [6]. Hence, local observables can be evaluated efficiently if all tensors are chosen isometric. As we require only upper bounds on the MERA refinement parameter, the apparent restriction to square lattices is not essential. The conditions stated above are met for all typical MERA structures considered in the literature so far. See Fig. 3(a) for a 2D MERA with $b = 2$, $C_o = 8$, $C_t = 2$, and $C_T = 1$, for which one can reach $C_r = 2$.

Efficiently mapping MERA to PEPS for $D > 1$.—Let us explain a general scheme for mapping MERA states for $D > 1$ dimensional systems efficiently to PEPS. The preconditions listed above are assumed to be given. A simple procedure to assign the MERA tensors to certain lattice sites is to put the tensors of cell $\boldsymbol{n} \in \mathcal{L}_\tau$ of layer τ to the site $\boldsymbol{r}_\tau(\boldsymbol{n}) = b^\tau \boldsymbol{n} \in \mathcal{V}_{\mathrm{phys}}$. The problem with this approach is that one generates stacks of tensors at certain lattice sites, i.e., there exist positions $\boldsymbol{r} \in \mathcal{V}_{\mathrm{phys}}$ to which a number of tensors is assigned that is not independent of the lattice size. For example, at site $\boldsymbol{r} = (0, \ldots, 0)$ a number of $\propto T = \log_b L$ tensors accumulate. Further stacks of tensors with height $\propto T'$ accumulate at lattice sites with coordinates $b^{T'}(1, \ldots, 1)$. It is necessary to avoid such stacks of tensors, because they imply in general that χ_{PEPS} diverges with the system-size. Stacks can be avoided by shifting the allowed tensor positions for different layers relative to each other. One possible such choice for $\boldsymbol{r}_\tau(\boldsymbol{n})$ is

$$\boldsymbol{r}_\tau(\boldsymbol{n}) = b^\tau \boldsymbol{n} + b^{\tau-1}\boldsymbol{e} \in \mathcal{V}_{\mathrm{phys}} \quad \text{with} \quad \boldsymbol{n} \in \mathcal{L}_\tau \tag{6}$$

and $\boldsymbol{e} := (1, \ldots, 1) \in \mathbb{Z}^D$ as demonstrated in Fig. 1. With this choice, two tensors can end up at the same site only if they belong to the same layer $\mathcal{L}_\tau$ and the same lattice cell $\boldsymbol{n}$ within that layer. The possible tensor positions of layers τ form disjoint sublattices $\mathcal{V}_\tau$ of the physical lattice.

$$\mathcal{V}_\tau := \{\boldsymbol{r}_\tau(\boldsymbol{n}) | \boldsymbol{n} \in \mathcal{L}_\tau\} \subset \mathcal{V}_{\mathrm{phys}}, \ \mathcal{V}_\tau \cap \mathcal{V}_{\tau' \neq \tau} = \varnothing.$$

All coordinates r_i of $\boldsymbol{r} \in \mathcal{V}_\tau$ have a b-adic valuation of $\tau - 1$, where the b-adic valuation $v_b(n)$ of an integer n is defined such that $v_b(n) = \tau$ iff τ is the largest integer such that $n \bmod b^\tau = 0$, for example, $v_2(12) = 2$. Avoiding stacks of tensors is not sufficient for an efficient PEPS encoding. In $D = 1$, all contraction lines are assigned to paths that necessarily stack up on the x axis, Fig. 1. This stacking of the paths can be avoided in $D > 1$ by assigning contraction lines between tensors of layers τ and τ' to paths that are restricted to edges from subgrids $\mathcal{E}_\tau$ and $\mathcal{E}_{\tau'}$ and that are shortest paths with respect to the L_1 distance on $\mathcal{V}_\tau \cup \mathcal{V}_{\tau'}$. Here, a grid $\mathcal{E}_\tau$ is defined as the subset of physical edges connecting nearest neighbors of the lattice $\mathcal{V}_\tau$ on straight lines; see Fig. 4.

$$\mathcal{E}_\tau := \{(\boldsymbol{r}, \boldsymbol{r} + \boldsymbol{e}_i) \in \mathcal{E}_{\mathrm{Phys}} | v_b(r_j) = \tau - 1 \ \forall \ j \neq i\}$$

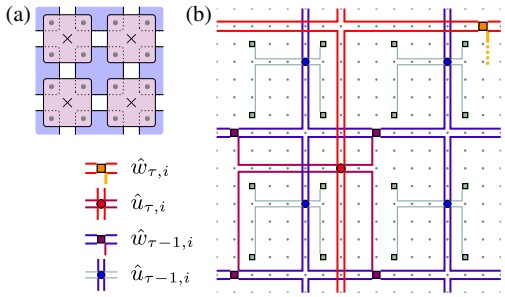

FIG. 3 (color online). (a) Unit cell of a specific 2D MERA state. With each layer, corresponding to a single RG step, unitary disentanglers are applied that reduce the entanglement between blocks of 2×2 sites with the rest of the system. Then, an isometry maps from those 2×2 sites (dots) into one (crosses). (b) Mapping of this MERA state to a PEPS. The diagram shows the assignment of two layers of the MERA, composed of disentanglers $\hat{u}$ and isometries $\hat{w}$, to the physical lattice.

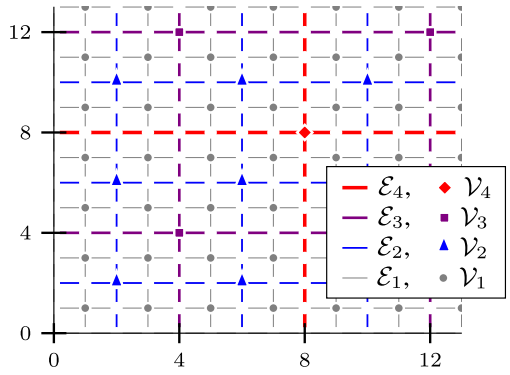

FIG. 4 (color online). Disjoint sublattices $\mathcal{V}_\tau \subset \mathcal{V}_{\text{phys}}$ to which MERA tensors are assigned and disjoint subsets of edges $\mathcal{E}_\tau \subset \mathcal{E}_{\text{phys}}$ to which MERA contraction lines are assigned for $b = 2$. In our construction, the paths assigned to contraction lines from tensors of layer 2 to tensors in layer 3 are, e.g., restricted to edges from $\mathcal{E}_2 \cup \mathcal{E}_3$.

with $[\boldsymbol{e}_i]_j = \delta_{i,j}$. Hence $\mathcal{E}_\tau \cap \mathcal{E}_{\tau'} = \varnothing \ \forall_{\tau \neq \tau'}$. For this choice of tensor positions and paths of MERA contraction lines an upper bound for the resulting PEPS refinement parameter χ_{PEPS} follows: Contraction lines assigned to an edge $e = (\boldsymbol{r}, \boldsymbol{r}') \in \mathcal{E}_\tau$ contract tensors of layer τ with tensors of layers τ' where $|\tau' - \tau| \leq C_T$. For a layer τ' with $\tau' > \tau$, tensors from at most $(2C_r)^D$ cells of $\mathcal{L}_{\tau'}$ around the cell corresponding to site $\boldsymbol{r}$ can have contraction line paths traversing edge e. From the layers τ' with $\tau' \leq \tau$, tensors of at most $(2C_r)^D \sum_{t=0}^{C_T} b^{Dt}$ cells can contribute. Thus, the number of contraction line paths traversing edge e and hence $\log_\chi(\chi_{\text{PEPS}})$ are bounded from above by

$$\log_\chi(\chi_{\text{PEPS}}) \leq (2C_r)^D (C_T + b^{D(C_T+1)}) C_t C_o. \quad (7)$$

As this upper bound is independent of the system size, the presented mapping of MERA to PEPS is efficient.

The scheme displayed in Fig. 3(b) for mapping the 2D MERA defined in Fig. 3(a) to a PEPS results in the PEPS refinement parameter $\chi_{\text{PEPS}} = \chi^6$. In the supplement [11], the notion of a refined PEPS is introduced which allows for a favorable scaling, $\chi_{\text{PEPS}}^{\text{refined}} = \chi^2$ in this case.

QCA violating the area law.—Let us point out that, also for $D > 1$, there exist efficiently contractible TNS that violate the entanglement area law; more details in [11]. Consider a QCA consisting of $2T$ layers, where in every layer, $(L/2)^D$ local unitary gates are applied to plaquettes of $2 \times \ldots \times 2$ sites each. For $T = (\log L)^{1/D}$, the computation cost for the evaluation of local observables with respect to such QCA is polynomial in L, namely $O(L^2 (\log L)^{1/D})$. At the same time, one finds for a suitable choice of the unitary gates and generic choices for subsystems $\mathcal{A}_L$ with $\text{Vol}\mathcal{A}_L \propto L^D$ an entanglement entropy of $S_{\mathcal{A}_L} = \Omega(L^{D-1} (\log L)^{1/D})$ which violates the area law.

Conclusion.—In this Letter, we have shown that MERA states for $D > 1$ can be efficiently encoded as PEPS. From a physical perspective, the result implies that real-space RG techniques, despite the scale-invariant features of the TNS they generate, give rise to states that can be encoded with local degrees of freedom. As a corollary, it follows that $D > 1$ MERA states obey the area law for the entanglement entropy [12,13]. Consequently, the refinement parameter χ needs to be scaled polynomially in the system size in order to describe $D > 1$ critical fermionic systems accurately. Otherwise, the real-space RG schemes addressed here [3,5] are imprecise for such systems.

We thank V. Nesme, A. Flesch, and G. Vidal for fruitful discussions. This work has been supported by the EU (MINOS, QESSENCE, COMPAS), and the EURYI.

[1] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975); F. J. Wegner, Phys. Rev. B **5**, 4529 (1972).
[2] L. Kadanoff, Physics **2**, 263 (1966).
[3] R. Jullien, J. Fields, and S. Doniach, Phys. Rev. Lett. **38**, 1500 (1977); S. D. Drell, M. Weinstein, and S. Yankielowicz, Phys. Rev. D **16**, 1769 (1977).
[4] Y.-Y. Shi, L.-M. Duan, and G. Vidal, Phys. Rev. A **74**, 022320 (2006).
[5] G. Vidal, Phys. Rev. Lett. **99**, 220405 (2007).
[6] G. Vidal, Phys. Rev. Lett. **101**, 110501 (2008).
[7] M. Fannes, B. Nachtergaele, and R. F. Werner, Commun. Math. Phys. **144**, 443 (1992); S. Rommer and S. Östlund, Phys. Rev. B **55**, 2164 (1997).
[8] H. Niggemann, A. Klümper, and J. Zittartz, Z. Phys. B **104**, 103 (1997); T. Nishino, K. Okunishi, Y. Hieida, N. Maeshima, and Y. Akutsu, Nucl. Phys. B **575**, 504 (2000); M. A. Martín-Delgado, M. Roncaglia, and G. Sierra, Phys. Rev. B **64**, 075117 (2001); F. Verstraete and J. I. Cirac, arXiv:cond-mat/0407066.
[9] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992); U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).
[10] C. V. Kraus, N. Schuch, F. Verstraete, and J. I. Cirac, arXiv:0904.4667; P. Corboz, G. Evenbly, F. Verstraete, and G. Vidal, Phys. Rev. A **81**, 010303(R) (2010); C. Pineda, T. Barthel, and J. Eisert, Phys. Rev. A **81**, 050303 (R) (2010); T. Barthel, C. Pineda, and J. Eisert, Phys. Rev. A **80**, 042333 (2009).
[11] See supplementary material at http://link.aps.org/supplemental/10.1103/PhysRevLett.105.010502 for a mapping to refined PEPS and for TNS that exceed the area law.
[12] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, Rev. Mod. Phys. **80**, 517 (2008); J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys. **82**, 277 (2010).
[13] G. Vidal, arXiv:quant-ph/0610099v1.
[14] M. M. Wolf, Phys. Rev. Lett. **96**, 010404 (2006); D. Gioev and I. Klich, Phys. Rev. Lett. **96**, 100503 (2006); T. Barthel, M.-C. Chung, and U. Schollwöck, Phys. Rev. A **74**, 022329 (2006); W. Li *et al.*, Phys. Rev. B **74**, 073103 (2006); M. Cramer, J. Eisert, and M. B. Plenio, Phys. Rev. Lett. **98**, 220603 (2007).
[15] For clarity we restrict to square lattices.
[16] The coarse-graining maps cells from $\mathcal{L}_{\tau'}$ to cells in $\mathcal{L}_{\tau > \tau'}$.
[17] Actually, a system-size independent upper bound on the combination occurring in Eq. (7) is sufficient.

Real-space renormalization yields finitely correlated states: EPAPS

Thomas Barthel,[1] Martin Kliesch,[1] and Jens Eisert[1,2]

[1]*Institute for Physics and Astronomy, Potsdam University, 14476 Potsdam, Germany*
[2]*Institute for Advanced Study Berlin, 14193 Berlin, Germany*

(Dated: March 01, 2010)

In this supplementary information, for the mapping of *multi-scale entanglement renormalization ansatz* (MERA) states for $D > 1$ spatial dimensions to *projected entangled pair states* (PEPS), it is explained how the resulting PEPS bond dimension can be reduced considerably by the use of *refined* PEPS. Furthermore, to show explicitly that 1D MERA states can in general not be mapped efficiently to 1D PEPS, a family of MERA states is constructed for which the entanglement entropy grows logarithmically in the linear system size for a suitable bipartition of the system. To show that, for $D > 1$, there exist other efficiently contractible schemes violating the area law, unlike MERA, we construct a family of efficiently contractible *tensor network states* (TNS) based on a unitary *quantum cellular automata*. For a specific choice of the composing tensors, one obtains instances that violate the area law.

I. LOWER BOND DIMENSIONS BY PEPS REFINEMENT

In the mapping of $D > 1$ MERA to PEPS, as described in the Letter, all tensors of a given lattice cell $\boldsymbol{n} \in \mathcal{L}_\tau$ of layer τ of the MERA are assigned to the same physical lattice site $\boldsymbol{r}_\tau(\boldsymbol{n}) \in \mathcal{V}_{\text{phys}}$.Therefore, a considerable number of contraction lines that start at the tensors of a given cell $\boldsymbol{n}$ may traverse the same edges around $\boldsymbol{r}_\tau(\boldsymbol{n})$ and cause hence a relatively high χ_{PEPS}. While this is unproblematic for the purpose of proving the existence of an efficient mapping, the situation can be improved for numerical purposes, e.g., by introducing for each site of the physical layer $b^{\delta\tau D} - 1$ auxiliary sites with $\delta\tau > 0$, resulting in *refined* lattices $\mathcal{V}'_{\text{phys}}$ and $\mathcal{V}'_\tau$. The corresponding *refined PEPS* has tensors for the physical sites and tensors for the auxiliary sites, where the latter ones carry no physical indices. The sites from $\mathcal{V}'_\tau$ allowed for tensors of layer τ of the MERA are then defined as

$$\boldsymbol{r}_\tau(\boldsymbol{n},\boldsymbol{m}) := b^{\tau+\delta\tau}\boldsymbol{n} + b^\tau \boldsymbol{m} + b^{\tau-1}\boldsymbol{e} \quad \in \mathcal{V}'_{\text{phys}} \qquad (1)$$
$$\text{with} \quad \boldsymbol{n} \in \mathcal{L}_\tau \quad \text{and} \quad m_i \in \{0,\dots,b^{\delta\tau}-1\}^D.$$

Lattice cells in layer τ are again labeled by $\boldsymbol{n} \in \mathcal{L}_\tau$ and $\boldsymbol{m}$ labels now the possible positions for tensors inside that cell. Due to the refinement of the physical lattice, the resulting PEPS consists of $|\mathcal{V}'_{\text{phys}}| = b^{\delta\tau D}|\mathcal{V}_{\text{phys}}|$ instead of $|\mathcal{V}_{\text{phys}}|$ tensors. A refined PEPS is transformed to a "normal" PEPS by contracting the PEPS tensors for the $b^{\delta\tau D} - 1$ auxiliary sites with the tensor for the corresponding physical site, resulting in $\chi_{\text{PEPS}} = (\chi^{\text{refined}}_{\text{PEPS}})^{b^{\delta\tau(D-1)}}$.

The scheme for mapping the 2D MERA defined in Fig. 1a to a PEPS results in the PEPS refinement parameter $\chi^{\text{refined}}_{\text{PEPS}} = \chi^2$ if one uses a refined PEPS with $\delta\tau = 1$ where isometries are located at $\boldsymbol{m} = (1,1)$ and disentanglers at $\boldsymbol{m} = (0,0)$, according to Eq. (1). Each edge of the grid is traversed by at most two contraction line paths in this case. Using, instead, the tensor coordinates according to the most simple scheme, described in the Letter, yields the $\chi_{\text{PEPS}} = \chi^6$.

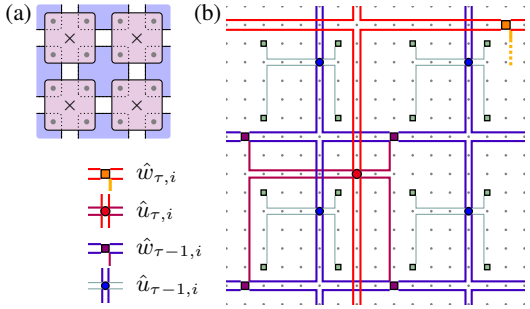

Figure 1: (a) Unit cell of a specific 2D MERA state with linear branching ration $b = 2$. With each layer, corresponding to a single renormalization step, unitary *disentanglers* are applied that reduce the entanglement between blocks of 2×2 sites with the rest of the system. Then, an isometry maps from those 2×2 sites (dots) into one (crosses). (b) Mapping of this MERA state to a refined PEPS with $\delta\tau = 1$. The diagram shows the assignment of two layers of the MERA, composed of disentanglers $\hat{u}$ and isometries $\hat{w}$, to the physical lattice. Each edge of the lattice is traversed by at most two contraction line paths. The resulting PEPS has $\chi_{\text{PEPS}} = \chi^2$.

II. 1D TTN AND MERA STATES CANNOT BE MAPPED EFFICIENTLY TO 1D PEPS

To show that a 1D MERA can in general not be mapped to a 1D PEPS with a bond dimension χ_{PEPS} that is independent of the system size L, we construct a family of *graph tree tensor network states* [1, 2] for which the entanglement entropy grows logarithmically with L for a suitable bipartition of the system; see, e.g., Refs. [3–5] for a numerical analysis. We choose a *tree tensor network* (TTN) state, so a MERA state without disentanglers, with (partial) isometries mapping from two qubits to one, i.e., $\chi = 2$ and $b = 2$. Each representative of the family of states is defined on $L = 2^T$ sites $\{0,\dots,L-1\}$ with a positive odd integer T. For simplicity, the TTN is embedded into the entire lattice of L sites, and each isometry is considered as a unitary having one input from the previous layer and one input $|0\rangle$; see Fig. 2. The state of the the top layer $\tau = T$ and the two-site gates $\hat{u}$ are defined as

$$|\psi_T\rangle := |0\rangle^{\otimes L} \quad \text{and} \quad \hat{u} := e^{-i\pi\hat{X}\otimes\hat{X}/4}, \qquad (2)$$

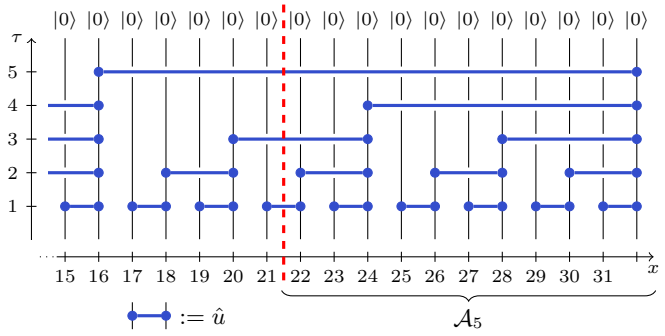

Figure 2: The 1D TTN state discussed in Sec. II for $T = \log L = 5$. For the chosen subsystems $\mathcal{A}_T$, the entanglement entropy is logarithmic in the system size, $S_{\mathcal{A}_T} = (T+1)/2$.

$\hat{X}$ and $\hat{Z}$ denoting Pauli matrices, and $|0\rangle$ being an eigenstate of $\hat{Z}$. For layers $\tau = 1, \ldots, T$ the state $|\psi_{\tau-1}\rangle$ is generated from $|\psi_\tau\rangle$ by applying gates $\hat{u}$ to sites $2^\tau(k-1/2)-1$ and $2^\tau k - 1$ for $k = 1, \ldots, 2^{T-\tau}$. The entanglement is computed for the subsystem $\mathcal{A}_T$ consisting of the last p_T sites, where $p_1 := 1$ and $p_{T+2} := 4p_T - 1$. Since all the gates, specified in Eq. (2), are mutually commuting, all gates that are supported entirely on $\mathcal{A}_T$ or entirely on its complement $\mathcal{A}_T^\perp$ can be disregarded for the computation of the entanglement entropy

$$S_{\mathcal{A}_T} = -\operatorname{Tr} \hat{\rho}_{\mathcal{A}_T} \log \hat{\rho}_{\mathcal{A}_T} \text{ with } \hat{\rho}_{\mathcal{A}_T} = \operatorname{Tr}_{\mathcal{A}_T^\perp} |\psi_T\rangle\langle\psi_T|.$$

The locations p_T of the bipartitions are chosen such that, for odd T, exactly $(T+1)/2$ gates act across the cut. Since each such gate generates one pair of maximally entangled qubits, one obtains $S_{\mathcal{A}_T} = (T+1)/2$ which is logarithmically divergent in the system size L and implies that any 1D PEPS encoding of the given graph state requires a χ_{PEPS} that diverges with L.

III. EFFICIENTLY CONTRACTIBLE TNS THAT VIOLATE THE AREA LAW

As shown in the Letter, unlike for $D = 1$ spatial dimensions, MERA states for $D > 1$ obey the entanglement area law and not a log-area law, $S_{\mathcal{A}_L} = \Omega(L^{D-1} \log L)$, as it occurs for critical fermionic models [6–12]. This raises the question of whether any efficiently contractible tensor network automatically yields an area law. This is however not the case. In order to show this, we construct, for a D-dimensional cubic lattice of L^D sites, a family of efficiently contractible TNS based on a unitary *quantum cellular automata* (QCA). For a specific choice of the constituting tensors, one obtains instances that violate the area law for generic bipartitions of the system.

Let us consider a QCA consisting of T layers $\tau = 1, \ldots, T$, where each layer consists of two sublayers. With the first sublayer, $\hat{K}_1$ is applied which consists of $(L/2)^D$ local unitary gates $\hat{s}$ supported on plaquettes of $2 \times \cdots \times 2$ sites each. The operator $\hat{K}_2$ for the second sublayer is identical to $\hat{K}_1$ except for a relative shifting of the gate positions by

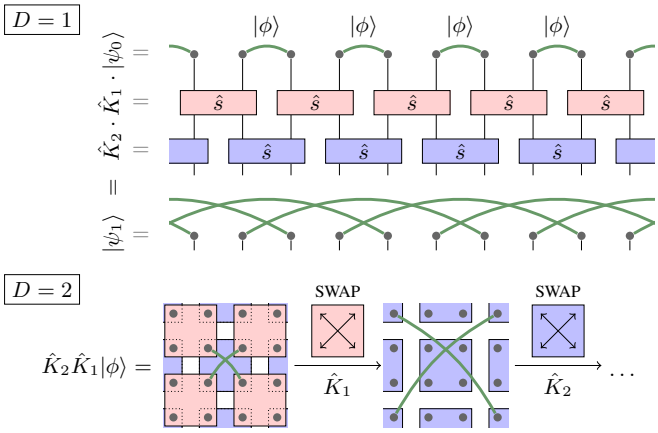

Figure 3: The graph state QCA discussed in Sec. III for $D = 1$ and $D = 2$. Each layer of the QCA moves the maximally entangled qubits that initially are located on next nearest neighbor sites two steps further apart from each other. This increases the entanglement entropy for a given bipartition of the system in every step. For an appropriate choice of the number of layers, the states violate the entanglement area law while still being efficiently contractible.

$(1, \ldots, 1)$. Therefore, periodic boundary conditions are imposed, and L is required to be even. The initial state $|\psi_0\rangle$ is a product state of $(L/2)^D$ plaquette states $|\phi\rangle$ for $2 \times \cdots \times 2$ sites each, where the plaquette positions coincide with those of the gates in $\hat{K}_2$. The plaquette states are product states of 2^{D-1} maximally entangled pairs of qubits sitting each at the ends of the plaquette diagonals, e.g., $|\phi\rangle = \hat{u}_{0,1}|0\rangle^{\otimes 2}$ for $D = 1$, where $\hat{u}$ is chosen according to Eq. (2) and the indices label the sites the gate acts on. For $D = 2$, the plaquette state is $|\phi\rangle = \hat{u}_{(0,0),(1,1)}\hat{u}_{(1,0),(0,1)}|0\rangle^{\otimes 4}$. The plaquette operators $\hat{s}$, composing the $\hat{K}_i$, are chosen as products of swap operators $\hat{S}_{i,j}|\sigma_i\sigma_j\rangle = |\sigma_j\sigma_i\rangle$ that act similarly on the qubits at the ends of the plaquette diagonals. For example, $\hat{s} = \hat{S}_{(0,0),(1,1)}\hat{S}_{(1,0),(0,1)}$ for $D = 2$.

The QCA layers and the initial state are invariant under translations by two sites and rotations by $\pi/4$ and so are all states $|\psi_\tau\rangle := (\hat{K}_2\hat{K}_1)^\tau|\psi_0\rangle$. Like $|\psi_0\rangle$, every state $|\psi_\tau\rangle$ is a product state of $2^{D-1}(L/2)^D$ maximally entangled qubit pairs. If two entangled qubits have positions $\boldsymbol{r} \pm \Delta\boldsymbol{r}$ in $|\psi_\tau\rangle$ there is exactly one corresponding entangled qubit pair at positions $\boldsymbol{r} \pm (1 + \frac{2\sqrt{D}}{|\Delta\boldsymbol{r}|})\Delta\boldsymbol{r}$ in $|\psi_{\tau+1}\rangle$. Applying one QCA layer after another, distances of entangled qubits increase by $4\sqrt{D}$ in each step, e.g., $\hat{K}_1\hat{K}_2\hat{u}_{(2,2),(3,3)} = \hat{u}_{(0,0),(5,5)}$, see Fig. 3.

For generic choices for bipartitions of the system into two parts, $\mathcal{A}_L \subset \mathcal{V}_{\text{phys}}$ and its complement, where $\mathcal{A}_L$ is connected and has a volume $\propto L^D$, the corresponding entanglement entropy will violate the area law if the number of layers, T, is chosen appropriately. Consider as an example the bipartition with $\mathcal{A}_L = \{0, \ldots, L/2-1\} \times \{0, \ldots, L-1\}^{D-1}$. The subsystem boundary is formed by the planes $\{\boldsymbol{r}|r_1 = 0\}$ and $\{\boldsymbol{r}|r_1 = L/2\}$. Each plane is crossed by a number of different pairs of entangled qubits that is proportional to its area and to

T, as long as $L/2 > 2T$. Consequently,

$$S_{\mathcal{A}_L}(T) = \Omega(L^{D-1}T). \tag{3}$$

For a choice $T \propto \log L$, this yields a log-area law $S_{\mathcal{A}_L} = \Omega(L^{D-1} \log L)$. But an upper bound on the computation cost for the evaluation of local observables with respect to QCA states of the given class (with arbitrary $\hat{u}$) is of order $O(2^{2T^D}T)$, i.e., $O(L^{2(\log L)^{D-1}} \log L)$ for $T \propto \log L$. This cost is not polynomial in L and the QCA are for this T hence not efficiently contractible in an obvious fashion. However, the computation cost is, of order $O(L^2(\log L)^{1/D})$ for the choice $T = (\log L)^{1/D}$, i.e., polynomial in L. The resulting entanglement entropy is $S_{\mathcal{A}_L} = \Omega(L^{D-1}(\log L)^{1/D})$ which violates the area law by the sublogarithmic factor $(\log L)^{1/D}$. Note also that even a QCA consisting of a single layer of $k \times \cdots \times k$ plaquettes, where k is allowed to grow as $k \propto (\log L)^{1/D}$, can also violate the entanglement area law, albeit only for specific bipartitions of the system, while being efficiently contractible.

[1] M. Hein, J. Eisert, and H. J. Briegel, Phys. Rev. A **69**, 062311 (2004).
[2] Y.-Y. Shi, L.-M. Duan, and G. Vidal, Phys. Rev. A **74**, 022320 (2006).
[3] G. Vidal, Phys. Rev. Lett. **101**, 110501 (2008).
[4] G. Evenbly and G. Vidal, arXiv:0710.0692 (2007).
[5] C. M. Dawson, J. Eisert, and T. J. Osborne, Phys. Rev. Lett. **100**, 130501 (2008).
[6] M. M. Wolf, Phys. Rev. Lett. **96**, 010404 (2006).
[7] D. Gioev and I. Klich, Phys. Rev. Lett. **96**, 100503 (2006).
[8] T. Barthel, M.-C. Chung, and U. Schollwöck, Phys. Rev. A **74**, 022329 (2006).
[9] W. Li, L. Ding, R. Yu, T. Roscilde, and S. Haas, Phys. Rev. B **74**, 073103 (2006).
[10] M. Cramer, J. Eisert, and M. B. Plenio, Phys. Rev. Lett. **98**, 220603 (2007).
[11] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, Rev. Mod. Phys. **80**, 517 (2008).
[12] J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys. **82**, 277 (2010).

Quasilocality and Efficient Simulation of Markovian Quantum Dynamics

Thomas Barthel and Martin Kliesch
Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany
Institute for Physics and Astronomy, University of Potsdam, 14476 Potsdam, Germany
(Received 21 November 2011; published 5 June 2012)

We consider open many-body systems governed by a time-dependent quantum master equation with short-range interactions. With a generalized Lieb-Robinson bound, we show that the evolution in this very generic framework is quasilocal; i.e., the evolution of observables can be approximated by implementing the dynamics only in a vicinity of the observables' support. The precision increases exponentially with the diameter of the considered subsystem. Hence, time evolution can be simulated on classical computers with a cost that is independent of the system size. Providing error bounds for Trotter decompositions, we conclude that the simulation on a quantum computer is additionally efficient in time. For experiments and simulations in the Schrödinger picture, our result can be used to rigorously bound finite-size effects.

DOI: 10.1103/PhysRevLett.108.230504 PACS numbers: 03.67.Ac, 02.60.Cb, 03.65.Yz, 89.70.Eg

In Lorentz-invariant theories, a maximum speed for the propagation of information is, by construction, the speed of light. In nonrelativistic quantum theory, the existence of a maximum propagation speed results more indirectly and for different reasons. For nonpathological models, this maximum speed is much smaller than the speed of light. The seminal paper by Lieb and Robinson [1] and further contributions like [2–13] cover isolated systems.

Here, we consider the evolution of a more general and, experimentally, extremely relevant class of systems—open quantum many-body systems governed by a quantum master equation [14,15] with short-range Liouvillians that are allowed to be time dependent. Prominent experimental examples are presented in Refs. [16–20], and recent theoretical advances on quantum computation, nonequilibrium steady states, and phase transitions in open systems can, for example, be found in Refs. [21–24]. Going beyond the existence of a finite maximum propagation speed and the existence of a well-defined thermodynamic limit [1,25], we show that the time evolution of such systems is quasilocal. This means that, up to an exponentially small error, the diameter of the support of any evolved local observable grows at most linearly in time, or, put differently, that the evolution of the local observable can be approximated to arbitrary precision by applying the propagator of a spatially truncated version of the Liouvillian, as seen in Fig. 1(b). For the special case of isolated systems, where the evolution is given by a unitary transformation, the corresponding question has been addressed in Ref. [9]. As a tool for the proof of quasilocality, we derive and employ a Lieb-Robinson-type bound very similar to the recent results of Poulin [26] and Nachtergaele *et al.* [25]. All constants in the bounds are given explicitly in terms of the system parameters.

The quasilocality of Markovian quantum dynamics has several crucial consequences. It implies that the evolution of observables with a finite spatial support can be simulated efficiently on classical computers, in the sense that the computation cost is independent of the system size, irrespective of the desired accuracy. This can, for example, be exploited in an exact diagonalization approach for a sufficiently large vicinity of the support of the considered observable, as illustrated in Fig. 1(b). For more sophisticated simulation techniques, we provide, in extension of Ref. [27], error bounds for Trotter decompositions [28] of the subsystem propagator into a circuit of local channels, as shown in Fig. 1(c). The Trotter error is polynomial in the time, at most linear in the size of the time step, and can hence be made arbitrarily small. Importantly, the subsystem Trotter decompositions allow for the efficient simulation of the time evolution on a quantum computer as envisaged by Feynman. For any required accuracy, the simulation can be implemented with a cost that is independent of the system size and polynomial in the time.

Experimental and numerical physicists who study nonequilibrium systems in the Schrödinger picture can use our result on quasilocality to rigorously bound finite-size effects. This is, for example, relevant for experiments with ultracold atoms in optical lattices [29] and numerical investigations employing time-dependent density-matrix renormalization group methods [30–33].

Setting.—Let us consider lattice systems, where each site $z \in \Lambda$ is associated with a local Hilbert space $\mathcal{H}_z$. Subsystem Hilbert spaces are denoted by $\mathcal{H}_V := \bigotimes_{z\in V}\mathcal{H}_z \forall_{V\subset\Lambda}$ and $\mathcal{H} := \mathcal{H}_\Lambda$. Let $\rho(t)$ denote the system state at time t. Markovian dynamics of an open quantum system, i.e., the evolution under a linear differential equation that generates a completely positive and trace-preserving map for ρ, can always be written in the form of a Lindblad equation [34–36]:

$$\partial_t\rho = -i[H,\rho] + \sum_\nu \left(L_\nu \rho L_\nu^\dagger - \frac{1}{2}(L_\nu^\dagger L_\nu \rho + \rho L_\nu^\dagger L_\nu) \right),$$

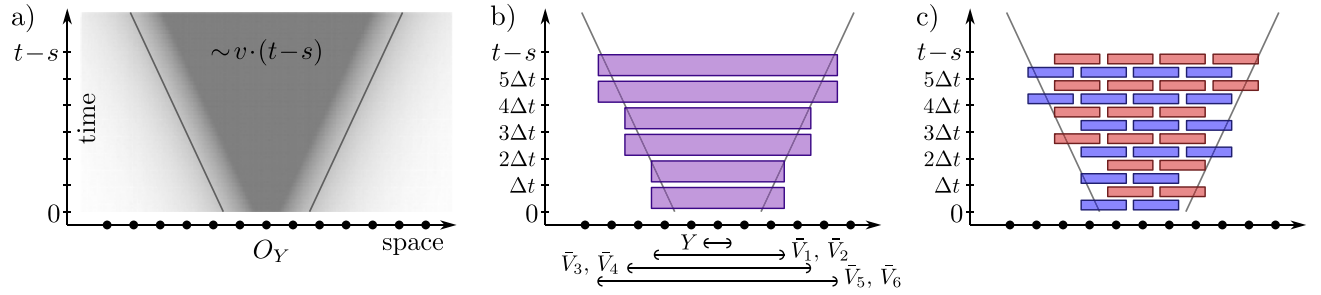

FIG. 1 (color online). (a) An evolved local operator $\tau(s,t)O_Y$ behaves almost like the identity outside its associated space-time cone. (b) Approximating $\tau(s,t)O_Y$ by application of subsystem propagators to O_Y. The errors decrease exponentially with the subsystem sizes. (c) For one-dimensional systems, approximating $\tau(s,t)O_Y$ by a Trotter decomposition yields an error scaling as $(t-s)^2\Delta t$. Note that the Trotter circuit can be trimmed off at the boundary of the Lieb-Robinson space-time cone.

where the arbitrary Lindblad operators L_ν and the Hermitian Hamiltonian H may depend on time. This equation captures, for example, in the framework of the Born-Markov approximation, the evolution of a system that interacts with an environment [14,15] and isolated systems as a special case. Let us switch from the Schrödinger picture, where expectation values are evaluated according to $\langle O\rangle_{s\to t} = \mathrm{Tr}[\rho(t)O]$ with $\rho(s)=\rho$ to the Heisenberg picture, where $\langle O\rangle_{s\to t} = \mathrm{Tr}[\rho O(s)]$ with $O(t)=O$. The corresponding time-dependence of an observable $O(s)\in\mathcal{B}(\mathcal{H})$ is then given by the quantum master equation $\partial_s O(s) = -\mathcal{L}(s)O(s)$, where $\mathcal{L}(t)\in\mathcal{B}(\mathcal{B}(\mathcal{H}))$ is a superoperator, the so-called Liouvillian, with the Lindblad representation

$$\mathcal{L}\,O = i[H,O] + \sum_\nu \left(L_\nu^\dagger O L_\nu - \frac{1}{2}(L_\nu^\dagger L_\nu O + O L_\nu^\dagger L_\nu)\right).$$

The set of Liouvillians with spatial support $V\subset\Lambda$ will be denoted by $\mathbb{L}_V \subset \mathcal{B}(\mathcal{B}(\mathcal{H}_V))$.

In order to be able to use Lieb-Robinson bound techniques, we need to restrict ourselves to Liouvillians with norm-bounded short-range interaction terms. Let us hence assume that $\mathcal{L}$ is a sum of local Liouville terms ℓ_Z with norm bound $|\ell|$, maximum range a, and a maximum number $\mathcal{Z}$ of nearest neighbors [37]. Specifically,

$$\mathcal{L}(t) = \sum_{Z\subset\Lambda} \ell_Z(t), \qquad \ell_Z(t)\in\mathbb{L}_Z, \tag{1}$$

$$|\ell| := \sup_{t,Z\subset\Lambda} \|\ell_Z(t)\|, \tag{2}$$

$$a := \sup_{Z:\ell_Z\neq 0} \mathrm{diam}(Z), \tag{3}$$

$$\mathcal{Z} := \max_{Z:\ell_Z\neq 0} |\{Z'\subset\Lambda | \ell_{Z'}\neq 0, Z'\cap Z\neq\varnothing\}|, \tag{4}$$

where $\mathrm{diam}(Z) := \max_{x,y\in Z} d(x,y)$ is the diameter of Z and d is a metric on the lattice Λ. In Eq. (2), we have used the superoperator norm [38,39] defined by $\|T\| := \sup_{O\in\mathcal{B}(\mathcal{H})}\|TO\|/\|O\|$. In the Heisenberg picture, this is the physically relevant norm as induced by the operator norm $\|O\|$. For notational convenience, we define for every subsystem $V\subset\Lambda$ the corresponding extension $\bar{V}$, volume $\mathrm{Vol}(V)$, and truncated Liouvillian $\mathcal{L}_V$,

$$\bar{V} := \bigcup_{\substack{Z:\ell_Z\neq 0\\ Z\cap V\neq\varnothing}} Z, \tag{5}$$

$$\mathrm{Vol}(V) := |\{Z\subset V | \ell_Z\neq 0\}|, \tag{6}$$

$$\mathcal{L}_V(t) := \sum_{Z\subset V} \ell_Z(t). \tag{7}$$

Propagators $\tau_V(s,t)$ are superoperators that map observables to time-evolved observables. They are defined as the unique solutions of

$$\partial_s \tau_V(s,t) = -\mathcal{L}_V(s)\tau_V(s,t), \quad \tau_V(t,t) = \mathrm{id} \quad \forall_{s\le t}. \tag{8}$$

With $\tau(s,t) := \tau_\Lambda(s,t)$, one has indeed $O(s) = \tau(s,t)O(t)$. Propagators obey the composition rule $\tau(r,s)\tau(s,t) = \tau(r,t)$ $\forall_{r\le s\le t}$. Furthermore [38], the derivative with respect to the second time argument is given by

$$\partial_t \tau_V(s,t) = \tau_V(s,t)\mathcal{L}_V(t), \tag{9}$$

and the propagators are norm-decreasing,

$$\|\tau(s,t)O\| \le \|O\| \quad \forall\ \mathcal{L}\in\mathbb{L}_\Lambda, s\le t, O\in\mathcal{B}(\mathcal{H}). \tag{10}$$

Quasilocality of the evolution.—Given an operator $O_Y\in\mathcal{B}(\mathcal{H}_Y)$ with support $Y\subset\Lambda$, we would like to show that the exactly time-evolved operator $\tau(r,t)O_Y$ with $r\le t$ can be approximated by the evolution with respect to a spatially truncated Liouvillian, i.e., by $\tau_{\bar{V}}(r,t)O_Y$ with $Y\subset V\subset\Lambda$. Indeed, our main result, Theorem 2, states that the approximation error is exponentially small, in the distance of $\Lambda\setminus V$ to the time-r slice of a space-time cone originating from the operator's support Y at time t, as depicted in Fig. 1(b). More precisely, the error decays exponentially in $d(Y,\Lambda\setminus V)/a - v(t-r)$, where $d(X,Y) := \inf_{x\in X, y\in Y} d(x,y)$ is the distance of two subsystems X, $Y\subset\Lambda$, and $v = e\mathcal{Z}|\ell|$ is the so-called Lieb-Robinson velocity.

To prove this, we can write the difference of the evolved operators in the form

$$\begin{aligned}\tau(r,t)O_Y - \tau_{\bar{V}}(r,t)O_Y &= -\int_r^t ds\,\partial s[\tau_{\bar{V}}(r,s)\tau(s,t)]O_Y \\ &= \int_r^t ds\,\tau_{\bar{V}}(r,s)\underbrace{[\mathcal{L}(s)-\mathcal{L}_{\bar{V}}(s)]}_{=\mathcal{L}_{\Lambda\setminus V}(s)}\tau(s,t)O_Y\end{aligned}$$

due to the fundamental theorem of calculus and Eqs. (8) and (9). Using the triangle inequality and the fact that the propagators are norm-decreasing, it follows that

$$\|\tau(r,t)O_Y - \tau_{\bar{V}}(r,t)O_Y\| \le \sum_{X\subset\Lambda\setminus V}\int_r^t ds\|\ell_X(s)\tau(s,t)O_Y\|. \tag{11}$$

In the case of unitary dynamics ($\ell_X(s)O = i[h_X, O]$), the integrand would be of the form $\|[h_X, \tau(s,t)O_Y]\|$, and the standard Lieb-Robinson bound [1–5] would be applicable. To proceed in our more general case, however, we use a Lieb-Robinson bound for Markovian quantum dynamics, similar to recent results in Refs. [25,26].

Theorem 1 (Lieb-Robinson bound for Markovian quantum dynamics).—Let the Liouvillian $\mathcal{L}(t) = \sum_{Z\subset\Lambda}\ell_Z(t)$ for the lattice Λ be of finite range a, with a finite maximum number $\mathcal{Z}$ of nearest neighbors, and $|\ell|$ as defined in Eqs. (1)–(4). Also, let $\mathcal{K}_X \in \mathbb{L}_X$, $O_Y \in \mathcal{B}(\mathcal{H}_Y)$, and $r \le t \in \mathbb{R}$. Then

$$\|\mathcal{K}_X\tau(r,t)O_Y\| \le \mathcal{V}_{X,Y}\|\mathcal{K}_X\|\|O_Y\|e^{v(t-r)-d(X,Y)/a}, \tag{12}$$

where $v := \exp(1)\mathcal{Z}|\ell|$ and $\mathcal{V}_{X,Y} := \min\{\frac{Vol(\bar{X})}{\mathcal{Z}}, \frac{Vol(\bar{Y})}{\mathcal{Z}}\}$.

The proof is given in the Supplemental Material [38]. With the Lindblad representation $\mathcal{K}_X O = i[k, O] + \sum_\nu[K_\nu^\dagger O K_\nu - \frac{1}{2}(K_\nu^\dagger K_\nu O + O K_\nu^\dagger K_\nu)]$ of the Liouvillian $\mathcal{K}_X$, one has in Eq. (12) that $\|\mathcal{K}_X\|/2 \le \|k\| + \sum_\nu\|K_\nu\|^2$. The theorem tells us that an evolved observable $\tau(r,t)O_Y$ remains basically unchanged when we evolve it with respect to a Liouvillian that is supported at a distance $R \gg v(t-r)$ away from Y, i.e., that $\tau(r,t)O_Y$ behaves like the identity outside the corresponding space-time cone. In the special case $\mathcal{K}_X O = i[O_X, O]$, Eq. (12) yields a Lieb-Robinson bound for $\|[O_X, \tau(r,t)O_Y]\|$ as in Ref. [26].

This theorem can now be employed to proceed from Eq. (11) in our proof of quasilocality. Let us restrict ourselves to the typical case of Liouvillians $\mathcal{L}(t)$ for which the number of terms $\ell_X(t)$ with distance $d(y,X)/a \in [n, n+1)$ from any site $y \in \Lambda$ is bounded by a power law,

$$\begin{aligned}&|R_{n,y}| \le Mn^\kappa \quad \forall_{y\in\Lambda, n\in\mathbb{N}_+}, \\ &R_{n,y} := \left\{X \subset \Lambda | \ell_X \ne 0, \frac{d(y,X)}{a} \in [n, n+1)\right\},\end{aligned} \tag{13}$$

for some constants $M, \kappa > 0$. Now, choose a point $y_0 \in Y$ that is closest to $\Lambda\setminus V$, i.e., $d(y_0, \Lambda\setminus V) = d(Y, \Lambda\setminus V)$. With $D := \lceil d(Y, \Lambda\setminus V)/a\rceil$, we can exploit that the support of every term in $\mathcal{L}_{\Lambda\setminus V}$ is an element of exactly one of the sets R_{n,y_0} with $n \ge D$ to obtain

$$\begin{aligned}&\|\tau(r,t)O_Y - \tau_{\bar{V}}(r,t)O_Y\| \\ &\le \sum_{n=D}^\infty\sum_{X\in R_{n,y_0}}\int_r^t ds\|\ell_X(s)\tau(s,t)O_Y\| \\ &\le \sum_{n=D}^\infty Mn^\kappa|\ell|\|O_Y\|\int_r^t ds e^{v(t-r)-n} \\ &\le M|\ell|\|O_Y\|\frac{e^{v(t-r)}}{v}\sum_{n=D}^\infty n^\kappa e^{-n}.\end{aligned}$$

In the second step, Theorem 1 and $\mathcal{V}_{XY} \le \mathrm{Vol}(\bar{X})/\mathcal{Z} \le 1$ have been used. With the bound $\sum_{n=D}^\infty n^\kappa e^{-n} \le 2eD^\kappa e^{-D}$ $\forall_{D>2\kappa+1}$, we arrive at the central result of this work.

Theorem 2 (Quasilocality of Markovian quantum dynamics).—Let the Liouvillian $\mathcal{L}(t) = \sum_{Z\subset\Lambda}\ell_Z(t)$ for the lattice Λ be of finite range a, with a finite maximum number $\mathcal{Z}$ of nearest neighbors, and $|\ell|$ as defined in Eqs. (1)–(4). Further, let constraint Eq. (13) be fulfilled for some constants $M, \kappa > 0$. Also, let $Y \subset V \subset \Lambda$, $O_Y \in \mathcal{B}(\mathcal{H}_Y)$, and $r \le t \in \mathbb{R}$. Then one has with $D := \lceil d(Y, \Lambda\setminus V)/a\rceil$

$$\begin{aligned}&\|\tau(r,t)O_Y - \tau_{\bar{V}}(r,t)O_Y\| \\ &\le \frac{2M}{\mathcal{Z}}\|O_Y\|D^\kappa e^{v\cdot(t-r)-D} \quad \forall_{D>2\kappa+1},\end{aligned} \tag{14}$$

where v is the Lieb-Robinson speed from Eq. (12).

The full dynamics can be approximated with exponential accuracy by subsystem dynamics. In a sense, the constraint Eq. (13) requires the lattice to have a finite spatial dimension. A $\mathcal{D}$-dimensional hypercubic lattice with finite-range interactions fulfills Eq. (13) with $\kappa = \mathcal{D} - 1$. An interesting observation is that short-range models on a Bethe lattice [40] have a finite Lieb-Robinson speed according to Theorem 1 but do not fulfill Eq. (13) and are thus not covered by Theorem 2. Hence, for such systems, it is conceivable that a quench of the Liouvillian starting at time $t = 0$ with a distance of at least aD from some point y causes a perceptible effect at y for a time $t^* \ll D/v$.

Trotter decomposition of the evolution.—The quasilocality of the dynamics, Theorem 2, implies that the evolution of observables with a finite spatial support can be simulated efficiently on classical computers, in the sense that the computation cost is independent of the system size, irrespective of the desired accuracy. However, exploiting this, in an exact diagonalization approach that stores the approximated time-evolved observable $\tau_{\bar{V}}(r,t)O_Y$ in a full basis of $\mathcal{H}_{\bar{V}}$ exactly, requires resources

that are exponential in the size $|\bar{V}|$ of the considered subsystem. There are more sophisticated numerical techniques; e.g., one can use matrix-product operators [41–43] for the representation of (an approximation to) $\tau_{\bar{V}}(r,t)O_Y$ or sampling algorithms. In such schemes, it is typically not possible to address the differential equation for $\tau_{\bar{V}}(r,t)O_Y$ directly, but one can use Trotter decompositions [28] instead, where propagators $\tau_{\bar{V}}(r,t)$ are decomposed into a circuit of local (diameter-a) channels.

Using the quasilocality, Theorem 2, and techniques as in Ref. [27], we can derive a Trotter decomposition with an error that is polynomial in time, at most linear in the time step, and, in extension of Ref. [27], system-size independent. Furthermore, implementing such a Trotter circuit on a quantum computer [27] yields a simulation that, additionally to being independent of the system size, is efficient in time. In this case, the physically relevant norm for superoperators T is the subsystem-seminorm

$$\|T\|_Y := \sup_{O_Y \in \mathcal{B}(\mathcal{H}_Y)} \|TO_Y\|/\|O_Y\|. \tag{15}$$

Theorem 3 (Efficient Trotter decomposition of time-evolved observables).—With the preconditions of Theorem 2, a sequence of times $t_0 \leq t_1 \leq \ldots \leq t_N$ and a sequence of subsystems $Y \subset V_1 \subset V_2 \subset \ldots \subset V_N \subset \Lambda$ such that $D_n := \lceil d(Y, \Lambda \setminus V_n)/a \rceil > 2\kappa + 1 \; \forall_n$, the Trotter decomposition

$$\tilde{\tau} := \prod_{n=1}^{N} \prod_{Z \subset \bar{V}_n : \ell_Z \neq 0} \tau_Z(t_{n-1}, t_n) \tag{16}$$

into propagators τ_Z for local Liouville terms ℓ_Z approximates the full system propagator $\tau(t_0, t_N)$ up to an error

$$\|\tau(t_0, t_N) - \tilde{\tau}\|_Y \leq \sum_{n=1}^{N} \left(\frac{2M}{\mathcal{Z}} D_n^{\kappa} e^{v(t_n - t_0) - D_n} + \varepsilon_n \right),$$

$$\varepsilon_n := (t_n - t_{n-1})^2 \mathcal{Z} \, \mathrm{Vol}(\bar{V}_n) |\ell|^2 e^{(t_n - t_{n-1})|\ell|} \tag{17}$$

with the Lieb-Robinson speed v from Eq. (12).

In the Trotter decomposition $\tilde{\tau}$, we used the convention $\prod_{n=1}^{N} T_n = T_1 T_2 \ldots T_N$, and the ordering of the channels τ_Z in the second product of Eq. (16) can be chosen arbitrarily. As in Ref. [27], one can use averaged Liouvillians, i.e., $\tau_Z(r,t) \mapsto e^{\int_r^t ds \ell_Z(s)}$, without changing the scaling of the error bound. Choosing a constant time step, $t_n = n\Delta t$, and subsystems V_n such that $D_n = D_0 + vn\Delta t$, for sufficiently large D_0, the bound (17) is dominated by the Trotter errors ε_n. The subsystems can be chosen such that $\mathrm{diam} V_n \leq \mathrm{diam}(Y) + aD_n$, as shown in Fig. 1(c). For this case, the total error is in $\mathcal{O}(\Delta t(\mathrm{diam}(Y)/a + D_0 + vt)^{\kappa+2})$. Higher-order Trotter-Suzuki decompositions [44] can be used to further improve the scaling in Δt.

To prove Theorem 3, one can first apply Theorem 2, the inequality $\|T_1 T_2 - \tilde{T}_1 \tilde{T}_2\| \leq \|T_1\| \|T_2 - \tilde{T}_2\| + \|T_1 - \tilde{T}_1\| \|\tilde{T}_2\|$, and Eq. (10) iteratively N times, to obtain

$$\|\tau(t_0, t_N) - \tau^V\|_Y \leq \frac{2M}{\mathcal{Z}} \sum_{n=1}^{N} D_n^{\kappa} e^{v(t_n - t_0) - D_n} \tag{18}$$

with $\tau^V := \prod_{n=1}^{N} \tau_{\bar{V}_n}(t_{n-1}, t_n)$. For every time-step propagator $\tau_{\bar{V}_n}(t_{n-1}, t_n)$, we can then employ a Trotter decomposition similar to Ref. [27], yielding

$$\|\tau_{\bar{V}}(r,t) - \prod_{Z \subset \bar{V}, \ell_Z \neq 0} \tau_Z(r,t)\|_Y \leq (t-r)^2 \mathcal{Z} \, \mathrm{Vol}(\bar{V}) |\ell|^2 e^{(t-r)|\ell|}. \tag{19}$$

See the Supplemental Material [38] for details. Combining Eqs. (18) and (19) with the triangle inequality proves Theorem 3.

Conclusion.—We have shown that the evolution of an observable with support Y under a quantum master equation with a short-range Liouvillian can be approximated by the evolution with respect to the truncation of the Liouvillian to a subsystem $V \supset Y$. The error decreases exponentially in the distance of Y from the complement of V. With this tool, we derived an error bound for Trotter decompositions of the propagator. Those results correspond to efficient simulation techniques for open-system dynamics on classical and quantum computers and provide rigorous bounds to finite-size effects.

We gratefully acknowledge inspiring discussions with J. Eisert, C. Gogolin, V. Nesme, and T. J. Osborne. This work has been supported by the EU (Minos, Qessence), the EURYI, the BMBF (QuOReP), and the Studienstiftung des Deutschen Volkes.

[1] E. H. Lieb and D. W. Robinson, Commun. Math. Phys. **28**, 251 (1972).
[2] O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics* (Springer-Verlag, Berlin, 1997), 2nd ed., Vol. 2.
[3] M. B. Hastings, Phys. Rev. B **69**, 104431 (2004).
[4] M. B. Hastings and T. Koma, Commun. Math. Phys. **265**, 781 (2006).
[5] B. Nachtergaele and R. Sims, Commun. Math. Phys. **265**, 119 (2006).
[6] S. Bravyi, M. B. Hastings, and F. Verstraete, Phys. Rev. Lett. **97**, 050401 (2006).
[7] J. Eisert and T. J. Osborne, Phys. Rev. Lett. **97**, 150404 (2006).
[8] T. J. Osborne, Phys. Rev. Lett. **97**, 157202 (2006).
[9] B. Nachtergaele and R. Sims, in *New Trends in Mathematical Physics*, Selected Contributions of the XVth International Congress on Mathematical Physics, edited by V. Sidoravicius (Springer, Heidelberg, 2009), pp. 591–614.

[10] C. K. Burrell and T. J. Osborne, Phys. Rev. Lett. **99**, 167201 (2007).
[11] C. K. Burrell, J. Eisert, and T. J. Osborne, Phys. Rev. A **80**, 052319 (2009).
[12] B. Nachtergaele, H. Raz, B. Schlein, and R. Sims, Commun. Math. Phys. **286**, 1073 (2008).
[13] N. Schuch, S. K. Harrison, T. J. Osborne, and J. Eisert, Phys. Rev. A **84**, 032309 (2011).
[14] E. B. Davis, *Quantum Theory of Open Systems* (Academic, London, 1976).
[15] R. Alicki and K. Lendi, *Quantum Dynamical Semigroups and Applications*, Lect. Notes Phys. Vol. 717 (Springer, Berlin, 2007).
[16] C. J. Myatt, B. E. King, Q. A. Turchette, C. A. Sackett, D. Kielpinski, W. M. Itano, C. Monroe, and D. J. Wineland, Nature (London) **403**, 269 (2000).
[17] L. Viola, E. M. Fortunato, M. A. Pravia, E. Knill, R. Laflamme, and D. G. Cory, Science **293**, 2059 (2001).
[18] S. Deléglise, I. Dotsenko, C. Sayrin, J. Bernu, M. Brune, J.-M. Raimond, and S. Haroche, Nature (London) **455**, 510 (2008).
[19] J. T. Barreiro, P. Schindler, O. Gühne, T. Monz, M. Chwalla, C. F. Roos, M. Hennrich, and R. Blatt, Nature Phys. **6**, 943 (2010).
[20] J. T. Barreiro, M. Müller, P. Schindler, D. Nigg, T. Monz, M. Chwalla, M. Hennrich, C. F. Roos, P. Zoller, and R. Blatt, Nature (London) **470**, 486 (2011).
[21] S. Diehl, A. Micheli, A. Kantian, B. Kraus, H. P. Buchler, and P. Zoller, Nature Phys. **4**, 878 (2008).
[22] F. Verstraete, M. M. Wolf, and J. I. Cirac, Nature Phys. **5**, 633 (2009).
[23] T. Prosen and E. Ilievski, Phys. Rev. Lett. **107**, 060403 (2011).
[24] T. Prosen, Phys. Rev. Lett. **107**, 137201 (2011).
[25] B. Nachtergaele, A. Vershynina, and V. A. Zagrebnov, Contemp. Math. **552**, 161 (2011).
[26] D. Poulin, Phys. Rev. Lett. **104**, 190401 (2010).
[27] M. Kliesch, T. Barthel, C. Gogolin, M. Kastoryano, and J. Eisert, Phys. Rev. Lett. **107**, 120501 (2011).
[28] H. F. Trotter, Proc. Am. Math. Soc. **10**, 545 (1959).
[29] I. Bloch, J. Dalibard, and W. Zwerger, Rev. Mod. Phys. **80**, 885 (2008).
[30] G. Vidal, Phys. Rev. Lett. **93**, 040502 (2004).
[31] S. R. White and A. E. Feiguin, Phys. Rev. Lett. **93**, 076401 (2004).
[32] A. Daley, C. Kollath, U. Schollwöck, and G. Vidal, J. Stat. Mech. (2004) P04005.
[33] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).
[34] G. Lindblad, Commun. Math. Phys. **48**, 119 (1976).
[35] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, J. Math. Phys. (N.Y.) **17**, 821 (1976).
[36] M. Wolf and J. I. Cirac, Commun. Math. Phys. **279**, 147 (2008).
[37] The results of this Letter follow similarly for systems with long-range interactions of sufficiently fast decay. For the sake of readability we refrain from presenting this more general scenario.
[38] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.108.230504 for a brief discussion of the employed superoperator norms and properties of the propagators, a proof of the dissipative Lieb-Robinson bound (Theorem 1), the employed bounds on some elementary series, and the Trotter decomposition for a single time step.
[39] J. Watrous, Quantum Inf. Comput. **5**, 58 (2005).
[40] H. A. Bethe, Proc. R. Soc. A **150**, 552 (1935).
[41] M. Zwolak and G. Vidal, Phys. Rev. Lett. **93**, 207205 (2004).
[42] I. P. McCulloch, J. Stat. Mech. (2007) P10014.
[43] M. J. Hartmann, J. Prior, S. R. Clark, and M. B. Plenio, Phys. Rev. Lett. **102**, 057202 (2009).
[44] M. Suzuki, J. Math. Phys. (N.Y.) **26**, 601 (1985).

Supplementary Material
(Quasi-locality and efficient simulation of Markovian quantum dynamics).

Thomas Barthel and Martin Kliesch
Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany and
Institute for Physics and Astronomy, University of Potsdam, 14476 Potsdam, Germany
(Dated: February 12, 2012)

In the following appendices we give a brief account of the operator and super-operator norms used in the main text, explain some properties of the propagators, prove the generalized Lieb-Robinson bound for Markovian quantum dynamics (Theorem 1), give the derivations for the employed bounds on two elementary series, and present the Trotter decomposition of the propagator for a single time step along the lines of Ref. [1].

A. Operator and super-operator norms

In this work, two of the *Schatten p-norms* [2] are employed. The ∞-*norm* of an operator $O \in \mathcal{B}(\mathcal{H})$ is defined as its largest singular value and is equal to the *operator norm*,

$$\|O\|_\infty = \|O\| \coloneqq \sup_{|\psi\rangle\in\mathcal{H}} \frac{\||O|\psi\rangle\|}{\||\psi\rangle\|}, \tag{1}$$

where $\||\psi\rangle\| = \sqrt{\langle\psi|\psi\rangle}$ denotes the *vector* 2-*norm*. The ∞-*norm* is the physically relevant norm for observables. The 1-*norm*, of an operator $O \in \mathcal{B}(\mathcal{H})$ is defined as the sum of its singular values and is equal to the *trace norm*,

$$\|O\|_1 = \|O\|_{\mathrm{tr}} \coloneqq \mathrm{Tr}\sqrt{O^\dagger O}. \tag{2}$$

It is the physically relevant norm for states, i.e., density matrices [3]. Those operator norms induce corresponding norms for super-operators $T \in \mathcal{B}(\mathcal{B}(\mathcal{H}))$. The $(\infty\to\infty)$-*norm* is defined as

$$\|T\| \coloneqq \|T\|_{\infty\to\infty} \coloneqq \sup_{O\in\mathcal{B}(\mathcal{H})} \frac{\|TO\|_\infty}{\|O\|_\infty} \tag{3}$$

and the $(1\to1)$-*norm* is

$$\|T\|_{1\to1} \coloneqq \sup_{O\in\mathcal{B}(\mathcal{H})} \frac{\|TO\|_1}{\|O\|_1}. \tag{4}$$

In order to switch between the Schrödinger and the Heisenberg picture, one needs to consider the adjoint $T^\dagger$ of a super-operator T, defined by

$$\langle A, TB\rangle_{\mathrm{HS}} = \langle T^\dagger A, B\rangle_{HS} \quad \forall_{A,B\in\mathcal{B}(\mathcal{H})}, \tag{5}$$

where $\langle\cdot,\cdot\rangle_{\mathrm{HS}}$ denotes the Hilbert-Schmidt inner product $\langle A, B\rangle_{\mathrm{HS}} \coloneqq \mathrm{Tr}(A^\dagger B)$. The $(1\to1)$-norm is *dual* to the $(\infty\to\infty)$-norm in the sense that

$$\begin{aligned}\|T\|_{\infty\to\infty} &= \sup_{\|O\|_\infty=1} \|TO\|_\infty \\ &= \sup_{\|O\|_\infty=1} \sup_{\|X\|_1=1} |\langle X, TO\rangle_{\mathrm{HS}}| \\ &= \sup_{\|X\|_1=1} \sup_{\|O\|_\infty=1} |\langle T^\dagger X, O\rangle_{\mathrm{HS}}| \\ &= \sup_{\|X\|_1=1} \left\|T^\dagger X\right\|_1 = \left\|T^\dagger\right\|_{1\to1}. \end{aligned} \tag{6}$$

This allows us to relate the appropriate norm of a propagator T in the Heisenberg picture to the norm of the corresponding propagator $T^\dagger$ in the Schrödinger picture. For more on properties of the norms, see, for example, Refs. [2, 4].

B. Properties of the propagators

The derivative of a propagator with respect to its second time argument is given by

$$\partial_t \tau_V(s,t) = \tau_V(s,t)\mathcal{L}_V(t). \tag{7}$$

Using the defining properties $\partial_s\tau_V(s,t) = -\mathcal{L}_V(s)\tau_V(s,t)$ and $\tau_V(t,t) = \mathrm{id}$, Eq. (7) follows from the equation

$$\begin{aligned} 0 = \partial_t\,\mathrm{id} &= \partial_t[\tau_V(t,s)\tau_V(s,t)] \\ &= [\partial_t\tau_V(t,s)]\tau_V(s,t) + \tau_V(t,s)[\partial_t\tau_V(s,t)] \end{aligned}$$

after applying $\tau_V(s,t)$ to it.

Let us explain why the propagators are norm-decreasing, i.e.,

$$\|\tau(s,t)O\| \le \|O\| \quad \forall\,\mathcal{L}\in\mathbb{L}_\Lambda,\, s\le t,\, O\in\mathcal{B}(\mathcal{H}). \tag{8}$$

The adjoint propagator $\tau^\dagger(s,t)$ (see Sect. A), describes the time-evolution in the Schrödinger picture, $\rho(t) = \tau^\dagger(s,t)\rho(s)$, where $\rho(t)$ denotes the system state at time t. First of all, we note that $\tau^\dagger(s,t)$ is a completely positive, trace-preserving (CPT) map since it can be written as a *product integral* [5],

$$\tau^\dagger(s,t) = \lim_{\Delta t_j\to0} \prod_j e^{\mathcal{L}^\dagger(t_j)\Delta t_j}.$$

Every factor $e^{\mathcal{L}^\dagger(t_j)\Delta t_j}$ is an exponential of a constant Liouvillian and is hence CPT. Thus, the finite products are CPT maps and, since the set of CPT maps is closed, also the limit $\tau^\dagger(s,t)$ is a CPT map. Then Eq. (8) follows from the norm duality $\|T\| \equiv \|T\|_{\infty\to\infty} = \|T^\dagger\|_{1\to1}$ [Eq. (6)] and $\|T^\dagger\|_{1\to1} = 1$ for all CPT maps $T^\dagger$. The latter has, for example, been shown in Ref. [1].

C. Proof of Theorem 1

With an argument similar to those in Refs. [6–12], we want to bound the norm of the operator

$$G(r) \coloneqq \mathcal{K}_X \tau(r,t) O_Y \tag{9}$$

under the preconditions of Theorem 1. G is the solution to the final value problem $G(t) = \mathcal{K}_X O_Y$,

$$\begin{aligned}\partial_r G(r) &= -\mathcal{K}_X \mathcal{L}(r) \tau(r,t) O_Y \\ &= -\mathcal{L}_{\Lambda\setminus X}(r) G(r) - \mathcal{K}_X \mathcal{L}_{\bar{X}}(r) \tau(r,t) O_Y,\end{aligned}$$

due to $\partial_s \tau(s,t) = -\mathcal{L}(s)\tau(s,t)$, $\mathcal{L} = \mathcal{L}_{\bar{X}} + \mathcal{L}_{\Lambda\setminus X}$, and $\mathcal{K}_X \mathcal{L}_{\Lambda\setminus X} = \mathcal{L}_{\Lambda\setminus X}\mathcal{K}_X$ for all Liouvillians $\mathcal{K}_X \in \mathbb{L}_X$. As can be checked by differentiation, a corresponding integral equation for $G(r)$ is

$$\begin{aligned}G(r) &= \tau_{\Lambda\setminus X}(r,t) G(t) \\ &\quad + \int_r^t \mathrm{d}s\, \tau_{\Lambda\setminus X}(r,s) \mathcal{K}_X \mathcal{L}_{\bar{X}}(s) \tau(s,t) O_Y.\end{aligned}$$

Using the triangle inequality, the norm-submultiplicativity, and the fact that the propagators are norm-decreasing, this yields the bound

$$\begin{aligned}\|G(r)\| &\le \|G(t)\| + \|\mathcal{K}_X\| \int_r^t \mathrm{d}s\, \|\mathcal{L}_{\bar{X}}(s)\tau(s,t)O_Y\| \\ &\le \|G(t)\| + \|\mathcal{K}_X\| \sum_{Z\subset\bar{X}} \int_r^t \mathrm{d}s\, \|\ell_Z(s)\tau(s,t)O_Y\|.\end{aligned} \tag{10}$$

Now a Picard iteration for the related quantity

$$C_X(r) \coloneqq \sup_{\mathcal{K}\in\mathbb{L}_X} \frac{\|\mathcal{K}\tau(r,t)O_Y\|}{\|\mathcal{K}\|} \tag{11}$$

can be used to obtain a bound for $\|G(r)\|$. Inserting Eq. (10) in Eq. (11) gives

$$\begin{aligned}C_X(r) &\le C_X(t) + \sum_{Z\subset\bar{X}} \sup_{s\in[r,t]} \|\ell_Z(s)\| \int_r^t \mathrm{d}s\, C_Z(s), \\ C_X(t) &\le \delta(X,Y)\,\|O_Y\|,\end{aligned} \tag{12}$$

where $\delta(X,Y) = 1$ for $X\cap Y \neq \varnothing$ and $\delta(X,Y) = 0$, otherwise. The second line follows from $\mathcal{K}_X O_Y = 0$ for Liouvillians $\mathcal{K}_X \in \mathbb{L}_X$ with $X\cap Y = \varnothing$, and $\|\mathcal{K}O_Y\| \le \|\mathcal{K}\|\,\|O_Y\|$ in general. Starting the Picard iteration for $C_X(r)$ with Eq. (12) and $Z_0 \coloneqq X$ leads to

$$C_X(r) \le \|O_Y\| \sum_{n=0}^{\infty} \frac{(t-r)^n}{n!} c_n \text{ with} \tag{13}$$

$$c_n = \sum_{Z_1\subset\bar{Z}_0} \sum_{Z_2\subset\bar{Z}_1} \cdots \sum_{Z_n\subset\bar{Z}_{n-1}} \delta(Z_n,Y) \prod_{i=1}^{n} \sup_{s\in[r,t]} \|\ell_{Z_i}(s)\|.$$

Now we can exploit that the Liouville terms are of finite range a and that they induce a finite maximum number

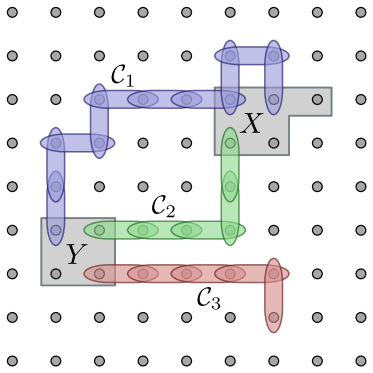

Figure 1: In the proof of Theorem 1, we need to bound a sum over all paths of length n starting in X, the support of $\mathcal{K}_X$, and ending in Y, the support of O_Y; see Eq. (13). A path corresponds to a sequence of local Liouville terms (ℓ_{Z_i}) with overlapping supports. In the depicted situation of a two-dimensional lattice with nearest-neighbor interaction, path $\mathcal{C}_1$ would contribute to the sum for $n = 10$ and $\mathcal{C}_2$ for $n = 5$. To simplify the calculation for upper bounds, the sums are extended to contain *all* paths starting in Y (if $\mathrm{Vol}(\bar{Y}) < \mathrm{Vol}(\bar{X})$; all paths starting from X, otherwise). Hence, in the bound for $n = 5$, also paths like $\mathcal{C}_3$ are taken into account.

$\mathcal{Z} = \max_{Z:\ell_Z\neq 0} \mathrm{Vol}(\bar{Z})$ of nearest neighbors. The sum in Eq. (13) runs over all paths from X to Y. Depending on whether $\mathrm{Vol}(\bar{X})$ or $\mathrm{Vol}(\bar{Y})$ is larger, the number of such paths with length n can be bounded by the number of all length-n paths starting in X or Y, respectively. See Fig. 1. This gives the simple bound

$$\begin{aligned}c_n &\le \mathcal{V}_{X,Y} (\mathcal{Z}|\ell|)^n, \text{ and thus,} \\ C_X(t) &\le \mathcal{V}_{X,Y}\,\|O_Y\| \sum_{n=D}^{\infty} \frac{\theta^n}{n!},\end{aligned} \tag{14}$$

where $\theta \coloneqq (t-r)\mathcal{Z}|\ell|$, $D = \lceil d(X,Y)/a \rceil$, and $\mathcal{Z}\mathcal{V}_{X,Y}$ is the minimum of the numbers of Liouville terms ℓ_{Z_i} supported in X and Y, $\mathcal{V}_{X,Y} = \min\{\mathrm{Vol}(\bar{X}), \mathrm{Vol}(\bar{Y})\}/\mathcal{Z}$. We have also used that $c_n = 0$ for all $n < D$, as one needs at least D Liouville terms of overlapping support to pass from the subsystem X to subsystem Y, such that $\delta(Z_n, Y) \neq 0$. Using induction, the sum in Eq. (14) can be bounded by $\sum_{n=D}^{\infty} \frac{\theta^n}{n!} \le e^{\theta e - D}$; see Sect. D. Hence, Theorem 1 follows,

$$\|G(t)\| \le \mathcal{V}_{X,Y}\,\|\mathcal{K}_X\|\,\|O_Y\|\, e^{(t-r)\mathcal{Z}|\ell|e-D}.$$

D. Bound on the partial exponential sum

In the following, we prove that

$$\underbrace{\sum_{n=N}^{\infty} \tfrac{x^n}{n!}}_{=:f_N(x)} \le \underbrace{e^{xe-N}}_{=:g_N(x)} \quad \forall x \ge 0,\ N \in \mathbb{N}_0. \tag{15}$$

Note first that, for $N = 0$,

$$f_0(x) = e^x \le e^{xe} = g_0(x) \quad \forall_{x\ge 0}.$$

The functions f_N and g_N obey the differential equations

$$\partial_x f_{N+1}(x) = f_N(x), \quad \partial_x g_{N+1}(x) = g_N(x) \quad \forall_{x,N}.$$

For all $N > 0$, the initial values $f_N(0) = 0$ and $g_N(0) = e^{-N}$ obviously obey $f_N(0) \leq g_N(0)\ \forall_{N>0}$. Consequently, $f_N(x) \leq g_N(x)\ \forall_{x\geq 0}$ implies $f_{N+1}(x) \leq g_{N+1}(x)\ \forall_{x\geq 0}$. This proves Eq. (15) inductively.

E. Bound on a sum of exponentials

In the following, it is shown that

$$\sum_{n=D}^{\infty} n^{\kappa} e^{-n} \leq 2eD^{\kappa}e^{-D} \quad \forall \kappa > 0,\ D > 2\kappa + 1 \in \mathbb{N}. \tag{16}$$

Due to the definition $\Gamma(a,x) \coloneqq \int_x^{\infty} \mathrm{d}t\, t^{a-1}e^{-t}$ of the incomplete Gamma function, one has

$$\sum_{n=D}^{\infty} n^{\kappa} e^{-n} \leq \Gamma(\kappa + 1, D - 1). \tag{17}$$

The bound

$$\Gamma(a,x) \leq Bx^{a-1}e^{-x} \quad \forall a > 1,\ B > 1,\ x > \tfrac{B(a-1)}{B-1}$$

of Natalini and Palumbo [13], reads for the choice $B = 2$

$$\Gamma(a,x) \leq 2x^{a-1}e^{-x} \quad \forall a > 1,\ x > 2(a-1).$$

Together with Eq. (17) one obtains

$$\sum_{n=D}^{\infty} n^{\kappa} e^{-n} \leq 2(D-1)^{\kappa}e^{-D+1} \quad \forall_{\kappa>0,\, D>2\kappa+1}$$

and hence the acclaimed Eq. (16).

F. Trotter expansion of a propagator

For two times $q \leq t$, we derive the Trotter error bound

$$\Big\|\tau_V(q,t) - \prod_{Z\subset V} \tau_Z(q,t)\Big\|_Y \leq (t-q)^2 \mathcal{Z}\mathrm{Vol}(V)|\ell|^2 e^{(t-q)|\ell|} \tag{18}$$

that is employed in the proof of Theorem 3. To this purpose, let us determine an upper bound for the right-hand side of

$$\big\|T^{q,t}_{\mathcal{L}+\ell} - T^{q,t}_{\mathcal{L}} T^{q,t}_{\ell}\big\|_Y \leq \big\|T^{q,t}_{\mathcal{L}+\ell} - T^{q,t}_{\mathcal{L}} T^{q,t}_{\ell}\big\|,$$

where $T^{q,t}_{\mathcal{K}}$ denotes the propagator for a Liouvillian $\mathcal{K}(t) \in \mathbb{L}$, $\mathcal{L}(t) \in \mathbb{L}$ obeys the preconditions of Theorem 3, and $\ell(t) \in \mathbb{L}_Z$ is a local Liouvillian term with support Z. We denote the inverse of a propagator $T^{q,t}_{\mathcal{K}}$ by $T^{t,q}_{\mathcal{K}}$. Using $\partial_q T^{q,t}_{\mathcal{K}} = -\mathcal{K}(q)T^{q,t}_{\mathcal{K}}$, $\partial_t T^{q,t}_{\mathcal{K}} = T^{q,t}_{\mathcal{K}}\mathcal{K}(t)$, $T^{r,s}_{\mathcal{K}} T^{s,t}_{\mathcal{K}} = T^{r,t}_{\mathcal{K}}$, $T^{t,t}_{\mathcal{K}} = \mathrm{id}$, and applying the fundamental theorem of calculus twice, one finds

$$\begin{aligned}
T^{q,t}_{\mathcal{L}+\ell} - T^{q,t}_{\mathcal{L}} T^{q,t}_{\ell} &= (T^{q,t}_{\mathcal{L}+\ell} T^{t,q}_{\ell} T^{t,q}_{\mathcal{L}} - \mathrm{id}) T^{q,t}_{\mathcal{L}} T^{q,t}_{\ell} \\
&= \int_q^t \mathrm{d}s\, \partial_s \big(T^{q,s}_{\mathcal{L}+\ell} T^{s,q}_{\ell} T^{s,q}_{\mathcal{L}}\big) T^{q,t}_{\mathcal{L}} T^{q,t}_{\ell} \\
&= \int_q^t \mathrm{d}s\, T^{q,s}_{\mathcal{L}+\ell} \big(\mathcal{L}(s) - T^{s,q}_{\ell}\mathcal{L}(s) T^{q,s}_{\ell}\big) T^{s,q}_{\ell} T^{s,t}_{\mathcal{L}} T^{q,t}_{\ell} \\
&= \int_q^t \mathrm{d}s \int_q^s \mathrm{d}r\, T^{q,s}_{\mathcal{L}+\ell} \partial_r \big(T^{s,r}_{\ell}\mathcal{L}(s) T^{r,s}_{\ell}\big) T^{s,q}_{\ell} T^{s,t}_{\mathcal{L}} T^{q,t}_{\ell} \\
&= \int_q^t \mathrm{d}s \int_q^s \mathrm{d}r\, T^{q,s}_{\mathcal{L}+\ell} T^{s,r}_{\ell} [\ell(r), \mathcal{L}(s)] T^{r,q}_{\ell} T^{s,t}_{\mathcal{L}} T^{q,t}_{\ell}.
\end{aligned}$$

The time arguments occurring in the integrand are ordered according to $q \leq r \leq s \leq t$. The norm of the propagators is $\big\|T^{s,t}_{\mathcal{K}}\big\| = 1\ \forall_{s\leq t}$. A bound for the norm of the inverse propagators can be obtained from their representations as time-ordered exponentials [1, 5], yielding $\big\|T^{t,s}_{\mathcal{K}}\big\| \leq \exp(\int_s^t \mathrm{d}r\, \|\mathcal{K}(t)\|)\ \forall_{s\leq t}$. With those properties, the triangle inequality, and the norm submultiplicativity,

$$\begin{aligned}
\big\|T^{q,t}_{\mathcal{L}+\ell} - T^{q,t}_{\mathcal{L}} T^{q,t}_{\ell}\big\| &\leq \int_q^t \mathrm{d}s \int_q^s \mathrm{d}r\, \|[\ell(r), \mathcal{L}(s)]\|\, e^{(s-q)|\ell|} \\
&\leq (t-q)^2 \mathcal{Z}|\ell|^2 e^{(t-q)|\ell|}.
\end{aligned}$$

This bound and the inequality $\|T_1T_2 - \tilde{T}_1\tilde{T}_2\| \leq \|T_1\|\,\|T_2 - \tilde{T}_2\| + \|T_1 - \tilde{T}_1\|\|\tilde{T}_2\|$ can now be used iteratively, separating one local propagator $T^{q,t}_{\ell_Z}$ after another. As $\mathcal{L}_V$ is a sum of $\mathrm{Vol}(V)$ terms ℓ_Z, Eq. (18) follows.

[1] M. Kliesch, T. Barthel, C. Gogolin, M. Kastoryano, and J. Eisert, Phys. Rev. Lett. **107**, 120501 (2011).

[2] R. Bhatia, *Matrix Analysis* (Springer-Verlag, New York, 1997).

[3] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[4] J. Watrous, Quantum Inf. Comput. **5**, 58 (2005).

[5] J. D. Dollard and C. N. Friedman, *Product integration with applications to differential equations* (Addison-Wesley, Reading, Mass., 1979).

[6] E. H. Lieb and D. W. Robinson, Commun. Math. Phys. **28**, 251 (1972).

[7] O. Bratteli and D. W. Robinson, *Operator algebras and quantum statistical mechanics, Volume 2*, 2nd ed. (Springer-Verlag, Berlin, 1997).

[8] M. B. Hastings, Phys. Rev. B **69**, 104431 (2004).
[9] M. B. Hastings and T. Koma, Commun. Math. Phys. **265**, 781 (2006).
[10] B. Nachtergaele and R. Sims, Commun. Math. Phys. **265**, 119 (2006).
[11] D. Poulin, Phys. Rev. Lett. **104**, 190401 (2010).
[12] B. Nachtergaele, A. Vershynina, and V. A. Zagrebnov, Contemp. Math. **552**, 161 (2011).
[13] P. Natalini and B. Palumbo, Math. Inequal. Appl. **3**, 6977 (2000).

Boson-Sampling in the light of sample complexity

C. Gogolin, M. Kliesch, L. Aolita, and J. Eisert

Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany

September 17, 2013

Boson-Sampling is a classically computationally hard problem that can — in principle — be efficiently solved with quantum linear optical networks. Very recently, a rush of experimental activity has ignited with the aim of developing such devices as feasible instances of quantum simulators. Even approximate Boson-Sampling is believed to be hard with high probability if the unitary describing the optical network is drawn from the Haar measure. In this work we show that in this setup, with probability exponentially close to one in the number of bosons, no symmetric algorithm can distinguish the Boson-Sampling distribution from the uniform one from fewer than exponentially many samples. This means that the two distributions are operationally indistinguishable without detailed a priori knowledge. We carefully discuss the prospects of efficiently using knowledge about the implemented unitary for devising non-symmetric algorithms that could potentially improve upon this. We conclude that due to the very fact that Boson-Sampling is believed to be hard, efficient classical certification of Boson-Sampling devices seems to be out of reach.

1 Introduction

Quantum information theory suggests that it should be possible to design physical devices performing information processing tasks that cannot be classically efficiently simulated. The most spectacular example of this type known to date is a fully-fletched Shor-class quantum computer, able to factorize numbers efficiently, hence solving a practically relevant problem for which no classical efficient algorithm is known [1]. Needless to say, the actual physical realisation of such a device is extraordinarily difficult for a number of reasons, the difficulty of protecting quantum systems from the unwanted effects of decoherence being only one of them. In the light of this observation, it has become a very important milestone to identify devices that can solve some problem that seems impossible to be realised classically, or — in the wording of a blog entry [2] — to achieve "quantum supremacy". This is a challenge equally interesting for experimentalists as well as for theorists: On one hand, it surely is still very difficult to achieve the necessary degree of control, on the other hand, it is a challenge for complexity theorists and theoretical computer scientists to show that a task at hand is computationally hard.

A seminal theoretical step in this direction has recently been achieved with the introduction of the *Boson-Sampling problem* [3]. In this problem, the task is the following: Given as input the unitary U, the number of modes m, and the number of photons n, together describing a quantum linear optical device (see Fig. 1), sample from the output distribution of this device. Ref. [3] establishes strong reasons to believe that classically sampling from this distribution up to a small error in 1-norm is computationally hard with high probability if the unitary U is chosen from the Haar measure and m is scaled suitably with n. The hardness proof

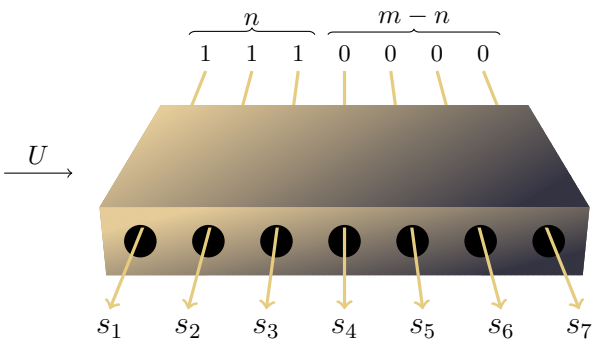

Figure 1: The Boson-Sampling quantum device receives as input a unitary U, applies it to m bosonic modes initialized with exactly one boson in each of the first n modes and the vacuum in the remaining $m-n$ modes and outputs the results $(s_1,\ldots,s_m)$ of local boson number measurements. The Boson-Sampling problem is to sample from the output distribution of this device where U is fixed and chosen in the beginning from the Haar measure on $U(m)$.

rests on the fact that approximating the probabilities of individual outcomes of such a device basically amounts to approximating the permanent of a submatrix of U [10], which, by a plausible complexity theoretic conjecture, is believed to be $\#P$ hard. The main result of Ref. [3] suggests that a 1-norm approximate efficient classical simulation of Boson-Sampling would imply a collapse of the polynomial hierarchy to the third level (compare also Ref. [4]). This has triggered a rush of exciting experimental activity [5–8], aiming at realizing instances of Boson-Sampling, accompanied by theoretical discussions about what errors one should expect in such quantum linear optical experiments [9].

In view of all this, a crucial question that arises is how to certify that a given experiment does actually solve the desired sampling problem, and how many repetitions of the experiment, i.e., samples from its output distribution, are needed for the certification. In contrast to a machine that is efficiently factoring large numbers and hence solves a problem in NP, i.e., produces an output that can be checked efficiently on a classical computer, no efficient certification scheme for Boson-Sampling is known and it is not clear whether such a scheme can exist.

As a first step towards a better understanding of the difficulty of certifying sampling devices we consider the task of deciding whether a device outputs samples from some given interesting probability distribution, for example the Boson-Sampling distribution, or the uniform one. We approach this decision problem in two complementary settings that differ in the amount of information the certifier is allowed to use besides the samples output by the device.

State discrimination (known probability distributions): The certifier works under the assumption that the sampling device outputs independent identically distributed samples and that the output distribution is either of two completely known distributions (for example one of them being the uniform one), i.e., he has access to all the probabilities.

Black box setting (unknown probability distribution): The certifier works under the assumption that the sampling device outputs independent identically distributed samples, but has no a priori knowledge about the output distribution apart from the sample space.

We show that known bounds on the sample complexity imply that in the first setting a number of samples scaling polynomially with the number of bosons is sufficient to guarantee distinguishability of the Boson-Sampling distribution from the uniform one. Notably, this

does not imply that polynomially many samples are sufficient to certify that a given device samples from the correct distribution, even if unlimited computational power and full a priori knowledge about m, n, and U implemented by the supposed device are assumed. The reason for this is simply that being able to decide which of two given distributions a device samples from under the promise that it does indeed sample from either of the two, does not necessarily imply that one is also able to exclude that the device samples from any distribution outside of a small region around the target distribution from the same number of samples.

The complexity theoretic conjecture under which Boson-Sampling is a hard sampling problem, namely that it is expected to be $\#P$ hard to approximate the permanent, implies that approximating the probabilities of the individual outputs of a Boson-Sampling device is also computationally hard. A classical certifier with limited computational power will hence have only very limited knowledge about the ideal output distribution of a supposed Boson-Sampling device. The state discrimination scenario is thus far from realistically capturing the challenge of classically efficiently certifying a real Boson-Sampling device in the laboratory.

The realistic situation much more closely resembles the black box setting (we will discuss this in more detail in Section 3). In this setting the certifier has no a priori knowledge about the output distribution. It is hence reasonable to demand that its decision should only depend on how frequent the outcomes appear. That is to say, knowing nothing about the probability distribution, the labels of the collected samples don't mean anything to the certifier, hence they should not influence his decision.

We formalize this in the notion of *symmetric probabilistic decision algorithms* and show that such algorithms can give meaningful outputs only if they receive sufficiently many samples. This is true not only for the task of distinguishing some distribution from the uniform one, but also in more general settings involving multiple sampling devices. The number of samples necessarily depends on how ***flat*** the distribution(s) are. We call a probability distribution over a finite samples space ϵ-***flat*** if the probability of the most likely outcome is upper bounded by ϵ or, equivalently, if its min entropy is larger than $\log_2 1/\epsilon$. More precisely, we show that the output distribution of any symmetric probabilistic algorithm receiving at most $\mathrm{O}((1/\epsilon)^{1/4})$ samples from each of N sampling devices is with probability $1 - \mathrm{O}(N^2\sqrt{\epsilon})$ independent of the distributions of the sampling devices if the distributions are ϵ-flat (Theorem 8).

We then show that the Boson-Sampling distribution is, for the interesting parameter regions and if U is chosen from the Haar measure, with overwhelmingly high probability exponentially flat (Theorem 12). Together, our findings imply that in the black box setting distinguishing the Boson-Sampling distribution from the uniform one requires exponentially many samples.

We emphasize that our analysis applies to the ideal situation without any experimental imperfections. It is important to mention that our results concerning the flatness, just like the hardness proof of Ref. [3], is probabilistic, in the sense of holding with an extremely high probability if the unitary describing the optical circuit is randomly chosen from the Haar measure and if n and m are sufficiently large. To end up with, we identify a class of imperfect linear optical experimental situations for which one can classically efficiently sample from the output distribution even up to a constant small error in 1-norm.

The rest of this work is organized as follows. First, in Section 2 we introduce the notation and recapitulate the setting considered in the Boson-Sampling problem. Next, in Section 3 we connect the problem of certifying a sampling devices with the decision problem of distinguishing its output distribution from the uniform distribution, explain the state discrimination and the black box setting in more detail, give upper and lower bounds on the sample complexity of this task and discuss what they mean for the original question of certifying Boson-Sampling. Section 4 and 5 contain our technical results concerning the sample complexity in the state

discrimination and black box setting respectively. In Section 6 we show that, with very high probability, the Boson-Sampling distribution is exponentially flat (Theorem 12). Finally, in Section 7 we identify a class of sampling experiments that, due to experimental imperfections, are classically efficiently simulatable in 1-norm.

2 Setting and notation

We use the Landau symbols O and Ω for asymptotic upper and lower bounds. Moreover, we employ the short hand notation $[j] \coloneqq \{1, \ldots, j\}$ for $j \in \mathbb{Z}^+$. The 1-norm and ∞-norm on (probability) vectors are denoted by $\|\cdot\|_1$ and $\|\cdot\|_\infty$.

We consider the output probability distribution $\mathcal{D}_U$ of the *Boson-Sampling device* with $n \geq 1$ bosons in $m \in \mathrm{poly}(n)$ modes, given by [3, 10]

$$\Pr_{\mathcal{D}_U}[S] \coloneqq |\langle 1_n| \varphi(U) |S\rangle|^2 = \frac{|\mathrm{Perm}(U_S)|^2}{\prod_{j=1}^m (s_j!)}, \tag{1}$$

with $S \in \Phi_{m,n}$ being the output sequence of the Boson-Sampling device where

$$\Phi_{m,n} \coloneqq \left\{(s_1, \ldots, s_m) : \sum_{j=1}^m s_j = n\right\} \tag{2}$$

is the sample space. The state vector $|S\rangle$ is the Fock space vector corresponding to S, $|1_n\rangle$ the initial state vector of the Boson-Sampling device with $1_n \coloneqq (1, \ldots, 1, 0, \ldots, 0)$, $\varphi(U)$ the Fock space (metaplectic) representation of the implemented unitary. The unitary matrix $U \in U(m)$ is the corresponding unitary in mode space, transforming vectors of bosonic operators. In turn, $U_S \in \mathbb{C}^{n\times n}$ is the matrix constructed from U by discarding all but the first n columns of U and then, for all $j \in [m]$, taking s_j copies of the j^{th} row of that matrix (deviating from the notation of Ref. [3]). The permanent Perm is defined similarly to how the determinant can be defined via the Leibniz formula, but without the alternating sign.

We refer to $\mathcal{D}_U$ as the *Boson-Sampling distribution*. The *Boson-Sampling problem* is: given as input to the algorithm n, m, and U, sample exactly or approximately from $\mathcal{D}_U$. We will also consider the *post-selected Boson-Sampling distribution* $\mathcal{D}_U^*$ which is obtained from $\mathcal{D}_U$ by discarding all output sequences S with more than one boson per mode, i.e., all S which are not in the set of *collision-free* sequences

$$\Phi_{m,n}^* \coloneqq \left\{S \in \Phi_{m,n} : \forall s \in S : s \in \{0, 1\}\right\}. \tag{3}$$

For the relevant scalings of m with n the post-selection can be done efficiently in the sense that on average at least a constant fraction of the outcome sequences is collision-free (Theorem 13.4 in Ref. [3]).

The main result of Ref. [3] is that under reasonable complexity theoretic conjectures, 1-norm approximate Boson-Sampling, i.e., sampling from a distribution that is close to the Boson-Sampling distribution $\mathcal{D}_U$ in 1-norm, is computationally hard, with high probability if the unitary U is chosen from the Haar measure μ_H, which we denote by $U \sim \mu_H$, and m increases sufficiently fast with n. In fact, the hardness result of approximate Boson-Sampling requires that $m \in \Omega(n^5)$, but it is conjectured that m growing faster than n^2 is sufficient. Importantly, the proof of this result considers only collision-free output sequences, so in fact approximately sampling from $\mathcal{D}_U^*$ is already hard and the hardness argument for Boson-Sampling only uses the structure of the distribution $\mathcal{D}_U$ on $\Phi_{m,n}^*$.

We will repeatedly use that the size

$$|\Phi_{m,n}| = \binom{m+n-1}{n} \tag{4}$$

of the sample space of Boson-Sampling, which grows faster than exponentially with n, fulfills the following bound: Let $m \leq c\, n^{\nu}$ for some $\nu \geq 1$ and $c \geq 0$, then

$$|\Phi_{m,n}| \leq \frac{(m+n-1)^n}{n!} \leq \left(\frac{(m+n-1)\,\mathrm{e}}{n}\right)^n \tag{5}$$

$$\leq \mathrm{e}^n\,(c\, n^{\nu-1} + 1 - 1/n)^n \leq (2\,(c+1)\,\mathrm{e})^n\, n^{(\nu-1)\,n}. \tag{6}$$

3 Boson-Sampling in the context of sample complexity

Our aim is to understand if and how an experimental implementation of Boson-Sampling can be certified to sample from the correct distribution. In particular, we address the question of whether a certification can be achieved from the samples output by the device only. This is natural because Boson-Sampling is an abstract sampling problem with a classical input (the number of bosons n, the number of modes m the unitary matrix U) and a classical output (samples from $\Phi_{m,n}$). The sampling problem as such is independent of the particular physical implementation. We call all information about a claimed Boson-Sampling device that is in principle available to the certifier in addition to the samples *a priori knowledge*.

The problem of certifying a device can be formalized as a decision problem. Whether a decision can be reached can be expressed as a statement about the existence or non existence of an algorithm which, possibly using parts or all of the a priori knowledge, and given a set of samples, accepts (device certified) in the "good" situation with probability at least $2/3$ and in the bad instances rejects (device not certified) with probability at least $2/3$. The probabilities to erroneously reject in a "good" instance (accept in a "bad" instance) are called errors of the first (second) kind. If such an algorithm exists we say it decides the problem, if no such algorithm exists we say that the problem can not be decided. If a sampling problem has a natural problem size, like the number of bosons n in the case of Boson-Sampling, it is natural to consider the scaling of the number of samples needed such that a deciding algorithm exists as a function of this problem size. The order of the number of samples needed by an algorithm is called its *sample complexity*. The sample complexity of a decision problem in turn is the minimal sample complexity of any algorithm that decides the problem. The choice of the value $2/3$ for the accept/reject probabilities is purely conventional. Any other constant finite bias in the accept/reject probabilities can be amplified to values arbitrarily close to one without changing the sample complexity.

The main hardness result of Ref. [3] covers all distributions that are 1-norm close to the ideal Boson-Sampling distribution. Hence, an algorithm that certifies a Boson-Sampling device must necessarily reject with probability at least $2/3$ whenever the device samples from a distribution further away than some small constant distance in 1-norm and it is desirable that it accepts with probability at least $2/3$ if the device samples from the ideal Boson-Sampling distribution. Such an algorithm must hence at least be able to decide whether a given device samples from the ideal Boson-Sampling distribution or the uniform distribution over the same sample space.

In the state discrimination setting, the sample complexity of this task is of order $\mathrm{O}(n^3)$ (Theorem 3), but it is certainly not realistic to assume that the certifier asked to decide this question has full knowledge of the ideal Boson-Sampling distribution. After all, it is the

very point of Boson-Sampling that approximating the probabilities of individual outcomes is a computationally hard problem. It is therefore important to investigate the sample complexity of this task under more reasonable restrictions on the a priori knowledge and computational power of the certifier.

Colloquially speaking, our results on this problem, which are formally stated below, give rise to a rather ironic situation: Instead of building a device that implements Boson-Sampling, for example by means of a quantum optical experiment, one could instead simply program a classical computer to efficiently sample from the uniform distribution over $\Phi^*_{m,n}$ and claim that the device samples from the post-selected Boson-Sampling distribution $\mathcal{D}^*_U$ for some U. If one chooses U from the Haar measure the chances of being caught cheating becomes significantly large only after one was asked for exponentially many samples. This implies that the findings of any experimental realisation of Boson-Sampling have to be interpreted with great care, as far as the notion "quantum supremacy" is concerned.

To be precise, our main result is a lower bound on the sample complexity of distinguishing the post selected Boson-Sampling distribution from the uniform one for symmetric probabilistic algorithms. We will give a precise definition of symmetric probabilistic algorithms in Section 5 (Definition 5), but essentially a probabilistic decision algorithm is called symmetric if its output distribution is invariant under relabeling the elements of the sample space.

Theorem 1 (Distinguishing the post selected Boson-Sampling distribution from the uniform one)**.** *If $U \sim \mu_H$, i.e., U is drawn from the Haar measure, and $m \in \Omega(n^\nu)$ with $\nu > 2$, then with probability supra-exponentially small in n no symmetric probabilistic algorithm can distinguish the post-selected Boson-Sampling distribution $\mathcal{D}^*_U$ from the uniform distribution on $\Phi^*_{m,n}$ from fewer than $\Omega(\mathrm{e}^{n/2})$ many samples.*

Proof. The theorem is an immediate corollary of our Theorems 8 and 13. □

Notice that the hardness results of Ref. [3] requires that $\nu > 5$, and $\nu > 2$ is known to be necessary for the proof strategy used there to work, so our theorem fully covers the interesting parameter range.

Without any a priori knowledge about the distribution the labels of the elements of the sample space have no meaning to the certifier. Thus, in the black box setting any decision reached following a non-symmetric algorithm seems arbitrary and cannot qualify as a conclusion, but at the same time, as Theorem 1 shows, symmetric algorithms are essentially useless to distinguish the post selected Boson-Sampling distribution from the uniform one.

We now argue that the above theorem is relevant for the problem of certifying a real Boson-sampling device. Importantly, in a realistic situation the certifier knows the specific unitary U implemented by the supposed Boson-Sampling device. In some particular cases, for example, when U is such that it has some special structure, e.g., such that some outcomes are particularly likely to occur or some have probability zero [11, 12], this knowledge could be used to construct a non-symmetric decision algorithm, thus opening up the possibility to drastically reduce the sample complexity. However, this seems very implausible in the interesting instances, i.e., the ones that are covered by the hardness proof, precisely due to the fact that it is believed to be $\#P$ hard to approximate the probabilities of individual outcomes for $U \sim \mu_H$.

We can make a similar statement about the full Boson-Sampling distribution:

Theorem 2 (Distinguishing the Boson-Sampling distribution from the uniform one)**.** *If $U \sim \mu_H$ and $m \in \Omega(n^\nu)$ with $\nu > 3$, then with probability supra-exponentially small in n no symmetric probabilistic algorithm can distinguish the Boson-Sampling distribution $\mathcal{D}_U$ from the uniform distribution over $\Phi_{m,n}$ from fewer than $\Omega(\mathrm{e}^{n/2})$ many samples.*

Proof. The theorem is an immediate corollary of Theorem 8 and 12. □

As said earlier, Ref. [3] requires that $\nu > 5$, however, it is believed that m growing faster than quadratically with n is sufficient for hardness, which leaves open a parameter range not covered by Theorem 2. At the same time we have good reasons to believe that this is merely a technicality and that $\nu > 2$ is already sufficient for the statements of Theorem 2 to be valid (see the discussion in Section 6 and Theorem 13).

In the latter case of the full Boson-Sampling distribution $\mathcal{D}_U$ the restriction to symmetric algorithms is arguably less natural, mainly because it is known that bosons tend to *bunch* or *cluster* [11–13]. That is, output sequences $(s_1, \ldots, s_m) \in \Phi_{m,n}$ with "collisions", i.e., ones in which at least one s_j is larger than one are, on average over $U \sim \mu_H$, more likely than in the uniform distribution over $\Phi_{m,n}$ (although not dramatically more likely, see Theorem 13.4 in Ref. [3]). This could potentially be used to distinguish the Boson-sampling distribution from the uniform one by a non-symmetric algorithm. We argue that this does not qualify as a certification of a provably hard task. This is because the proof of Ref. [3] only considers the distribution on the Boson-Sampling distribution on the collision-free sector $\Phi^*_{m,n}$. Hence, checking that the output distribution shows the correct bunching cannot help to corroborate that the output distribution is covered by the hardness proof of Ref. [3].

This is related to another subtlety that is important to correctly understand the meaning of our results. The hardness result of Ref. [3] covers all distributions that are at most a small distance away in 1-norm from the ideal Boson-Sampling distribution. Hence, a device that certifies that a black box samples from a distribution that is covered by the hardness results of Ref. [3] does not necessarily need to accept with probability at least $2/3$ on the ideal Boson-Sampling distribution, it is in principle sufficient if it does so on some distribution inside this 1-norm ball. The 1-norm ball includes distributions that are not exponentially flat and which can be distinguished from the uniform distribution from polynomially large number of samples using a symmetric algorithm [14]. Symmetric certification algorithms with polynomial sample complexity for these distributions thus cannot be excluded by our results.

There is a further subtlety: Consider a device that with probability $1 - \epsilon$ outputs a sample from an ideal Boson-Sampling device and with probability ϵ outputs a specific sample that encodes the solution to an NP-complete problem. The output distribution of this device would be ϵ close to the Boson-Sampling distribution in 1-norm. At the same time, it can be certified from $\mathrm{O}(1/\epsilon)$ many samples, using a simple but non-symmetric algorithm, that the device is implementing a hard sampling problem by simply identifying the special outcome and checking that it is indeed a solution to the NP-complete problem. Even though this distribution is covered by the hardness results of Ref. [3], one would hardly say that its (certifiable) hardness is a consequence of the hardness of Boson-Sampling.

One might also consider the following alternative certification scenario. Assume one already has a certified Boson-Sampling device, then one could try to certify another device by comparing the samples they output. Again, our technical results, Theorem 8 and 12, imply that with high probability this cannot be done using a symmetric algorithm and less than exponentially many samples.

Finally, it is important to note that our findings do not contradict the results of Ref. [3].

4 Upper bounds on the sample complexity in the state discrimination setting

In the state discrimination setting the certifier has the promise that the given sampling device samples from one of two known distributions $\mathcal{P}$ or $\mathcal{Q}$. In particular he has knowledge of the sample space and all the probabilities that each of the two candidate distributions assign to the elements of this space. The certifier's aim is to minimize the probability of wrongfully answering $\mathcal{Q}$ if the true distribution is $\mathcal{P}$ (error of the first kind) and that of wrongfully an-

swering $\mathcal{P}$ if the true distribution is $\mathcal{Q}$ (error of the second kind). This minimization can be done in various different ways. For example one can minimize the (weighted) sum of the two probabilities, or minimize one while the other is kept constant or suppressed exponentially in the number of samples with a predefined rate. The asymptotic behavior of the number of samples needed in these situations has been extensively studied in both the classical [15, 16] and quantum setting [17–20], in which the probability distributions are replaced by quantum states. See also the introduction of Ref. [21] for a short review and [22] for further references.

Only recently, in Ref. [21], bounds on the error probabilities for finite sample sizes were derived. They hold in both the classical and the quantum setting, but here we will only need the classical versions. They imply that in the state discrimination setting the Boson-Sampling distribution can be distinguished from the uniform distribution from a polynomial number of samples:

Theorem 3 (Lower bound on the sample complexity of distinguishing the Boson-Sampling distribution from the uniform one in the state discrimination setting)**.** *Let $\epsilon > 0$ and $m \leq c\, n^\nu$ for some $c \geq 0, \nu \geq 1$. Then for any $\gamma > 0$ there exists a constant $C > 0$, such that for and any instance of Boson-Sampling with n bosons in m modes, whose distribution is at least ϵ far from the uniform distribution in 1-norm, there is an algorithm that distinguishes the former from the latter from $C\, n^{2+\gamma}$ many samples.*

Note that the above theorem covers all instances of Boson-Sampling that can potentially be hard to sample from approximately in 1-norm.

Proof. For $0 \leq t \neq 1$ and two probability distributions over a finite sample space Φ we define the *t-Rényi relative entropy* of $\mathcal{P}$ given $\mathcal{Q}$

$$S_t(\mathcal{P}\|\mathcal{Q}) \coloneqq \begin{cases} \dfrac{1}{t-1} \ln \displaystyle\sum_{S \in \mathrm{supp}(\mathcal{P}) \cap \mathrm{supp}(\mathcal{Q})} \Pr_{\mathcal{P}}[S]^t \Pr_{\mathcal{Q}}[S]^{1-t} & \text{if } \mathrm{supp}(\mathcal{P}) \subseteq \mathrm{supp}(\mathcal{Q}) \\ \infty & \text{otherwise} \end{cases}, \tag{7}$$

where $\mathrm{supp}(\mathcal{P}) = \{S \in \Phi : \Pr_{\mathcal{P}}[S] \neq 0\}$ and we set $\ln 0 = -\infty$. As $\mathcal{P}$ is normalized the limit $t \to 1$ exists [21] and $S(\mathcal{P}\|\mathcal{Q}) \coloneqq \lim_{t\to 1} S_t(\mathcal{P}\|\mathcal{Q})$ is called *relative entropy* of $\mathcal{P}$ given $\mathcal{Q}$.

Let $\beta_{l,\alpha}$ be the optimal achievable error of the second kind in the state discrimination setting after receiving l samples when the error of the first kind is upper bounded by α. Theorem 3.3 in Ref. [21] implies that for all $l, \alpha > 0$

$$\frac{1}{l} \ln \beta_{l,\alpha} \leq -S(\mathcal{P}\|\mathcal{Q}) + \frac{1}{\sqrt{l}} 4\sqrt{2} \ln(\alpha^{-1}) \ln \eta - \frac{2 \ln 2}{l}, \tag{8}$$

where

$$\eta = 1 + \mathrm{e}^{S_{3/2}(\mathcal{P}\|\mathcal{Q})/2} + \mathrm{e}^{-S_{1/2}(\mathcal{P}\|\mathcal{Q})/2}. \tag{9}$$

Since

$$\eta \leq 2 + \mathrm{e}^{S_{3/2}(\mathcal{P}\|\mathcal{Q})/2} \leq \mathrm{e}^{S_{3/2}(\mathcal{P}\|\mathcal{Q})/2 + \ln 3} \tag{10}$$

this implies

$$\frac{1}{l} \ln \beta_{l,\alpha} \leq -S(\mathcal{P}\|\mathcal{Q}) + \frac{1}{\sqrt{l}} 4\sqrt{2} \ln(\alpha^{-1}) \left(S_{3/2}(\mathcal{P}\|\mathcal{Q})/2 + \ln 3\right). \tag{11}$$

Theorem 1.15 in Ref. [23] implies the first of the following inequalities

$$S(\mathcal{P}\|\mathcal{Q}) \geq \frac{1}{2}\|\mathcal{P}-\mathcal{Q}\|_1^2 \geq \frac{1}{2}\epsilon^2, \tag{12}$$

the second is implied by the assumptions of the Theorem. Moreover, if $\mathcal{Q}$ is the uniform distribution over Φ, then

$$S_{3/2}(\mathcal{P}\|\mathcal{Q}) \leq S_2(\mathcal{P}\|\mathcal{Q}) = \ln(|\Phi| \sum_{S\in\Phi} \Pr_{\mathcal{P}}[S]^2) \leq \ln|\Phi|. \tag{13}$$

Hence, for $\Phi = \Phi_{m,n}$ and if $m \leq c\,n^\nu$ we have by Eq. (6)

$$S_{3/2}(\mathcal{P}\|\mathcal{Q}) \leq n\ln(2\,(c+1)\,\mathrm{e}) + n\,(\nu-1)\ln(n). \tag{14}$$

This implies that for $\alpha = 1/3$, $\mathcal{P}$ the Boson-Sampling distribution with $m \leq c\,n^\nu$, and $\mathcal{Q}$ the uniform distribution over $\Phi_{m,n}$, a number of samples l scaling like $l \in \Omega(n^{2+\gamma})$, for any $\gamma > 0$, is sufficient to make the right hand side of Eq. (11) negative, and thereby $\beta_{l,1/3} \leq 1/3$ for sufficiently large n. □

5 Sample complexity in the black box setting

In this section we give lower bounds on the sample complexity of decision problems in the black box setting. Apart from the scenario relevant for the certification of Boson-Sampling, in which the certifier is given a black box and is asked to distinguish the two cases where it samples from the Boson-Sampling distribution or the uniform one, we will also cover scenarios where the certifier is given two or more black boxes and is asked to decide whether they sample from the same or form different probability distributions (see also Ref. [14]).

The lower bounds are ultimately a consequence of a variant of the birthday paradox for ϵ-flat probability distributions. We call a probability distribution $\mathcal{P}$ over a finite sample space ϵ-*flat* if $\|\mathcal{P}\|_\infty \leq \epsilon$, i.e., all probabilities are smaller than ϵ, or equivalently if $\mathcal{P}$ has min entropy $H_\infty \geq -\log_2\epsilon$.

Lemma 4 (Non-uniform non-identically distributed birthday paradox)**.** *The probability $\bar{p}(l,|\Phi|,\epsilon)$ that l samples drawn independently from not necessarily identical ϵ-flat distributions over a finite sample space Φ are all different fulfills*

$$\forall l \leq 1 + 1/(2\epsilon): \quad \bar{p}(l,|\Phi|,\epsilon) \geq 2^{-l^2\epsilon}. \tag{15}$$

Proof. The probability $\bar{p}(l,|\Phi|,\epsilon)$ that all samples are different is bounded by

$$\bar{p}(l,|\Phi|,\epsilon) \geq \prod_{j=1}^{l-1}(1-j\epsilon). \tag{16}$$

If $1/2 \leq 1-\epsilon\,(l-1) \leq 1$, we have

$$\forall j \in [l-1]: \quad 1-j\,\epsilon \geq 2^{-2\,j\,\epsilon}. \tag{17}$$

This implies that for sufficiently large l

$$\bar{p}(l, |\Phi|, \epsilon) \geq \prod_{j=1}^{l-1}(1 - j\,\epsilon) \geq \prod_{j=1}^{l-1} 2^{-2\,j\,\epsilon} \tag{18}$$

$$= 2^{-2\sum_{j=1}^{l-1} j\,\epsilon} = 2^{-l\,(l-1)\epsilon} \geq 2^{-l^2\epsilon}. \tag{19}$$

□

Now, we consider the situation of N black boxes that sample each from one of N probability distributions $(\mathcal{P}^{(j)})_{j=1}^{N}$ over the same finite sample space Φ. For $j \in [N]$ and $l \in \mathbb{Z}^+$, let $\mathcal{S}^{(j)} \coloneqq (S_1^{(j)}, \ldots, S_l^{(j)}) \in \Phi^l$ be sequences of samples of length l from each of the distributions respectively. We will keep the discussion in this chapter general but will later mostly be interested in the case $N = 1$.

The certifier works under the assumption that the sampling device outputs independent identically distributed samples. Hence, the order of the samples in each sequence should not influence the certifier's decision. Moreover, in the black box setting the certifier is assumed to have no a priori knowledge about the distribution. If in addition the decision problem of the certifier is invariant under a relabeling of the sample space, its decision should be independent of which element of the sample space is assigned which label. If this is not the case it cannot qualify as a conclusion reached based on the samples. Therefore, for tasks such as deciding whether a given black box is sampling from the uniform distribution or not, or deciding whether a number of black boxes sample from the same or from different distributions the certifier should follow a *symmetric probabilistic algorithm*.

Definition 5 (Symmetric probabilistic algorithm)**.** *An algorithm that takes as input for each $j \in [N]$ a sequence of samples $\mathcal{S}^{(j)} \subset \Phi^l$ and probabilistically outputs either "accept" or "reject" is called a* symmetric probabilistic algorithm *if its output distribution is invariant under permuting the samples in each sequences*

$$\forall j \in [N]: \qquad (S_1^{(j)}, \ldots, S_l^{(j)}) \mapsto (S_{\tau_j(1)}^{(j)}, \ldots, S_{\tau_j(l)}^{(j)}), \quad \tau_j \in \mathrm{Sym}([l]), \tag{20}$$

and relabeling of the sample space Φ, i.e., the action of $\mathrm{Sym}(\Phi)$ on all $\mathcal{S}^{(j)}$ simultaneously

$$\forall j \in [N]: \qquad (S_1^{(j)}, \ldots, S_l^{(j)}) \mapsto (\kappa(S_1^{(j)}), \ldots, \kappa(S_l^{(j)})), \quad \kappa \in \mathrm{Sym}(\Phi). \tag{21}$$

Following Ref. [14] we define the *fingerprint tensor* $C((\mathcal{S}^{(j)})_{j=1}^{N}) \in \mathbb{N}^{(l+1)\times\cdots\times(l+1)}$ of the sequences of samples, such that for all $k_1, \ldots, k_N \in \{0, 1, \ldots, l\}$, $C_{k_1,\ldots,k_N}$ is the number of elements in Φ that for all $j \in [N]$ appear exactly k_j times in the j-th sequence of samples $\mathcal{S}^{(j)}$. Obviously $\sum_{k_1,\ldots,k_N=0}^{l} C_{k_1,\ldots,k_N} = |\Phi|$. For $N = 1$ this construction results in the fingerprint vector

$$C_{k_1} \coloneqq |\{S' \in \Phi : |\{S \in \mathcal{S}^{(1)} : S = S'\}| = k_1\}| \tag{22}$$

and for $N = 2$ the result is the fingerprint matrix

$$\begin{aligned} C_{k_1,k_2} \coloneqq |\{S' \in \Phi :& |\{S^{(1)} \in \mathcal{S}^{(1)} : S^{(1)} = S'\}| = k_1 \\ &\text{and } |\{S^{(2)} \in \mathcal{S}^{(2)} : S^{(2)} = S'\}| = k_2\}|. \end{aligned} \tag{23}$$

For example, if $\Phi = [6]$, $S^{(1)} = (1, 5, 1, 1, 2)$ and $S^{(2)} = (2, 6, 1, 4, 6)$, then the fingerprint

matrix is given by

$$C = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{24}$$

The fingerprint tensor encodes all the information contained in the samples that is invariant under permuting the labels of the sample space or reordering the samples in each sequence. That is, the sequences of samples can be reconstructed exactly from the fingerprint up to a permutation of the samples in each sequence and a global relabeling of the sample space [14]. This proves:

Lemma 6 (Symmetric algorithms and the fingerprint (see also Ref. [14]))**.** *For every symmetric probabilistic algorithm $\mathcal{A}$ there is exists an algorithm $\mathcal{A}'$ which has the same output distribution as $\mathcal{A}$, but takes as input the fingerprint of the sequences of samples.*

Denote by $\mathcal{D}_C((\mathcal{P}^{(j)})_{j=1}^N, l)$ the probability distribution on fingerprint tensors induced by drawing l samples from each $\mathcal{P}^{(j)}$, and then constructing the corresponding fingerprint tensor, and when we write $C \sim \mathcal{D}_C((\mathcal{P}^{(j)})_{j=1}^N, l)$ we mean C drawn from $\mathcal{D}_C((\mathcal{P}^{(j)})_{j=1}^N, l)$. For each $|\Phi|$, N, and l there is a unique *trivial fingerprint* tensor $\tilde{C}$ that characterizes the situation where no sample appears more than once. For $N = 1$ this is the vector

$$\tilde{C} = (|\Phi| - l, l, 0, \ldots, 0) \in \mathbb{N}^{l+1}, \tag{25}$$

and for $N = 2$ it is the matrix

$$\tilde{C} := \begin{pmatrix} |\Phi| - 2\,l & l & 0 & \cdots & 0 \\ l & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix} \in \mathbb{N}^{(l+1)\times(l+1)}. \tag{26}$$

Due to the birthday paradox, fingerprint tensors constructed from few samples are trivial with high probability.

Lemma 7 (Fingerprint tensors from few samples)**.** *Let $N \in \mathbb{Z}^+$ and $(P^{(j)})_{j=1}^N$ be ϵ-flat probability distributions over a finite sample space Φ. If $l \in \mathrm{O}((1/\epsilon)^{1/4})$ many samples are drawn from each $\mathcal{P}^{(j)}$ then*

$$\Pr_{C\sim\mathcal{D}_C((\mathcal{P}^{(j)})_{j=1}^N,l)}[C \neq \tilde{C}] \in \mathrm{O}(N^2\sqrt{\epsilon}). \tag{27}$$

Proof. Let $a > 0$ and $l \leq a\,(1/\epsilon)^{1/4}$ and denote for each $j \in [N]$ by $\mathcal{S}^{(j)}$ the sequence of l samples drawn from $\mathcal{P}^{(j)}$. Since all the $\mathcal{P}^{(j)}$ are ϵ-flat probability distributions over Φ we have

$$\Pr_{C\sim\mathcal{D}_C((\mathcal{P})_{j=1}^N,l)}\left[\exists S' \in \Phi : \sum_{j=1}^{N} |\{S \in \mathcal{S}^{(j)} : S = S'\}| > 1\right] = \bar{p}(N\,l, |\Phi|, \epsilon), \tag{28}$$

with $\bar{p}$ as in Lemma 4. For sufficiently small ϵ Lemma 4 yields

$$\bar{p}(N\,l, |\Phi|, \epsilon) \leq 2^{-(N\,a)^2\sqrt{\epsilon}} \tag{29}$$

and

$$\Pr_{C\sim\mathcal{D}_C((\mathcal{P})_{j=1}^N,l)}[C \neq \tilde{C}] \leq 1 - 2^{-(N\,a)^2\sqrt{\epsilon}} \leq (N\,a)^2\sqrt{\epsilon}. \tag{30}$$

□

Similar results can be obtained for all scalings $l \in \mathrm{O}((1/\epsilon)^\alpha)$ with $\alpha < 1/2$, but $\alpha = 1/4$ is good enough for our purposes and yields the particularly simple result stated above.

Theorem 8 (Symmetric algorithms and ϵ-flat distributions)**.** *For every symmetric probabilistic algorithm there exists a trivial output distribution such that the output distribution of the algorithm after receiving at most $\mathrm{O}((1/\epsilon)^{1/4})$ many samples from each of N black boxes sampling from ϵ-flat distributions is with probability $1 - \mathrm{O}(N^2\sqrt{\epsilon})$ equal to the trivial output distribution and hence, in particular, does not depend on which ϵ-flat distributions were used to generate the samples.*

Proof. By Lemma 6 any symmetric probabilistic algorithm is equivalent to an algorithm that only receives the fingerprint of the samples as input. If all input distributions are ϵ-flat, then by Lemma 7, if at most $\mathrm{O}((1/\epsilon)^{1/4})$ samples are drawn from each distribution, the probability that their fingerprint is non-trivial is of order $\mathrm{O}(N^2\sqrt{\epsilon})$. The result follows and the trivial output distribution is the output distribution corresponding to samples with the trivial fingerprint. □

By strengthening Lemma 7, as pointed out after its proof, a result similar to Theorem 8 can be obtained for the number of samples scaling like $\mathrm{O}((1/\epsilon)^\alpha)$ for all $\alpha < 1/2$.

6 The Boson-Sampling distribution is flat

In this section we show that the Boson-Sampling distribution is extremely flat with high probability. The strategy is as follows: First we relate the probability measure induced on the matrices U_S described in Section 2 to a Gaussian measure $\mu_{G_S(\sigma)}$. Then we use measure concentration for $\mu_{G_S(\sigma)}$ to prove ϵ-flatness.

A crucial step in the proof of the main result of Ref. [3] is to show that if m is sufficiently large compared to n and $U \sim \mu_H$, i.e., U is chosen from the Haar measure μ_H on $U(m)$, then, for any fixed $S \in \Phi^*_{m,n}$, the measure on $\mathbb{C}^{n\times n}$ induced by the map $g_S = (U \mapsto U_S)$ is close to $\mu_{G(1/\sqrt{m})}$, where $\mu_{G(\sigma)}$ is the measure obtained by choosing the real and imaginary part of every entry of an $n \times n$ matrix independently from a Gaussian distribution with mean zero and standard deviation σ.

Lemma 9 (Theorem 5.2 in Ref. [3])**.** *Let $f : \mathbb{C}^{n\times n} \to [0, 1]$ be measurable and $\delta > 0$ with the property that $m \geq (n^5/\delta)\ln^2(n/\delta)$. Then*

$$\forall S \in \Phi^*_{m,n} : \quad \mathop{\mathbb{E}}_{U\sim\mu_H} f(U_S) \leq (1 + \mathrm{O}(\delta)) \mathop{\mathbb{E}}_{X\sim\mu_{G(1/\sqrt{m})}} f(X). \tag{31}$$

It is known that $m \geq c\,n^\nu$ with $\nu > 2$ and $0 < c \in \mathrm{O}(1)$ is necessary for closeness of $\mu_H \circ g_S^{-1}$ and $\mu_{G(1/\sqrt{m})}$. As this is a crucial ingredient to the proof of hardness of Ref. [3], we will from now on assume that $m \geq c\,n^\nu$ with $\nu > 2$ and $0 < c \in \mathrm{O}(1)$.

Lemma 9 is not strong enough for our purpose, as we must be able to control all of $\Phi_{m,n}$ and not only the collision-free subspace. Fortunately, the above lemma extends naturally to all $S \in \Phi_{m,n}$, but first we need some notation: For every sequence S, let $\tilde{S}$ be the sequence obtained from S by removing all the zeros, i.e,

$$\tilde{S} = (\tilde{s}_1, \ldots, \tilde{s}_{|\tilde{S}|}) \coloneqq (s \in S : s > 0). \tag{32}$$

Let $\mu_{G_S(\sigma)}$ be the probability measure on $\mathbb{C}^{n\times n}$ obtained by drawing the real and imaginary part of every entry of a $|\tilde{S}| \times n$ matrix independently from a Gaussian distribution with mean zero and standard deviation σ and then for all $j \in [|\tilde{S}|]$ taking $\tilde{s}_j$ copies of the j^{th} row of this matrix.

Lemma 10 (Multiplicative error bound)**.** *Let $f : \mathbb{C}^{n\times n} \to [0,1]$ be measurable and $\delta > 0$ with the property that $m \geq (n^5/\delta)\ln^2(n/\delta)$. Then for all $S \in \Phi_{m,n}$*

$$\mathop{\mathbb{E}}_{U\sim\mu_H} f(U_S) \leq (1 + \mathrm{O}(\delta)) \mathop{\mathbb{E}}_{X\sim\mu_{G_S(1/\sqrt{m})}} f(X). \tag{33}$$

Proof. Let $S \in \Phi_{m,n}$, $\tilde{S}$ as in Eq. (32) and $m' \coloneqq |\tilde{S}|$. Define v to be the sequence containing $\tilde{s}_j$ times the integer j for every $j \in [m']$ in increasing order and w the sequence containing the positions of each of the first of the repeated rows in U_S, i.e.,

$$v \coloneqq (\underbrace{1,\ldots,1}_{\tilde{s}_1}, \underbrace{2,\ldots,2}_{\tilde{s}_2}, \ldots, \underbrace{m',\ldots,m'}_{\tilde{s}'_m}) \in (\mathbb{Z}^+)^n, \tag{34}$$

$$w \coloneqq (1, 1+\tilde{s}_1, 1+\tilde{s}_1+\tilde{s}_2, \ldots, 1+\sum_{j=1}^{m'-1}\tilde{s}_j) \in (\mathbb{Z}^+)^{m'}. \tag{35}$$

The sequence v defines a linear embedding $\eta : \mathbb{C}^{m'\times n} \to \mathbb{C}^{n\times n}$ component wise by

$$\eta(Y)_{i,j} \coloneqq Y_{v_i,j} \quad \forall i,j \in [n], \tag{36}$$

i.e., $\eta(Y)$ has s_j copies of the j-th row of Y. The sequence w defines a linear projection $\pi : \mathbb{C}^{n\times n} \to \mathbb{C}^{m'\times n}$ by

$$\pi(X)_{i,j} \coloneqq X_{w_i,j} \quad \forall i \in [m'],\ j \in [n], \tag{37}$$

in particular, $\pi(U_S)$ contains only the first out of each series of the repeated rows in U_S. Note that $\eta \circ \pi : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ is a projection onto the subspace of matrices that have the same repetition structure as U_S. Let

$$f_S \coloneqq f \circ \eta \circ \pi, \tag{38}$$

then $f_S(U_S) = f(U_S)$ only depends on the first of the repeated rows in U_S and is independent of all the other rows. Since the Haar measure is permutation-invariant,

$$\mathop{\mathbb{E}}_{U\sim\mu_H} f_S(U_S) = \mathop{\mathbb{E}}_{U\sim\mu_H} f_S(U_{1_n}). \tag{39}$$

Hence, Lemma 9 yields the inequality in the calculation

$$\begin{aligned} \mathop{\mathbb{E}}_{U\sim\mu_H} f(U_S) &= \mathop{\mathbb{E}}_{U\sim\mu_H} f_S(U_{1_n}) && (40)\\ &\leq (1+\mathrm{O}(\delta)) \mathop{\mathbb{E}}_{X\sim\mu_{G(1/\sqrt{m})}} f_S(X) && (41)\\ &= (1+\mathrm{O}(\delta)) \mathop{\mathbb{E}}_{X\sim\mu_{G_S(1/\sqrt{m})}} f(X), && (42) \end{aligned}$$

which finishes the proof.

□

In addition to the multiplicative error bound we need a concentration result for the Gaussian measure $\mu_{G_S(\sigma)}$.

Lemma 11 (Concentration of the Gaussian measure $\mu_{G_S(\sigma)}$)**.** *For all $n, m \in \mathbb{Z}^+$, all $S \in \Phi_{m,n}$ and all $\xi > 0$*

$$\Pr_{X\sim\mu_{G_S(\sigma)}}\left[\max_{j,k\in[n]} |x_{j,k}| \geq \xi\right] \leq 1 - \left(1 - \mathrm{Erfc}\left(\frac{\xi}{\sqrt{2}\,\sigma}\right)\right)^{n^2}. \tag{43}$$

Proof. For Gaussian random variables we have

$$\forall \xi > 0,\ j,k \in [n]: \quad \Pr_{X\sim\mu_{G(\sigma)}}[|x_{j,k}| \geq \xi] = \mathrm{Erfc}\left(\frac{\xi}{\sqrt{2}\,\sigma}\right) \tag{44}$$

where

$$\mathrm{Erfc}\left(\frac{\xi}{\sqrt{2}\,\sigma}\right) := 2\int_{\xi}^{\infty} \frac{\mathrm{e}^{-\frac{x^2}{2\,\sigma^2}}}{\sqrt{2\,\pi\,\sigma^2}}\,\mathrm{d}x \tag{45}$$

is the complementary error function. This implies that

$$\forall \xi > 0: \quad \Pr_{X\sim\mu_{G(\sigma)}}[\forall j,k \in [n] : |x_{j,k}| \leq \xi] = \left(1 - \mathrm{Erfc}\left(\frac{\xi}{\sqrt{2}\,\sigma}\right)\right)^{n^2}. \tag{46}$$

It is also true that

$$\begin{aligned} \forall S \in \Phi_{m,n}, \xi > 0: \quad & \Pr_{X\sim\mu_{G_S(\sigma)}}[\forall j,k \in [n] : |x_{j,k}| \leq \xi] \\ & \geq \Pr_{X\sim\mu_{G(\sigma)}}[\forall j,k \in [n] : |x_{j,k}| \leq \xi], \end{aligned} \tag{47}$$

because the additional dependency of the entries of $X \sim \mu_{G_S(\sigma)}$ only decreases the chance of having an exceptionally large entry. □

Theorem 12 (Flatness of the Boson-Sampling distribution)**.** *Let $\nu > 3$. Then for every $m \in \Omega(n^\nu)$*

$$-\ln\left(\Pr_{U\sim\mu_H}\left[\exists S \in \Phi_{m,n} : \Pr_{\mathcal{D}_U}[S] \geq \mathrm{e}^{-2\,n}\right]\right) \in \mathrm{O}\left(n^{\nu-2-1/n}\right). \tag{48}$$

Proof. Using the union bound (also known as Boole's inequality) we obtain that for every

$\epsilon > 0$

$$\Pr_{U\sim\mu_H}\left[\exists S \in \Phi_{m,n} : \Pr_{\mathcal{D}_U}[S] \geq \epsilon\right] \tag{49}$$

$$\leq \sum_{S\in\Phi_{m,n}} \Pr_{U\sim\mu_H}\left[\Pr_{\mathcal{D}_U}[S] \geq \epsilon\right] \tag{50}$$

$$\leq |\Phi_{m,n}| \max_{S\in\Phi_{m,n}} \Pr_{U\sim\mu_H}\left[\Pr_{\mathcal{D}_U}[S] \geq \epsilon\right] \tag{51}$$

$$= |\Phi_{m,n}| \max_{S\in\Phi_{m,n}} \Pr_{U\sim\mu_H}\left[\frac{|\operatorname{Perm}(U_S)|^2}{\prod_{j=1}^m (s_j!)} \geq \epsilon\right]. \tag{52}$$

Applying Lemma 10 for the S that yields the maximum with $\delta = n$ and the indicator function

$$f(U_S) = \begin{cases} 1 & \text{if } \frac{|\operatorname{Perm}(U_S)|^2}{\prod_{j=1}^m (s_j!)} \geq \epsilon \\ 0 & \text{otherwise} \end{cases} \tag{53}$$

yields

$$\begin{aligned} &\Pr_{U\sim\mu_H}\left[\exists S \in \Phi_{m,n} : \Pr_{\mathcal{D}_U}[S] \geq \epsilon\right] \\ &\leq (1+\mathrm{O}(n))\,|\Phi_{m,n}| \max_{S\in\Phi_{m,n}} \Pr_{X\sim\mu_{G_S(1/\sqrt{m})}}\left[\frac{|\operatorname{Perm}(X)|^2}{\prod_{j=1}^m (s_j!)} \geq \epsilon\right]. \end{aligned} \tag{54}$$

Recall that the permanent $\operatorname{Perm}(X)$ of a matrix $X = (x_{j,k}) \in \mathbb{C}^{n\times n}$ is defined as

$$\operatorname{Perm}(X) \coloneqq \sum_{\tau\in\operatorname{Sym}([n])} \prod_{j=1}^n x_{j,\tau(j)}, \tag{55}$$

where $\operatorname{Sym}([n])$ is the symmetric group acting on $[n]$. This implies that

$$\frac{|\operatorname{Perm}(X)|^2}{\prod_{j=1}^m (s_j!)} \leq |\operatorname{Perm}(X)|^2 \leq (n!)^2 \left(\max_{j,k\in[n]} |x_{j,k}|\right)^{2n}. \tag{56}$$

Hence, for every $S \in \Phi_{m,n}$ and every $\epsilon > 0$

$$\Pr_{X\sim\mu_{G_S(1/\sqrt{m})}}\left[\frac{|\operatorname{Perm}(X)|^2}{\prod_{j=1}^m (s_j!)} \geq \epsilon\right] \leq \Pr_{X\sim\mu_{G_S(1/\sqrt{m})}}\left[\max_{j,k\in[n]} |x_{j,k}| \geq \left(\frac{\sqrt{\epsilon}}{n!}\right)^{1/n}\right]. \tag{57}$$

Now we use Lemma 11 with

$$\xi = \left(\frac{\sqrt{\epsilon}}{n!}\right)^{1/n}, \tag{58}$$

and Eq. (6) to arrive at

$$\begin{aligned}&\Pr_{U\sim\mu_H}\left[\exists S\in\Phi_{m,n}:\Pr_{\mathcal{D}_U}[S]\geq\epsilon\right]\\&\leq(1+\mathrm{O}(n))\,(2\,(c+1)\,\mathrm{e})^n\,n^{(\nu-1)n}\left(1-\left(1-\mathrm{Erfc}\sqrt{\frac{c}{2}\frac{\epsilon^{1/n}\,n^\nu}{(n!)^{2/n}}}\right)^{n^2}\right).\end{aligned}\tag{59}$$

Bounding the complementary error function by [24]

$$\mathrm{Erfc}\,(x)\leq\mathrm{e}^{-x^2},\tag{60}$$

we obtain

$$1-(1-\mathrm{Erfc}(x))^{n^2}\leq1-\left(1-\mathrm{e}^{-x^2}\right)^{n^2}=1-\sum_{k=0}^{n^2}\binom{n^2}{k}(-\mathrm{e}^{-x^2})^k\tag{61}$$

$$=\sum_{k=1}^{n^2}\binom{n^2}{k}\mathrm{e}^{-x^2k}\,(-1)^{k-1}\leq\sum_{k=1}^{n^2}(n^2\mathrm{e}/k)^k\,\mathrm{e}^{-x^2k}\tag{62}$$

$$=\sum_{k=1}^{n^2}(n^2\,\mathrm{e}^{-x^2+1})^k.\tag{63}$$

If x is sufficiently large such that

$$n^2\,\mathrm{e}^{-x^2+1}\leq\frac{1}{2}<1,\tag{64}$$

the geometric series converges and

$$\sum_{k=1}^{n^2}(n^2\mathrm{e}^{-x^2+1})^k\leq\frac{n^2\mathrm{e}^{-x^2+1}}{1-n^2\mathrm{e}^{-x^2+1}}\tag{65}$$

$$\leq2\,n^2\,\mathrm{e}^{-x^2+1}.\tag{66}$$

Hence, for the bound (59) to become meaningful it is sufficient that the argument of the square root in the error function grows slightly faster than linear with n. Because of the bound $n!\leq\mathrm{e}^{1-n}\,n^{n+1/2}$ (a variant of Stirling's approximation) we have for the argument of the square root in Eq. (59)

$$\frac{c}{2}\frac{\epsilon^{1/n}n^\nu}{(n!)^{2/n}}\geq\frac{c}{2}\frac{\epsilon^{1/n}n^\nu}{\mathrm{e}^{2/n-2}n^{2+1/n}}=\frac{c}{2}\frac{\epsilon^{1/n}}{\mathrm{e}^{2/n-2}}n^{\nu-2-1/n},\tag{67}$$

and with the convenient choice $\epsilon=\mathrm{e}^{-2n}$ it follows that for all $\nu>3$

$$\begin{aligned}&\Pr_{U\sim\mu_H}\left[\exists S\in\Phi_{m,n}:\Pr_{\mathcal{D}_U}[S]\geq\mathrm{e}^{-2n}\right]\\&\in\mathrm{O}\left(n^3\,(2\,(c+1)\,\mathrm{e})^n\,n^{(\nu-1)n}\exp(-c\,\mathrm{e}^{-2/n}\,n^{\nu-2-1/n}/2)\right).\end{aligned}\tag{68}$$

□

The above proof of Theorem 12 yields the result only for $\nu > 3$. This is a consequence of the $n!$ prefactor introduced in the extremely crude bound on the permanent used in Eq. (56). In fact, it is known that [3]

$$\mathop{\mathbb{E}}_{X\sim\mu_{G(1/\sqrt{m})}}[|\operatorname{Perm}(X)|^2] = 2^n\, n!\, m^{-n}, \tag{69}$$

so it seems likely that, the inequality in Eq. (56) can be replaced by an inequality that is fulfilled with high probability and has a $\sqrt{n!}$ prefactor instead of the $n!$.

For all S in the collision-free subspace $\Phi^*_{m,n}$ we can show the improved bound:

Theorem 13 (Flatness of the Boson-Sampling distribution on the collision-free subspace)**.** *Let $\nu > 1$. Then for every $1 > \epsilon > 0$ and $m \in \Omega(n^\nu)$*

$$-\ln\left(\Pr_{U\sim\mu_H}\left[\exists S \in \Phi^*_{m,n} : \Pr_{\mathcal{D}_U}[S] \geq \epsilon\right]\right) \in \mathrm{O}\left((\nu-1)\, n \ln n\right) - 2\ln(1/\epsilon), \tag{70}$$

and in particular

$$-\ln\left(\Pr_{U\sim\mu_H}\left[\exists S \in \Phi^*_{m,n} : \Pr_{\mathcal{D}_U}[S] \geq n^{-n/2}\right]\right) \in \mathrm{O}\left((\nu-2)\, n \ln n\right). \tag{71}$$

Proof. It is known that [3, 25]

$$\mathop{\mathbb{E}}_{X\sim\mu_{G(1/\sqrt{m})}}[|\operatorname{Perm}(X)|^4] = 2^{2n}(n!)^2\,(n+1)\, m^{-2n}. \tag{72}$$

Hence, by using Markov's inequality for the positive random variable $|\operatorname{Perm}(X)|^4$ with $m = cn^\nu$ we find that for every $\epsilon > 0$

$$\Pr_{X\sim\mu_{G(1/\sqrt{m})}}[|\operatorname{Perm}(X)|^2 \geq \epsilon] \leq 2^{2\,n}\,(n!)^2\,(n+1)\, c^{-2n}\, n^{-2\,\nu\,n}\epsilon^{-2}. \tag{73}$$

Using again the bound $n! \leq \mathrm{e}^{1-n}\, n^{n+1/2}$ this implies

$$\Pr_{X\sim\mu_{G(1/\sqrt{m})}}[|\operatorname{Perm}(X)|^2 \geq \epsilon] \leq n(n+1)2^{2\,n}\,\mathrm{e}^{2-2\,n}\, c^{-2\,n}\, n^{2\,(1-\nu)\,n}\,\epsilon^{-2}. \tag{74}$$

Hence, by Eq. (6)

$$\begin{aligned} &|\Phi_{m,n}| \max_{S\in\Phi^*_{m,n}} \Pr_{X\sim\mu_{G_S(1/\sqrt{m})}}\left[\frac{|\operatorname{Perm}(X)|^2}{\prod_{j=1}^m (s_j!)} \geq \epsilon\right] \\ &\leq (2\,(c+1)\,\mathrm{e})^n\, n\,(n+1)\, 2^{2\,n}\,\mathrm{e}^{2-2\,n}\, c^{-2\,n}\, n^{(1-\nu)\,n}\,\epsilon^{-2}. \end{aligned} \tag{75}$$

Inserting this into Eq. (54) and taking the logarithm yields the first bound, the choice $\epsilon = n^{-n/2}$ the second bound. □

A derivation of a similar bound for all $S \in \Phi_{m,n}$ would prove the statement of Theorem 12 under a weaker condition on ν. We conjecture that the statement of Theorem 12 is true for all $\nu > 2$.

7 Efficiently simulatable instances in 1-norm

In this section we finally ask the question in what settings one can expect an efficient classical simulation to be feasible even up to a small error in 1-norm. After all, any experiment will not realise the precise ideal Boson-Sampling setting, but instead an imperfect approximation thereof. This may provide room for the efficient classical simulation of the output distribution actually obtained. Subsequently, we will identify a setting of this kind, which resembles those implementable with present-day linear optical circuits. It is not claimed that the discussed scenario exactly matches realistic experiments, but it does share many features. We will show that efficient classical 1-norm approximate sampling is possible under the following conditions:

Condition 1: The input state $|1_n\rangle$ is replaced by a Gaussian product state ρ [26, 27]. Sources that produce such states are common in quantum optical implementations. In practice, many single photon sources provide approximately coherent states or mixed Gaussian states instead of states for which the probability of having more than a single photon is zero. If single photon sources are being generated by heralding [28] a source of (Gaussian) two-mode squeezed states, the argument presented here is still valid: After all, the entire statistics, including the heralding events, is then classically simulatable.

Condition 2: The unitary $\varphi(U)$ with $U \in U(m)$, specifying the optical network, is replaced by a Gaussian completely positive map $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, a Gaussian channel [29]. Such operations cover the ideal unitary case $\rho \mapsto \varphi(U)\,\rho\,\varphi(U)^\dagger$ as well as situations involving losses in the linear optical network and aberrations due to mode matching issues. Gaussian completely positive maps are a very accurate modelling of present linear optical experiments.

Condition 3: Projection onto Fock states is replaced by measurements described by dichotomic POVMs $\{\Pi_0, \Pi_1\}$ with $\Pi_0 + \Pi_1 = \mathbb{I}$ ("bucket detector"), where the Wigner function corresponding to the no click event Π_0 for some fixed $R > 0$ is given by

$$W_{\Pi_0}(r) = \begin{cases} 1/(2\,\pi) & \text{if } |r| < R \\ 0 & \text{otherwise} \end{cases}. \tag{76}$$

This is an idealised model for imperfect photon detectors used in experiments that distinguishes the presence and the absence of photons, taking into account losses and dark counts. In the latter aspect the model considered here departs the furthest from actual experiments: While Condition 1 and 2 are usually satisfied to an extraordinarily large extent in quantum optical experiments, Condition 3 constitutes a rather crude approximation of an imperfect detector such as a realistic avalanche photodiode. Still, it is noteworthy that these conditions are sufficient to arrive at an efficient classical simulation. Needless to say, other detector models with positive Wigner functions for the POVM elements work equally well.

For a trace class operator A acting on a system with m modes, its Wigner function $W_A : \mathbb{R}^{2m} \to \mathbb{R}$ is defined as

$$W_A(r) \coloneqq \frac{1}{\pi^m} \operatorname{Tr}[w(r)\,\Pi^{\otimes m} w(r)^\dagger A]. \tag{77}$$

Here, Π is the single mode parity operator, $\{w(r)\}$ the family of Weyl operators, and $r \in \mathbb{R}^{2m}$ a vector collecting the $2m$ phase space variables. A state is Gaussian if and only if its Wigner function is Gaussian [26, 27]. Gaussian channels transform states with a Gaussian Wigner function into states with a Gaussian Wigner function. The Jamiolkowski isomorphs of such maps are Gaussian states.

Expressing Hilbert-Schmidt scalar products as integrals over Wigner functions, one finds for the dark count rate

$$\langle 0|\Pi_1|0\rangle = 1 - \langle 0|\Pi_0|0\rangle = 1 - 2\int_0^R r\,\mathrm{e}^{-r^2}\,\mathrm{d}r = \mathrm{e}^{-R^2}. \tag{78}$$

The Wigner function of the coherent state $|1\rangle_c$ that contains 1 photon on average is given by

$$W_{|1\rangle_c\langle 1|_c}(r) = \frac{1}{\pi}\mathrm{e}^{-(r_1-1)^2-r_2^2}. \tag{79}$$

With this one finds for the effective detector efficiency

$$\begin{aligned} &1 - \langle 1|_c\Pi_0|1\rangle_c = \\ &1 - \frac{1}{2\sqrt{\pi}}\int_{-R}^{R}\mathrm{d}p\,\mathrm{e}^{-p^2}\left(\mathrm{Erf}(1+(R^2-p^2)^{1/2}) - \mathrm{Erf}(1-(R^2-p^2)^{1/2})\right). \end{aligned} \tag{80}$$

For $R = 1.6$, say, one gets a reasonable dark count rate of $\langle 0|\Pi_1|0\rangle = 0.0773$ and $\langle 1|_c\Pi_0|1\rangle_c = 0.7104$, so an effective detector efficiency of 0.2896. These values are not that far off from those achieved in current experiments (see, e.g., Ref. [30, 34]).

In the setting considered here, the Wigner functions $W_{|0\rangle\langle 0|}$ and W_ω of the two single mode input states from which the initial state is constructed, that of the partial transposed of the Jamiolkowski isomorphs of all gates $W_{f_j^\Gamma},\ j \in [m^2]$, as well as that of the POVM elements W_{Π_0} and W_{Π_1} are non negative. Therefore, the algorithms of Refs. [31, 32] can be applied. The detailed error analysis of Refs. [32, 33] implies the following.

Observation 14 (Efficient sampling in 1-norm for imperfect detectors)**.** *For any number of modes m, any $R > 0$, any Gaussian product input state ρ, in which each mode is prepared in either the vacuum $|0\rangle\langle 0|$ or an arbitrary Gaussian state ω, any linear optical network, and any dichotomic detector with POVM elements Π_0 and Π_1 as defined in Eq.* (76)*, one can sample from the output distribution over $S \in \Phi_{m,n}$, to an error ϵ in 1-norm with effort $\mathrm{O}(\mathrm{poly}(m/\epsilon))$.*

That is to say, one can efficiently simulate the output distribution of imperfect linear optical networks and imperfect detectors of the above type even up to a small error in 1-norm. We suggest that it should be an important and constructive enterprise to exactly flesh out how far one can go with approximating realistic experimental devices, while still being able to provably efficiently simulate the output distribution.

8 Conclusion and outlook

In this work, we have revisited the Boson-Sampling problem from the perspective of sample complexity. We have arrived at the ironic conclusion that no symmetric probabilistic algorithm can distinguish the Boson-Sampling distribution from the mere uniform distribution on the collision-free subspace, unless exponentially many samples are available. The specifics of the problem if a priori knowledge is available have been discussed carefully. We have also addressed the question to what extend imperfect, approximate physical realisations of the Boson-Sampling problem can be classically efficiently simulated up to a constant error in 1-norm. As such, our work emphasizes the challenge of identifying ways to certify the correct working of such quantum simulators. Our results indicate that even though, unquestionably, the Boson-Sampling distribution has an intricate structure that makes sampling from it a classically hard problem, this structure seems inaccessible by classical means. To develop

a portfolio of methods for *certifying the correct functioning of quantum simulators* seems timelier than ever. Probably, quantum methods are indispensable to achieve that goal.

The question of the precise boundary of classically simulatable quantum processes remains wide open and interesting, and is also enjoying an increasing amount of attention [35, 36], not least because of the rather loud claims made in the context of the discussion on the functioning of the D-Wave processor and their careful assessment [37–39]. It is the hope that the present work can contribute to a thoughtful scientific reasoning on identifying the boundary of classically simulatable processes in general, and at the same time contribute to clarifying in what precise sense quantum devices such as Boson-Samplers are indeed more powerful than classical devices.

Obviously, technically, our argument leaves significant room for improvement. It would be interesting to see, for example, whether, or to what extent, a priori knowledge on the distribution can be used or what other important features of the Boson-Sampling distribution may be identified. It would also be important to see how the hardness argument can be partially *derandomised*, and the Haar-measure random unitaries replaced by appropriate *unitary designs* or related concepts derived from *quantum expanders*.

We also hope that our work can be read as yet another invitation to the enterprise of looking at the sample complexity of tasks in quantum theory. Quite generally, all information that is ever available in any quantum mechanical experiment is obtained from samples from a certain distribution. These samples may be used to infer about important features or properties of the underlying quantum state — or even about the very identity of the state in the first place. The quantum state tomography problem — the inference about an unknown quantum state from measurement data alone — should be phrased as a sampling problem. Indeed, the *tomography problem* has already been faithfully viewed as a sampling problem and the sample complexity lower bounded, both in the context of *quantum compressed sensing* [40] and in notions of *reliable quantum state tomography* [41]. A similar mindset has been taken in foundational arguments explaining the apparent emergence of ensembles of quantum statistical mechanics based on microscopic unitary evolution [42]: Indeed, one may argue that if by sampling alone, one cannot operationally distinguish a situation from the one predicted by a statistical ensemble, then the apparent emergence may be considered explained. It is the hope that the methods discussed in this work suggest further applications along these lines.

9 Acknowledgments

We warmly thank Fernando G. S. L. Brandao, Earl T. Campbell, and Rodrigo Gallego for insightful discussions and Scott Aaronson and Alex Arkhipov for useful criticism. We thank the EU (Q-Essence, REQS Marie Curie IEF No 299141), the ERC (TAQ), the EURYI, the BMBF (QuOReP), and the Studienstiftung des Deutschen Volkes for support.

References

[1] P. W. Shor, SIAM Journal of Computing **26**, 1484 (1997).

[2] J. Preskill, *Quantum supremacy now?*, blog entry on July 22, 2012, in Quantum Frontiers `http://quantumfrontiers.com/2012/07/22/supremacy-now/`.

[3] S. Aaronson and A. Arkhipov, Proceedings of ACM Symposium on the Theory of Computing, STOC, pp. 333-342 (Association for Computing Machinery, New York, 2011).

[4] M. J. Bremner, R. Jozsa, and D. J. Shepherd, Proc. R. Soc. A **467**, 459 (2011).

[5] J. B. Spring, B. J. Metcalf, P. C. Humphreys, W. S. Kolthammer, X.-M. Jin, M. Barbieri, A. Datta, N. Thomas-Peter, N. K. Langford, D. Kundys, J. C. Gates, B. J. Smith, P. G. R. Smith, and I. A. Walmsley, Science **339**, 798 (2013).

[6] M. Tillmann, B. Dakic, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, Nature Photonics, 10.1038/nphoton.2013.102 (2013).

[7] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White, Science **339**, 794 (2013).

[8] A. Crespi, R. Osellame, R. Ramponi, D. J. Brod, E. F. Galvao, N. Spagnolo, C. Vitelli, E. Maiorino, P. Mataloni, and F. Sciarrino, Nature Photonics, doi:10.1038/nphoton.2013.112 (2013).

[9] P. P. Rohde, Phys. Rev. A **86**, 052321 (2012).

[10] S. Scheel, quant-ph/0406127.

[11] M. C. Tichy, M. Tiersch, F. de Melo, F. Mintert, and A. Buchleitner, Phys. Rev. Lett. **104**, 220405 (2010).

[12] M. C. Tichy, M. Tiersch, F. Mintert, and A. Buchleitner, New J. Phys. **14** 093015 (2012).

[13] C. K. Hong, Z. Y. Ou, and L. Mandel, Phys. Rev. Lett. **59**, 2044 (1987).

[14] T. Batu, L. Fortnow, R. Rubinfeld, W. D. Smith, and P. White, Proceedings 41st Annual Symposium on Foundations of Computer Science, 259 (2000).

[15] H. Chernoff, Ann. Math. Stat. **23**, 493 (1952).

[16] R. E. Blahut, IEEE Trans. Inf. Theo. **20**, 405 (1974).

[17] C. W. Helström, *Quantum detection and estimation theory* (Academic Press, New York, 1976).

[18] A. S. Holevo, Theor. Prob. Appl. **23**, 411 (1978).

[19] K. M. R. Audenaert, J. Calsamiglia, L. Masanes, R. Munoz-Tapia, A. Acin, E. Bagan, and F. Verstraete, Phys. Rev. Lett. **98**, 160501 (2007).

[20] F. Hiai and D. Petz, Comm. Math. Phys. **143**, 99 (1991).

[21] K. M. R. Audenaert, M. Mosonyi, and F. Verstraete, J. Math. Phys. **53**, 122205 (2012).

[22] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Series in Telecommunication (John Wiley and Sons, New York, 1991).

[23] M. Ohya and D. Petz, *Quantum Entropy and its Use* (Springer, Heidelberg, 1993).

[24] N. Y. Ermolova and S.-G. Haggman, Proc. 12 European Signal Proces. Conf., Vienna, Austria, 1087 (2004).

[25] `http://mathoverflow.net/questions/45822/anti-concentration-bound-for-permanents-of-gaussian-matrices` (April 08, 2013).

[26] J. Eisert and M. B. Plenio, Int. J. Quant. Inf. **1**, 479 (2003).

[27] C. Weedbrook, S. Pirandola, R. Garcia-Patron, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Rev. Mod. Phys. **84**, 621 (2012).

[28] P. J. Mosley, J. S. Lundeen, B. J. Smith, P. Wasylczyk, A. B. U'Ren, C. Silberhorn, and I. A. Walmsley, Phys. Rev. Lett. **100**, 133601 (2008).

[29] J. Eisert and M. M. Wolf, *Gaussian quantum channels*, in *Quantum Information with Continous Variables of Atoms and Light*, pp. 23-42 (Imperial College Press, London, 2007).

[30] M. Stipčević, H. Skenderović, and D. Gracin, Opt. Exp. **18**, 16, 17448 (2010).

[31] A. Mari and J. Eisert, Phys. Rev. Lett. **109**, 230503 (2012).

[32] V. Veitch, N. Wiebe, C. Ferrie, and J. Emerson, New J. Phys. **15**, 013037 (2013).

[33] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, New J. Phys. **14**, 113011 (2012).

[34] J. S. Lundeen, A. Feito, H. Coldenstrodt-Ronge, K. L. Pregnell, C. Silberhorn, T. C. Ralph, J. Eisert, M. B. Plenio, and I. A. Walmsley, Nature Physics **5**, 27 (2009).

[35] M. J. Hoban, J. J. Wallman, H. Anwar, N. Usher, R. Raussendorf, D. E. Browne, arXiv:1304.2667 .

[36] N. Wiebe, C. Granade, C. Ferrie, D. G. Cory, arXiv:1309.0876.

[37] J. A. Smolin and G. Smith, arXiv:1305.4904.

[38] L. Wang, T. F. Roennow, S. Boixo, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer, arXiv:1305.5837.

[39] `http://www.scottaaronson.com/blog/?p=1400` (June 12, 2013).

[40] S. T. Flammia, D. Gross, Y.-K. Liu, and J. Eisert, New J. Phys. **14**, 095022 (2012).

[41] M. Christandl and R. Renner, Phys. Rev. Lett. **109**, 120403 (2012).

[42] C. Ududec, N. Wiebe, and J. Emerson, arXiv:1208.3419v3.

Reliable quantum certification for photonic quantum technologies

Leandro Aolita[1], Christian Gogolin[1,2,3], Martin Kliesch[1], and Jens Eisert[1]

[1]Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany
[2]ICFO-Institut de Ciències Fotòniques, Mediterranean Technology Park, 08860 Castelldefels (Barcelona), Spain
[3]Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, 85748 Garching, Germany

A major roadblock for large-scale photonic quantum technologies is the lack of practical reliable certification tools. We introduce an experimentally friendly — yet mathematically rigorous — certification test for experimental preparations of arbitrary m-mode pure Gaussian states, pure non-Gaussian states generated by linear-optical circuits with n-boson Fock-basis states as inputs, and states of these two classes subsequently post-selected with local measurements on ancillary modes. The protocol is efficient in m and the inverse post-selection success probability for all Gaussian states and all mentioned non-Gaussian states with constant n. We follow the mindset of an untrusted prover, who prepares the state, and a skeptic certifier, with classical computing and single-mode homodyne-detection capabilities only. No assumptions are made on the type of noise or capabilities of the prover. Our technique exploits an extremality-based fidelity bound whose estimation relies on non-Gaussian state nullifiers, which we introduce on the way as a byproduct result. The certification of many-mode photonic networks, as those used for photonic quantum simulations, boson samplers, and quantum metrology, is now within reach.

Many-body quantum devices promise exciting applications in ultra-precise quantum metrology[1], quantum computing[2–4], and quantum simulators[5–9]. In the quest for their large-scale realisation, impressive progress on a variety of quantum technologies has recently been made[6–9]. Among them, optical implementations play a key role. For example, sophisticated manipulations of multi-qubit entangled states of up to eight parametrically down-converted photons[10,11] have been demonstrated and continuous-variable entanglement among 60 stable[12] and up to 10000 flying[13] modes has been verified in optical set-ups. In addition, small-sized simulations of BosonSampling[14–17] and Anderson localisation in quantum walks[18,19] have been performed with on-chip integrated linear-optical networks.

This fast pace of advance, however, makes the problem of *reliable certification* an increasingly pressing issue[20–24]. From a practical viewpoint, further experimental progress on many-body quantum technologies is nowadays hindered by the lack of practical certification tools. At a fundamental level, certifying many-body quantum devices is ultimately about testing quantum mechanics in regimes where it has never been tested before.

Tomographic characterisation of quantum states requires the measurement of exponentially many observables. Compressed-sensing techniques[25] reduce, for states approximated by low-rank density matrices, the requirements significantly, but still demand exponentially many measurements. Efficient certification techniques, requiring only polynomially many measurements, for universal quantum computation[26,27] and a restricted model of computation with one pure qubit[28] exist in the form of quantum interactive proofs. However, these require either a fully fledged fault-tolerant universal quantum computer[26,27] or an experimentally non-trivial measurement-based quantum device[28]. In addition, these methods involve sequential interaction rounds with the device[26–28]. In contrast, permutationally invariant tomography[29], Monte-Carlo fidelity estimation[30–32], and Clifford-circuit benchmarking techniques[40] provide experimentally friendly alternatives for the efficient certification of preparations of permutationally invariant[29] and qubit stabiliser or W states[30–32,40], respectively. Nevertheless, none of these methods addresses continuous-variable systems, not even in Gaussian states.

Here, we introduce an experimentally friendly technique for the direct certification of continuous-variable state preparations without estimating the prepared state itself. First, we discuss intuitively and define rigorously reliable quantum-state certification tests. We do this for two notions of certification, differing in that in one of them robustness against preparation errors is mandatory. Then, we present a certification test, based on single-mode homodyne detection, for arbitrary m-mode pure Gaussian states, non-Gaussian states resulting from Gaussian unitary operators acting on Fock-basis states with n photons, and states prepared by post-selecting states in either of the two classes with measurements on $a < m$ ancillary modes in arbitrary local bases. This covers, for instance, Gaussian quantum simulations such as those of refs. 12 and 13 as well as the non-Gaussian ones of refs. 6, 10, 11, 14–19. Furthermore, so-called de-Gaussified (photon-subtracted) Gaussian states[41–44] as well as all non-Gaussian states accessible to qumode-encoded qubit[45,46] or finite-squeezing qumode[47,48] quantum computers also lie within the range of applicability of our method. The protocol is efficient in m and, for the cases with post-selection, in the inverse polynomial post-selection success probability, for all Gaussian states and all mentioned non-Gaussian states with constant n.

With a high probability, our test rejects all experimental preparations with a fidelity with respect to the chosen target

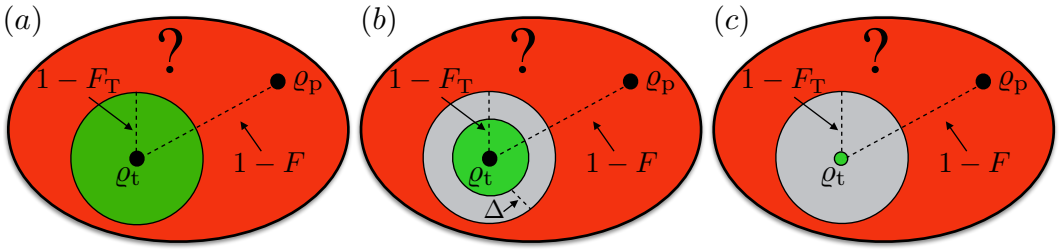

Figure 1. **Different certification paradigms**. (*a*) Naive approach: To certify an untrusted experimental preparation ϱ_{p} of the target state ϱ_{t}, a certifier Arthur would like to run a statistical test that, for all ϱ_{p}, decides whether the fidelity F between ϱ_{p} and ϱ_{t} is greater or equal than a pre-specified threshold $F_{\mathrm{T}} < 1$ (green region, accept), or smaller than it (red region, reject). However, due to the preparations at the boundary of the two regions and experimental uncertainties, a test able to make such a decision does not exist. (*b*) The ideal scenario: A more realistic certification notion is to ask that the test rejects every ϱ_{p} for which $F < F_{\mathrm{T}}$ (red region) and accepts every ϱ_{p} for which $F \geq F_{\mathrm{T}} + \Delta$ (green region), for some given $\Delta < 1 - F_{\mathrm{T}}$. Here, a buffer region of width Δ (in grey) is introduced within which the behaviour of the test can be arbitrary, but, in return, the certification is now feasible. This type of certification is thus robust against experimental infidelities as large as $1 - F_{\mathrm{T}} - \Delta$. (*c*) The practical scenario: Finally, the least one can demand is that the test rejects every ϱ_{p} for which $F < F_{\mathrm{T}}$ (red region) and accepts at least ϱ_{t} (green point). The former condition is sometimes called *soundness* and the latter one *completeness*. Here, no acceptance is guaranteed for any ϱ_{p} with $F \geq F_{\mathrm{T}}$ (grey region) other than ϱ_{t} itself, but any ϱ_{p} accepted by the test necessarily features $F \geq F_{\mathrm{T}}$. This certification notion is not robust against state deviations, but it can be more practical. In addition, in practice, the resulting tests succeed also in accepting many $\varrho_{\mathrm{p}} \neq \varrho_{\mathrm{t}}$ for which $F \geq F_{\mathrm{T}}$.

state lower than a desired threshold and accepts if the preparation is sufficiently close to the target. That is, the protocol is robust against small preparation errors. We upper-bound the failure probability in terms of the number of experimental runs and calculate the necessary number of measurement settings. Our method is built upon a fidelity lower bound, based on a natural extremality property, that is interesting in its own right. Finally, the experimental estimation of this bound relies on non-Gaussian state nullifiers, which we introduce on the way.

Results

We present our results in terms of photons propagating through optical networks, but our methods apply to any bosonic platform with equivalent dynamics. We consider a sceptic certifier, Arthur, with limited quantum capabilities, who wishes to ascertain whether an untrusted quantum prover, Merlin, presumably with more quantum capabilities, can indeed prepare certain quantum states that Arthur cannot. This mindset is reminiscent to that of quantum interactive-proof systems[26–28] of computer science, but our method has the advantage that no interaction apart from the measurements of the certifier on the single-run experimental preparations from the prover is required.

In particular, we consider the situation where Merlin possesses at least a network of active single-mode squeezers and displacers as well as passive beam-splitters and phase-shifters, sufficient to efficiently implement any m-mode Gaussian unitary[33–36], plus single-photon sources. Arthur's resources, in contrast, are restricted to classical computational power augmented with single-mode measurements. With that, he can characterise each of his single-mode measurement channels up to any desired constant precision. The task is for Merlin to provide him with copies of an m-mode pure *target state* ϱ_{t} of Arthur's choice. We assume that Merlin follows independent and identical state-preparation procedures on each experimental run, described by the density matrix ϱ_{p}. We refer to ϱ_{p} as a *preparation* of the target state ϱ_{t}. His preparation is unavoidably subject to imperfections and he might even be dishonest and try to trick Arthur. Thus, Arthur would like to run a test, with his own measurement devices, to *certify* whether ϱ_{p} is indeed a bona fide preparation of ϱ_{t}.

To measure how good a preparation ϱ_{p} of ϱ_{t} is, we use the fidelity between ϱ_{p} and ϱ_{t}, defined as

$$F := F(\varrho_{\mathrm{t}}, \varrho_{\mathrm{p}}) := \mathrm{Tr}\left[(\sqrt{\varrho_{\mathrm{t}}}\varrho_{\mathrm{p}}^{\dagger}\sqrt{\varrho_{\mathrm{t}}})^{1/2}\right]^2 = \mathrm{Tr}\left[\varrho_{\mathrm{t}}\varrho_{\mathrm{p}}\right], \quad (1)$$

where the last equality holds because ϱ_{t} is assumed to be pure. As we see below, our measurement schemes directly estimate fidelities. However, all our results can also be adapted to the trace distance $D := D(\varrho_{\mathrm{t}}, \varrho_{\mathrm{p}})$, which can be defined via the 1-norm distance in state space as $D(\varrho_{\mathrm{t}}, \varrho_{\mathrm{p}}) := \mathrm{Tr}[|\varrho_{\mathrm{t}} - \varrho_{\mathrm{p}}|]/2$. This is due to the fact that D can be bounded from both sides in terms of F through the well-known inequalities $1 - F^2 \leq D \leq \sqrt{1 - F^2}$, where the first inequality holds because ϱ_{t} is pure.

Let us first discuss what properties an experimental test must fulfil to qualify as a state certification protocol. Different certification paradigms are schematically represented in Fig. 1. We start with the formal definition of certification in the sense of Fig. 1 (c).

Definition 1 (Quantum state certification)**.** *Let $F_{\mathrm{T}} < 1$ be a threshold fidelity and $\alpha > 0$ a maximal failure probability. A test, which takes as input a classical description of ϱ_{t} and copies of a preparation ϱ_{p} and outputs "accept" or "reject" is a* certification test *for ϱ_{t} if, with probability at least $1-\alpha$, it both rejects every ϱ_{p} for which $F(\varrho_{\mathrm{t}}, \varrho_{\mathrm{p}}) < F_{\mathrm{T}}$ and accepts $\varrho_{\mathrm{p}} = \varrho_{\mathrm{t}}$. We say that any ϱ_{p} accepted by such a test is a* certified preparation *of ϱ_{t}.*

To specify the target states we need to introduce some no-

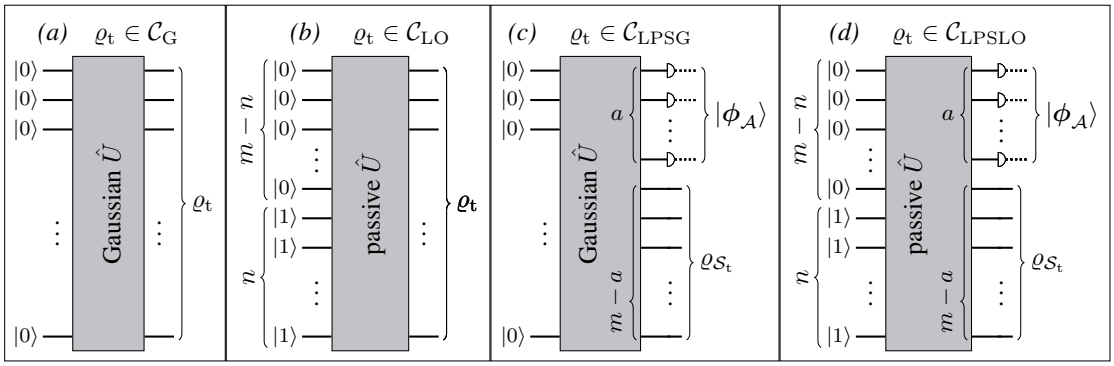

Figure 2. **Classes of target states**. (*a*) $\mathcal{C}_{\rm G}$ is the class composed of all m-mode pure Gaussian states. These can be prepared by applying an arbitrary Gaussian unitary $\hat{U}$ (possibly involving multi-mode squeezing) to the m-mode vacuum state $|\mathbf{0}\rangle$. (*b*) The class $\mathcal{C}_{\rm LO}$ includes all m-mode pure non-Gaussian states produced at the output of an arbitrary linear-optical network, which implements a passive Gaussian unitary $\hat{U}$ (without squeezing), with the Fock-basis state $|\mathbf{1}_n\rangle$ containing one photon in each of the first n modes and zero in the remaining $m-n$ ones as input. As the order of the modes is arbitrary, choosing the first n modes as the populated ones does not constitute a restriction. (*c*) The third class, $\mathcal{C}_{\rm LPSG}$, encompasses all $(m-a)$-mode pure non-Gaussian states obtained by projecting a subset $\mathcal{A}$ of $a<m$ modes of an m-mode pure Gaussian state $\varrho_{\rm t}\in\mathcal{C}_{\rm G}$ onto an arbitrary pure tensor-product state $|\boldsymbol{\phi}\rangle_{\mathcal{A}}$. In practice, this is done probabilistically by measuring $\mathcal{A}$ in a local basis that contains $|\boldsymbol{\phi}\rangle_{\mathcal{A}}$ and post-selecting only the events in which $|\boldsymbol{\phi}\rangle_{\mathcal{A}}$ is measured. Thus, the a modes in $\mathcal{A}$ are used as ancilas, whereas the effective system is given by the subset $\mathcal{S}$ containing the other $m-a$ modes, which carries the final target state. For concreteness, but without any loss of generality, in the plot, the ancillary modes are chosen to be the last a ones. (*d*) Analogously, the class $\mathcal{C}_{\rm LPSLO}$ is that of all $(m-a)$-mode pure non-Gaussian states obtained by projecting the ancillary modes of an m-mode pure linear-optical network state $\varrho_{\rm t}\in\mathcal{C}_{\rm LO}$ onto an arbitrary pure tensor-product state $|\boldsymbol{\phi}\rangle_{\mathcal{A}}$. These four classes cover the target states considered in the vast majority of quantum photonic experiments.

tation. We denote m-mode *Fock basis states* by $|\mathbf{n}\rangle$, with $\mathbf{n}\coloneqq(n_1,n_2,\ldots,n_m)$ being the sequence of photon numbers $n_j\geq 0$ in each mode $j\in[m]$, where the short-hand notation $[m]\coloneqq\{1,2,\ldots,m\}$ is introduced, and call $n\coloneqq\sum_{j=1}^m n_j$ the *total input photon number*. In particular, we will pay special attention to Fock basis states $|\mathbf{1}_n\rangle$ with exactly one photon in each of the first n modes and the vacuum in the remaining $m-n$ ones, i.e., those for which $\mathbf{n}=\mathbf{1}_n$, with

$$\mathbf{1}_n\coloneqq(\underbrace{1,\ldots,1}_{n\text{ times}},\underbrace{0,\ldots,0}_{m-n\text{ times}}). \tag{2}$$

Note that $|\mathbf{1}_0\rangle$ is the Gaussian *vacuum state* $|\mathbf{0}\rangle$. We denote the *photon number operator* corresponding to mode j by $\hat{n}_j$ and the *total photon number operator* by $\hat{n}\coloneqq\sum_{j=1}^m\hat{n}_j$.

In addition, for post-selected target states, we denote by $\mathcal{A}\coloneqq\{\mathcal{A}_j\}_{j\in[a]}$, where each element $\mathcal{A}_j\in[m]$ labels a different mode, the subset of $a\coloneqq|\mathcal{A}|<m$ modes on which the post-selection measurements are made. We then identify the remaining $m-a$ modes as the system subset $\mathcal{S}$, which carries the post-selected target state $\varrho_{\mathcal{S}_{\rm t}}$. The subindex $\mathcal{S}$ emphasises that $\varrho_{\mathcal{S}_{\rm t}}$ represents an $(m-a)$-mode post-selected target state and distinguishes it from m mode target states without post-selection, which we denote simply as $\varrho_{\rm t}$. We denote by $|\boldsymbol{\phi}\rangle_{\mathcal{A}}\coloneqq|\phi_1\rangle_{\mathcal{A}_1}|\phi_2\rangle_{\mathcal{A}_2}\ldots|\phi_a\rangle_{\mathcal{A}_a}$, with $\{|\phi_j\rangle_{\mathcal{A}_j}\}_{j\in a}$ an arbitrary pure normalised state of mode $\mathcal{A}_j$, an a-mode product state on the modes $\mathcal{A}$. We use the short-hand notations $\langle\boldsymbol{\phi}|_{\mathcal{A}}\,\varrho_{\rm t}\,|\boldsymbol{\phi}\rangle_{\mathcal{A}}\coloneqq\mathrm{Tr}_{\mathcal{A}}\left[\varrho_{\rm t}(\mathbb{1}_{\mathcal{S}}\otimes|\boldsymbol{\phi}\rangle_{\mathcal{A}}\langle\boldsymbol{\phi}|_{\mathcal{A}})\right]$, where $\mathrm{Tr}_{\mathcal{A}}$ indicates partial trace over the Fock space of $\mathcal{A}$, $\mathbb{1}_{\mathcal{S}}$ denotes the identity on $\mathcal{S}$, and $\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\rm t})\coloneqq\mathrm{Tr}\left[\langle\boldsymbol{\phi}|_{\mathcal{A}}\,\varrho_{\rm t}\,|\boldsymbol{\phi}\rangle_{\mathcal{A}}\right]$ is the *post-selection success probability*, i.e., the probability of measuring $|\boldsymbol{\phi}\rangle_{\mathcal{A}}$ in a projective measurement on $\mathcal{A}$. Without loss of generality, we consider throughout only the non-trivial case $\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\rm t})\neq 0$. Thus, we consider exclusively post-selected target states of the form

$$\varrho_{\mathcal{S}_{\rm t}}\coloneqq\frac{\langle\boldsymbol{\phi}|_{\mathcal{A}}\,\varrho_{\rm t}\,|\boldsymbol{\phi}\rangle_{\mathcal{A}}}{\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\rm t})}. \tag{3}$$

With the notation introduced, we derive our results for:

1) Arbitrary m-mode pure *Gaussian states*, given by the class

$$\mathcal{C}_{\rm G}\coloneqq\{\varrho_{\rm t}=\hat{U}\,|\mathbf{0}\rangle\langle\mathbf{0}|\,\hat{U}^\dagger\colon\hat{U}\text{ Gaussian unitary}\}, \tag{4}$$

2) m-mode pure *linear-optical network states* from the class

$$\mathcal{C}_{\rm LO}\coloneqq\{\varrho_{\rm t}=\hat{U}\,|\mathbf{1}_n\rangle\langle\mathbf{1}_n|\,\hat{U}^\dagger\colon\hat{U}\text{ passive unitary}\}, \tag{5}$$

3) arbitrary $(m-a)$-mode pure *locally post-selected Gaussian states*, given by the class

$$\mathcal{C}_{\rm LPSG}\coloneqq\{\varrho_{\mathcal{S}_{\rm t}}\colon\varrho_{\rm t}\in\mathcal{C}_{\rm G}\}, \tag{6}$$

4) and $(m-a)$-mode pure *locally post-selected linear-optical network states* from the class

$$\mathcal{C}_{\rm LPSLO}\coloneqq\{\varrho_{\mathcal{S}_{\rm t}}\colon\varrho_{\rm t}\in\mathcal{C}_{\rm LO}\}. \tag{7}$$

The class $\mathcal{C}_{\rm G}$ is crucial within the realm of "continuous-variable" quantum optics and quantum information processing. It encompasses, for instance, "twin-beam" (two-mode squeezed vacuum) states under passive networks, which are used to simulate, upon coincidence detection, multi-qubit

states[6]. The class $\mathcal{C}_{\rm LO}$ includes all the settings sometimes referred to as "discrete variable" linear-optical networks. This class covers, among others, the targets of several recent experimental simulations with on-chip integrated linear-optical networks[14–19]. The third class, $\mathcal{C}_{\rm LPSG}$, is the one of locally post-selected Gaussian states. This class includes crucial non-Gaussian resources for quantum information and quantum optics. For instance, when the post-selection is in the Fock basis, it encompasses de-Gaussified photon-subtracted squeezed Gaussian states[41–44]. Furthermore, if apart from Fock-basis measurements, the post-selection uses also quadrature homodyne measurements, $\mathcal{C}_{\rm LPSG}$ contains all the states accessible to finite-squeezing cluster-state qumode quantum computers[47,48]. The last class, $\mathcal{C}_{\rm LPSLO}$, of locally post-selected linear-optical network states, covers, for the case where the post-selection is in the Fock basis and n is proportional to m, all the states prepared by probabilistic schemes of the type of refs. 45 and 46 for universal qumode-encoded qubit quantum computation. Naturally, $\mathcal{C}_{\rm LPSLO}$ also includes both photon -added or -subtracted linear-optical network states.

The basis of the our certification scheme is a technique for the estimation of the quantity

$$F^{(n)} := 1 - \left\langle (\hat{n} - n) \prod_{j=1}^{n} \hat{n}_j \right\rangle_{\hat{U}^\dagger \varrho_{\rm p} \hat{U}}, \tag{8}$$

with n the total input photon number. As shown in the Methods section, for all target states $\varrho_{\rm t} \in \mathcal{C}_{\rm G} \cup \mathcal{C}_{\rm LO}$, $F^{(n)}$ is a lower bound on the fidelity F and, moreover, $F^{(n)} = F = 1$ if $\varrho_{\rm p} = \varrho_{\rm t}$ (see also Methods and Section S2.A of the SI for analogous bounds for the post-selected target states). This bound is a consequence of a natural extremality notion: the smaller the expectation value $\langle (\hat{n} - n) \prod_{j=1}^{n} \hat{n}_j \rangle_{\hat{U}^\dagger \varrho_{\rm p} \hat{U}}$ is, the closer are $|\mathbf{1}_n\rangle\langle\mathbf{1}_n|$ and $\hat{U}^\dagger \varrho_{\rm p} \hat{U}$ and, therefore, the closer are the preparation $\varrho_{\rm p}$ and the target state $\varrho_{\rm t}$. Our test $\mathcal{T}$, summarised in Box 1, yields an estimate $F^{(n)*}$ of $F^{(n)}$. If $F^{(n)*}$ is sufficiently above the threshold $F_{\rm T}$, the preparation $\varrho_{\rm p}$ is accepted. Otherwise it is rejected. The estimate $F^{(n)*}$ is obtained via a measurement scheme that depends on the specific target state. In the Gaussian case $n = 0$ the measurement scheme $\mathcal{M}_{\rm G}$ can be used, while linear-optical network states with $n > 0$ require the scheme $\mathcal{M}_{\rm LO}$. $\mathcal{M}_{\rm G}$ and $\mathcal{M}_{\rm LO}$ are both summarised in the Methods section and described in detail in Boxes S1 and S2, respectively, in Section S2 in the Supplementary Information (SI). In addition, in Section S2.B of the SI we adapt $\mathcal{T}$ to post-selected target states $\varrho_{\mathcal{S}_{\rm t}} \in \mathcal{C}_{\rm LPSG} \cup \mathcal{C}_{\rm LPSLO}$, and provide the corresponding adapted measurement schemes in Section S2.C of the SI.

Our theorems guarantee that the test from Box 1 is indeed a certification test and give a bound on the scaling of the number of samples that are needed for the test. In order to state them we introduce some notation related to mode space descriptions of linear-optical networks first. Any Gaussian unitary transformation $\hat{U}$ on Hilbert space can be represented by an affine symplectic transformation in mode space, i.e., by a *symplectic matrix* $\mathbf{S} \in {\rm Sp}(2m, \mathbb{R})$ followed by a phase-space *displacement* $\mathbf{x} \in \mathbb{R}^{2m}$ (see equation (26) in the Methods section), where the real symplectic group ${\rm Sp}(2m, \mathbb{R})$ contains all real $2m \times 2m$ matrices that preserve the canonical phase-space commutation relations[33,34]. By virtue of the Euler decomposition[33,35], $\mathbf{S}$ can be implemented with single-mode squeezing operations and passive mode transformations. We denote the *maximum single-mode squeezing* of $\mathbf{S}$ by $s_{\max}$ and define the *mode range* $d \leq m$ to be the maximal number of input modes to which each output mode is coupled (for details see Section S1 of the SI). Also, it will be useful to define

$$\kappa := 2 \min\{d^2, m\}. \tag{9}$$

The displacement $\mathbf{x}$ can be implemented by a single-mode displacer at each mode $j \in [m]$, with amplitude (x_{2j-1}, x_{2j}), where x_k, for $k \in [2m]$, is the k-th component of $\mathbf{x}$. The *vector 2-norm* is denoted by $\|\cdot\|_2$, i.e., $\|\mathbf{x}\|_2 := \left(\sum_{k=1}^{2m} x_k^2\right)^{1/2}$.

We take σ_i to be a uniform upper bound on the variances of any product of i *phase space quadratures* in the state $\varrho_{\rm p}$. If $\varrho_{\rm p}$ is Gaussian, σ_1 and σ_2 are functions of the single mode squeezing parameters of $\varrho_{\rm p}$. In addition, we call $\sigma_{\leq i} := \max_{k \leq i}\{\sigma_k\}$ the *maximal i-th variance* of $\varrho_{\rm p}$. Finally, we use the Landau symbol O to denote asymptotic upper bounds.

Box 1 (Certification test $\mathcal{T}$).
1) Arthur chooses a threshold fidelity $F_{\rm T} < 1$, a maximal failure probability $\alpha > 0$, and an estimation error $0 < \varepsilon \leq (1 - F_{\rm T})/2$.
2) Arthur provides Merlin with the classical specification n, $\mathbf{S}$, and $\mathbf{x}$ of the target state $\varrho_{\rm t}$ and requests a sufficient number of copies of it.
3) If $n = 0$, Arthur measures $2m\kappa$ two-mode correlations and $2m$ single-mode expectation values specified by the measurement scheme $\mathcal{M}_{\rm G}$ (see the Methods section), which can be done with $m + 3$ single-mode homodyne settings.
If $n > 0$, he measures $\mathrm{O}\left(m(4d^2 + 1)^n\right)$ multi-body correlators, each one involving between 1 and $2n + 1$ modes, specified by the measurement scheme $\mathcal{M}_{\rm LO}$ (see the Methods section), which can be done with at most $\binom{m}{n} 2^{n+1}$ single-mode homodyne settings.
4) By classical post-processing (see the Methods section), he obtains a fidelity estimate $F^{(n)*}$ such that $F^{(n)*} \in [F^{(n)} - \varepsilon, F^{(n)} + \varepsilon]$ with probability at least $1 - \alpha$, where $F^{(n)}$ is the lower bound to F given by expression (8).
5) If $F^{(n)*} < F_{\rm T} + \varepsilon$, he rejects. Otherwise, he accepts.

Theorem 2 (Quantum certification of Gaussian states)**.** *Let $F_{\rm T} < 1$ be a threshold fidelity, $\alpha > 0$ a maximal failure probability, and $0 < \varepsilon \leq (1 - F_{\rm T})/2$ an estimation error. Let $\varrho_{\rm t} \in \mathcal{C}_{\rm G}$ have maximum single-mode squeezing $s_{\max} \geq 1$, mode range $d \leq m$, and displacement $\mathbf{x}$. Test $\mathcal{T}$ from Box 1 is a certification test for $\varrho_{\rm t}$ and requires at most*

$$\mathrm{O}\left(\frac{s_{\max}^4 \left(2\sigma_1^2 \|\mathbf{x}\|_2^2 m^3 + \sigma_2^2 \kappa^3 m^4\right)}{\varepsilon^2 \ln(1/(1-\alpha))} \right) \tag{10}$$

copies of a preparation $\varrho_{\rm p}$ with first and second variance bounds $\sigma_1 > 0$ and $\sigma_2 > 0$, respectively.

Theorem 3 (Quantum certification of linear-optical network states)**.** *Let $F_{\rm T} < 1$ be a threshold fidelity, $\alpha > 0$ a maximal failure probability, and $0 < \varepsilon \leq (1 - F_{\rm T})/2$ an estimation error. Let $\varrho_{\rm t} \in \mathcal{C}_{\rm LO}$ have mode range $d \leq m$. Test $\mathcal{T}$ from Box 1 is a certification test for $\varrho_{\rm t}$ and requires at most*

$$\mathrm{O}\left(\frac{\sigma^2_{\leq 2(n+1)} m^4 (\lambda\, d^6\, n\, m)^n}{\varepsilon^2 \ln(1/(1-\alpha))}\right) \tag{11}$$

copies of a preparation $\varrho_{\rm p}$ with maximal $2(n+1)$-th variance $\sigma_{\leq 2(n+1)}$, where $\lambda > 0$ is an absolute constant.

The proofs of all our theorems are provided in the SI. The treatments of the classes $\mathcal{C}_{\rm LPSG}$ or $\mathcal{C}_{\rm LPSLO}$ follow as corollaries of Theorems 2 and 3, respectively, and are also provided in the SI (see Section S2.D there). Expressions (10) and (11) are highly simplified upper bounds on the total number of copies of $\varrho_{\rm p}$ that $\mathcal{T}$ requires. For more precise expressions see equations (S55) and (S77) of the SI. Note that neither of the two theorems requires any energy cut-off or phase-space truncation. While our bound in equation (11) is inefficient in n, both for the Gaussian and linear-optical cases, the number of copies of $\varrho_{\rm p}$ scales polynomially with all other parameters, in particular with m. Thus, arbitrary m-mode target states from the classes $\mathcal{C}_{\rm G}$ and $\mathcal{C}_{\rm LO}$ with constant n, are certified by $\mathcal{T}$ efficiently.

Interestingly, since states in $\mathcal{C}_{\rm LO}$ in general display negative Wigner functions, sampling from their measurement probability distributions cannot be efficiently done by the available classical sampling methods[37,38]. Furthermore, for Fock-state measurements, these distributions define BosonSampling, for which hardness results exist[39] for m asymptotically lower-bounded by n^5.

Also, note that there are no restrictions on $\varrho_{\rm p}$ except that, in practice, to apply the theorems, one needs bounds on σ_1, σ_2, and $\sigma_{\leq 2(n+1)}$. These variances are properties of $\varrho_{\rm p}$ and are therefore a priori unknown to Arthur. However, he can reasonably estimate them from his measurements. Note that, for random variables that can take any real value, assuming that the variances are bounded is a fundamental and unavoidable assumption to make estimations from samples; and it is one that can be contrasted with the measurement results.

To end up with, we consider certification in the robust sense of Fig. 1 (b):

Definition 4 (Robust quantum state certification)**.** *Let $F_{\rm T} < 1$ be a threshold fidelity, $\alpha > 0$ a maximal failure probability, and $\Delta < 1 - F_{\rm T}$ a fidelity gap. A test, which takes as input a classical description of the target state $\varrho_{\rm t}$ and copies of a preparation $\varrho_{\rm p}$ and outputs "accept" or "reject" is a* robust certification test *for $\varrho_{\rm t}$ if, with probability at least $1 - \alpha$, it both rejects every $\varrho_{\rm p}$ for which $F(\varrho_{\rm t}, \varrho_{\rm p}) < F_{\rm T}$ and accepts every $\varrho_{\rm p}$ for which $F(\varrho_{\rm t}, \varrho_{\rm p}) \geq F_{\rm T} + \Delta$. We say that any $\varrho_{\rm p}$ accepted by such a test is a* certified preparation *of $\varrho_{\rm t}$.*

This definition is more stringent than Definition 1 in that it guarantees that preparations sufficiently close to $\varrho_{\rm t}$ are necessarily accepted, rendering the certification robust against state deviations with infidelities as large as $1-(F_{\rm T}+\Delta)$. We show below that our test $\mathcal{T}$ from Box 1 is actually a robust certification test.

To this end, we first write $\varrho_{\rm p}$ as

$$\varrho_{\rm p} = F\varrho_{\rm t} + (1-F)\varrho_{\rm t}^{\perp}, \tag{12}$$

where $\varrho_{\rm t}^{\perp}$ is an operator orthogonal to $\varrho_{\rm t}$ with respect to the Hilbert-Schmidt inner product, i.e., such that $\mathrm{Tr}[\varrho_{\rm t}\,\varrho_{\rm t}^{\perp}] = 0$. As $\varrho_{\rm t}$ is assumed to be pure, it follows immediately that $\varrho_{\rm t}^{\perp}$ is actually a state. In fact, multiplying by $\varrho_{\rm t}$ and taking the trace on both sides of equation (12), one readily sees that the decomposition (12) is just another way to express the fidelity (1). We define the *photon mismatch* $\tilde{n}^{\perp}$ between $\varrho_{\rm t}$ and $\varrho_{\rm p}$ as

$$\tilde{n}^{\perp} := \langle (\hat{n} - n) \prod_{j=1}^{n} \hat{n}_j \rangle_{\hat{U}^{\dagger}\varrho_{\rm t}^{\perp}\hat{U}}. \tag{13}$$

The photon mismatch gives the expectation value that Arthur would obtain if he had access to $\varrho_{\rm t}^{\perp}$, applied the inverse of Merlin's network to it, and then measured $(\hat{n} - n)\prod_{j=1}^{n}\hat{n}_j$. For the ideal case $\varrho_{\rm p} = \varrho_{\rm t}$, it clearly holds that $\tilde{n}^{\perp} = 0$.

Theorem 5 (Robust quantum certification)**.** *Under the same conditions as in Theorems 2 and 3, test $\mathcal{T}$ from Box 1 is a robust certification test with fidelity gap*

$$\Delta := \max\left\{\frac{2\varepsilon + (1-F_{\rm T})(\tilde{n}^{\perp} - 1)}{\tilde{n}^{\perp}}, 2\varepsilon\right\}, \tag{14}$$

where $\tilde{n}^{\perp}$ is the photon mismatch.

As expected, the gap cannot be smaller than twice the estimation error for any photon mismatch. Notice also that in the limit $\tilde{n}^{\perp} \to \infty$ it holds that $\Delta \to 1 - F_{\rm T}$, so that the certification becomes less robust with increasing $\tilde{n}^{\perp}$. As $\tilde{n}^{\perp}$ decreases from infinity to one, the gap decreases to its minimal value $\Delta = 2\varepsilon$, where it remains for all $0 \leq \tilde{n}^{\perp} \leq 1$. We emphasise that $\tilde{n}^{\perp}$ depends on $\varrho_{\rm t}^{\perp}$. Thus it cannot be directly estimated from measurements on $\varrho_{\rm p}$ alone. However, for any $\tilde{n}^{\perp} < \infty$, Theorem 5 guarantees the existence of an entire region of states around $\varrho_{\rm t}$ that are rightfully accepted. Furthermore, in the experimentally relevant situations, $\tilde{n}^{\perp}$ is expected to be small. In this case, Theorem 5 provides a lower bound on the size of the region of accepted states.

Finally, a statement equivalent to Theorem 5 for target states $\varrho_{\mathcal{S}_{\rm t}} \in \mathcal{C}_{\rm LPSG} \cup \mathcal{C}_{\rm LPSLO}$ follows as an immediate corollary of it and is presented in Section S2.E in the SI.

Discussion

Large-scale photonic quantum technologies promise important scientific advances and technological applications. So far, considerably more effort has been put into their realisation than into the verification of their correct functioning and reliability. This imposes a serious obstacle for further experimental advance, specifically in the light of the speed at which progress towards many-mode architectures takes place. Here,

we have presented a practical reliable certification tool for a broad family of multi-mode bosonic quantum technologies.

We have proven theorems that upper-bound the number of experimental runs sufficient for our protocol to be a certification test. Our theorems provide large-deviation bounds from a simple extremality-based fidelity lower-bound that is interesting in its own right. Importantly, our theorems hold only for statistical errors, but the stability analyses on which they rely (see Lemmas S6 and S9 in the SI) holds regardless of the nature of the errors. As a matter of fact, in Section S5 in the SI, we show that our fidelity estimates are robust also against systematic errors.

From a more practical viewpoint, our test allows one to certify the state preparations of most current optical experiments, in both the "continuous-variable" and the "discrete-variable" setting. This is achieved under the minimal possible assumptions: namely, only that the variances of the measurement outcomes are finite. Thus, the certification is as unconditional as the fundamental laws of statistics allow. In particular, no assumption on the type of quantum noise is made. Despite the rigorous bounds on the estimation errors and failure probabilities, our methods are both experimentally friendly and resource efficient.

Notably, our test can for instance be applied to the certification of optical circuits of the type used in BosonSampling: There, m-mode Fock-basis states of n photons are subjected to a linear-optical network described by a random unitary $\hat{U}$ drawn from the Haar measure[39] and, subsequently, each output mode is measured in the Fock basis. While the question of the certification of the classical outcomes of such samplers without assumptions on the device is still largely open[20,21], with the methods described here the pre-measurement non-Gaussian quantum outputs of BosonSampling devices[14–17] can be certified reliably and, for constant n, even efficiently. In this sense, this work goes significantly beyond previously proposed schemes to rule out particular cheating strategies by the prover[21–24]. Furthermore, a variety of non-Gaussian states paradigmatic in quantum optics and quantum information are also covered by our protocol (see Section S2 in the SI for details). These include, for instance, de-Gaussified photon-subtracted multi-mode Gaussian states[41–44], multi-mode squeezed Gaussian states post-selected through photon-number or quadrature measurements, as in finite-squeezing cluster-state qumode quantum computers[47,48], and linear-optical network outputs post-selected though photon-number measurements, ranging from photon -added or -subtracted linear-optical network states to all the states preparable with Knill-Laflamme-Milburn-like schemes[45,46]. For all such states, our test is efficient in the inverse post-selection success probability $1/\mathbb{P}(\phi_{\mathcal{A}}|\varrho_{\mathrm{t}})$.

The present method constitutes a step forward in the field of photonic quantum certification, with potential implications on the certification of other many-body quantum-information technologies. Apart from that of BosonSamplers and optical schemes with post-selection, the efficient and reliable certification of large-scale photonic networks as those used, for instance, for multi-mode Gaussian quantum-information processing[12,13], non-Gaussian Anderson-localisation simulations[18,19], and quantum metrology[1], with a constant number of input photons, is now within reach.

Methods

Fidelity lower bound. In this section, we formalise the extremality notion and derive a lower bound on the fidelity F. All target states are of the form

$$\varrho_{\mathrm{t}} = \hat{U}\,|\mathbf{n}\rangle\langle\mathbf{n}|\,\hat{U}^{\dagger}, \tag{15}$$

where $\hat{U}$ is an arbitrary Gaussian unitary and $|\mathbf{n}\rangle$ an arbitrary Fock-basis state. First, we derive a general fidelity lower bound and then consider the linear-optical $\varrho_{\mathrm{t}} \in \mathcal{C}_{\mathrm{LO}}$ and Gaussian $\varrho_{\mathrm{t}} \in \mathcal{C}_{\mathrm{G}}$ cases separately. Analogous bounds for the post-selected target states are provided further below in the Measurement Scheme and Section S2.A of the SI.

We start recalling that

$$|\mathbf{n}\rangle = \prod_{j=1}^{m} \frac{1}{\sqrt{n_j!}} (\hat{a}_j^{\dagger})^{n_j}\,|\mathbf{0}\rangle\,, \tag{16}$$

where $a_j^{\dagger}$ is the creation operator of the j-th mode. Its Hermitian conjugated $\hat{a}_j$ is the corresponding annihilation operator. These operators satisfy $[\hat{a}_j, \hat{a}_{j'}^{\dagger}] = \delta_{j,j'}$, where $\delta_{j,j'}$ denotes the *Kronecker delta* of j and j', and $\hat{n}_j = \hat{a}_j^{\dagger}\hat{a}_j$, for all $j, j' \in [m]$. The fidelity (1) can be written as $F = F(|\mathbf{n}\rangle\langle\mathbf{n}|\,, \tilde{\varrho}_{\mathrm{p}})$, where $\tilde{\varrho}_{\mathrm{p}} \coloneqq \hat{U}^{\dagger}\varrho_{\mathrm{p}}\hat{U}$ is the Heisenberg representation of ϱ_{p} with respect to $\hat{U}^{\dagger}$. With this, equation (16), and the cyclicality property of the trace, we obtain that

$$F = \mathrm{Tr}\left[|\mathbf{0}\rangle\langle\mathbf{0}|\,\tilde{\varrho}_{\mathrm{p},\mathbf{n}}\right] = F(|\mathbf{0}\rangle\langle\mathbf{0}|\,, \tilde{\varrho}_{\mathrm{p},\mathbf{n}}), \tag{17}$$

where

$$\tilde{\varrho}_{\mathrm{p},\mathbf{n}} \coloneqq \prod_{j'=1}^{m} \frac{1}{\sqrt{n_{j'}!}} (\hat{a}_{j'})^{n_{j'}}\,\tilde{\varrho}_{\mathrm{p}} \prod_{j=1}^{m} \frac{1}{\sqrt{n_j!}} (\hat{a}_j^{\dagger})^{n_j}. \tag{18}$$

To lower-bound $F(|\mathbf{0}\rangle\langle\mathbf{0}|\,, \tilde{\varrho}_{\mathrm{p},\mathbf{n}})$, we consider the average total photon-number $\langle\hat{n}\rangle_{\tilde{\varrho}_{\mathrm{p},\mathbf{n}}} \coloneqq \mathrm{Tr}[\hat{n}\tilde{\varrho}_{\mathrm{p},\mathbf{n}}]$ of $\tilde{\varrho}_{\mathrm{p},\mathbf{n}}$. We write $\mathbb{1}$ for the *identity operator*. From the facts $\mathbb{1} - |\mathbf{0}\rangle\langle\mathbf{0}| \leq \hat{n}$ and $\tilde{\varrho}_{\mathrm{p},\mathbf{n}} \geq 0$, it follows that

$$\begin{aligned} \langle\hat{n}\rangle_{\tilde{\varrho}_{\mathrm{p},\mathbf{n}}} &= \mathrm{Tr}\left[\sum_{\mathbf{n}} n\,|\mathbf{n}\rangle\langle\mathbf{n}|\,\tilde{\varrho}_{\mathrm{p},\mathbf{n}}\right] \\ &\geq \mathrm{Tr}\left[(\mathbb{1} - |\mathbf{0}\rangle\langle\mathbf{0}|)\tilde{\varrho}_{\mathrm{p},\mathbf{n}}\right] \\ &= 1 - F \end{aligned} \tag{19}$$

and hence,

$$F \geq F^{(\mathbf{n})} \coloneqq 1 - \langle\hat{n}\rangle_{\tilde{\varrho}_{\mathrm{p},\mathbf{n}}}. \tag{20}$$

This bound justifies the natural extremality intuition mentioned: The lower the average number of photons of $\tilde{\varrho}_{\mathrm{p},\mathbf{n}}$ is, the closer to the vacuum it must be and, therefore, the closer

$\varrho_{\rm p}$ to $\varrho_{\rm t}$. Notice that, for $\varrho_{\rm p} = \varrho_{\rm t}$, the inequality in equation (19) becomes an equality and therefore bound (20) is saturated, as announced earlier.

Next, we define the operator valued Pochhammer-Symbol

$$p_t(\hat{n}_j) \coloneqq \hat{n}_j(\hat{n}_j - 1)(\hat{n}_j - 2)\cdots(\hat{n}_j - t), \tag{21}$$

for any integer $t \geq 0$, and $p_{-1}(x) \coloneqq \mathbb{1}$. In Section S6.A in the SI we show that

$$(\hat{a}_j^\dagger)^{n_j}\hat{n}_j(\hat{a}_j)^{n_j} = p_{n_j}(\hat{n}_j), \tag{22a}$$

and

$$(\hat{a}_j^\dagger)^{n_j}(\hat{a}_j)^{n_j} = p_{n_j-1}(\hat{n}_j). \tag{22b}$$

Inserting equation (18) into equation (20), using the cyclicity property of the trace, grouping the operators of each mode together, using equations (22) and that $p_t(\hat{n}_j) = p_{t-1}(\hat{n}_j)\,(\hat{n}_j - t)$, we obtain the general fidelity lower bound

$$F \geq F^{(\mathbf{n})} = 1 - \frac{1}{\mathbf{n}!}\left\langle (\hat{n} - n)\prod_{j=1}^{m} p_{n_j-1}(\hat{n}_j)\right\rangle_{\tilde{\varrho}_{\rm p}}, \tag{23}$$

where $\mathbf{n}! \coloneqq n_1!n_2!\ldots n_m!$. In order to specialise to the linear-optical case $\varrho_{\rm t} \in \mathcal{C}_{\rm LO}$, we simply take $\mathbf{n} = \mathbf{1}_n$, i.e., $n_j = 1$ for all $j \in [n]$ and $n_j = 0$ otherwise. With this, $F^{(\mathbf{n})}$ in equation (23) simplifies to precisely the bound $F^{(n)}$ in equation (8). Finally, to restrict it to the Gaussian case $\varrho_{\rm t} \in \mathcal{C}_{\rm G}$, we take $n_j = 0$ for all $j \in [m]$. This yields the particularly simple expression

$$F \geq F^{(0)} \coloneqq 1 - \langle \hat{n}\rangle_{\tilde{\varrho}_{\rm p}}. \tag{24}$$

Arthur does not have enough quantum capabilities to directly estimate $\langle \hat{n}\rangle_{\tilde{\varrho}_{\rm p}}$ by undoing the operation $\hat{U}$ on Merlin's outputs and then measuring $\hat{n}$ in the Fock state basis. However, we show in the next section that he can efficiently obtain $\langle \hat{n}\rangle_{\tilde{\varrho}_{\rm p}}$, as well as the expectation values in equations (23) and (8), from the results of single-mode homodyne measurements.

Measurement scheme. First, we introduce some notation. By $\hat{q}_j$ and $\hat{p}_j$ we denote, respectively, the conjugated position and momentum *phase-space quadrature operators* of the j-th mode in the canonical convention[33,34], i.e., with the commutation relations $[\hat{q}_j, \hat{p}_{j'}] = \mathrm{i}\,\delta_{j,j'}$. The *particle number operator* of the j-th mode can be written in terms of the phase-space quadratures as $\hat{n}_j = \hat{q}_j^2 + \hat{p}_j^2 - 1/2$. In addition, it will be convenient to group all quadrature operators into a $2m$-component column vector $\hat{\mathbf{r}}$, with elements

$$\hat{r}_{2j-1} \coloneqq \hat{q}_j \quad \text{and} \quad \hat{r}_{2j} \coloneqq \hat{p}_j. \tag{25}$$

As already mentioned, the action of $\hat{U}$ on mode space is given by a symplectic matrix $\mathbf{S} \in \mathrm{Sp}(2m, \mathbb{R})$ and a displacement vector $\mathbf{x} \in \mathbb{R}^{2m}$. More precisely, under a Gaussian unitary $\hat{U}$, $\hat{\mathbf{r}}$ transforms according to the affine linear map[33]

$$\hat{\mathbf{r}} \mapsto \hat{U}^\dagger\hat{\mathbf{r}}\hat{U} = \mathbf{S}\hat{\mathbf{r}} + \mathbf{x}. \tag{26}$$

Equivalently, the right-hand side of this equation defines the Heisenberg representation of $\hat{\mathbf{r}}$ with respect to $\hat{U}$. In addition, it will be useful to denote the Heisenberg representation of $\hat{\mathbf{r}}$ with respect to $\hat{U}^\dagger$ by $\hat{\tilde{\mathbf{r}}} \coloneqq \hat{U}\hat{\mathbf{r}}\hat{U}^\dagger$. Thanks to equation (26), we can write $\hat{\tilde{\mathbf{r}}}$ in terms of the symplectic matrix $\mathbf{S}$ and displacement vector $\mathbf{x}$ that define $\hat{U}$, as

$$\hat{\tilde{\mathbf{r}}} = \mathbf{S}^{-1}(\hat{\mathbf{r}} - \mathbf{x}). \tag{27}$$

The symbols $\hat{r}^2 \coloneqq \hat{\mathbf{r}}^T\hat{\mathbf{r}}$ and $\hat{\tilde{r}}^2 \coloneqq \hat{\tilde{\mathbf{r}}}^T\hat{\tilde{\mathbf{r}}}$ will represent, respectively, the scalar products of $\hat{\mathbf{r}}$ and $\hat{\tilde{\mathbf{r}}}$ with themselves. Also, we will use the same notation for the Heisenberg representations of each quadrature operator with respect to $\hat{U}^\dagger$, i.e., $\hat{\tilde{q}}_j \coloneqq \hat{U}^\dagger\hat{q}_j\hat{U}$ and $\hat{\tilde{p}}_j \coloneqq \hat{U}^\dagger\hat{p}_j\hat{U}$.

Next, for $\beta \in \{0, n, \mathbf{n}\}$, we express our fidelity bounds in the general form

$$F^{(\beta)} = 1 - \left\langle \hat{N}^{(\beta)}\right\rangle_{\varrho_{\rm p}}, \tag{28}$$

where $\hat{N}^{(\beta)}$ is an observable decomposed explicitly in terms of the local observables to which Arthur has access. We start with the Gaussian case $\varrho_{\rm t} \in \mathcal{C}_{\rm G}$. To express the bound (24) as in equation (28), we first write the total photon-number operator as

$$\hat{n} = \sum_{j=1}^{m}\hat{n}_j = \sum_{j=1}^{m}(\hat{q}_j^2 + \hat{p}_j^2 - \frac{1}{2}) = \hat{r}^2 - \frac{m}{2}. \tag{29}$$

This, in combination with equation (24), yields

$$\hat{N}^{(0)} \coloneqq \hat{\tilde{r}}^2 - \frac{m}{2}. \tag{30}$$

Note that, due to equation (27), each component of $\hat{\tilde{\mathbf{r}}}$ is a linear combination of at most $2m$ components of $\hat{\mathbf{r}}$. This implies that Arthur can obtain $\langle \hat{\tilde{r}}^2\rangle_{\varrho_{\rm p}}$ by measuring at most $2m$ single-quadrature expectation values of the form $\langle \hat{r}_k\rangle_{\varrho_{\rm p}}$ and $4m^2$ *second moments* of the form $\Gamma^{(1)}_{k,k'} \coloneqq \langle \frac{1}{2}(\hat{r}_k\hat{r}_{k'} + \hat{r}_{k'}\hat{r}_k)\rangle_{\varrho_{\rm p}}$. He can then classically efficiently combine them as dictated by $\mathbf{S}$ and $\mathbf{x}$ in equation (27). In Section S1.A of the SI, we give the details of this measurement procedure, which we call $\mathcal{M}_{\rm G}$, and show that measuring $m\,\kappa$ second moments, instead of $4m^2$, is actually enough. Furthermore, in Section S4.A of the SI, we show that only $m + 3$ experimental settings suffice.

Now, proceeding in a similar fashion with the generic bound (23), we obtain

$$\hat{N}^{(\mathbf{n})} \coloneqq \frac{1}{\mathbf{n}!}\left(\hat{\tilde{r}}^2 - \frac{m + 2n}{2}\right)\prod_{j=1}^{m} p_{n_j-1}\left(\hat{\tilde{q}}_j^2 + \hat{\tilde{p}}_j^2 - \frac{1}{2}\right). \tag{31}$$

Note that the observable in equation (30) is contained as the special case $n = 0$. For target states in the class $\mathcal{C}_{\rm LO}$, $\hat{U}$ is assumed to be a passive Gaussian unitary. Such unitaries preserve the area in phase space, i.e., if $\varrho_{\rm t} \in \mathcal{C}_{\rm LO}$ it holds that $\hat{\tilde{r}}^2 = \hat{r}^2$ (for details, see Section S1.B in the SI). Hence,

using this and specialising to the case $\mathbf{n} = \mathbf{1}_n$, equation (31) simplifies to

$$\hat{N}^{(n)} := \left(\hat{r}^2 - \frac{m+2n}{2}\right) \prod_{j=1}^{n} \left(\hat{\tilde{q}}_j^2 + \hat{\tilde{p}}_j^2 - \frac{1}{2}\right). \tag{32}$$

Again by virtue of equation (27), Arthur can now obtain the expectation values of the observables in equations (31) and (32) by measuring $2j$-*th moments* of the form $\Gamma^{(j)}_{k_1,l_1,\ldots,k_j,l_j} := \langle \frac{1}{2^j}(\hat{r}_{k_1}\hat{r}_{l_1} + \hat{r}_{l_1}\hat{r}_{k_1}) \cdots (\hat{r}_{k_j}\hat{r}_{l_j} + \hat{r}_{l_j}\hat{r}_{k_j}) \rangle_{\varrho_\mathrm{p}}$ and then classically recombining them, which — for constant n — he can do efficiently. In Section S1.B of the SI, we give the details of the measurement procedure to obtain $F^{(n)}$, which we call $\mathcal{M}_\mathrm{LO}$. In particular, we show that, to obtain $\langle \hat{N}^{(n)} \rangle_{\varrho_\mathrm{p}}$, estimating a total of $\mathrm{O}\left(m(4d^2+1)^n\right)$ $2j$-th moments, with $j \in [n+1]$, is enough. Also, we list which moments are the relevant ones in terms of $\varrho_\mathrm{t} \in \mathcal{C}_\mathrm{LO}$. Furthermore, in Section S4.B of the SI, we show that only $\binom{m}{n} 2^{n+1}$ experimental settings suffice.

Finally, in the SI, we derive a bound analogous to that of equations (28) with (31) for post-selected target states $\varrho_{\mathcal{S}_\mathrm{t}}$. More precisely, we show that the fidelity $F_\mathcal{S} := F(\varrho_{\mathcal{S}_\mathrm{t}}, \varrho_{\mathcal{S}_\mathrm{p}})$ between $\varrho_{\mathcal{S}_\mathrm{t}}$ and an arbitrary, unknown $(m-a)$-mode system preparation $\varrho_{\mathcal{S}_\mathrm{p}}$ is lower bounded as

$$F_\mathcal{S} \geq F_\mathcal{S}^{(\mathbf{n})} = 1 - \left\langle \hat{N}_\mathcal{S}^{(\mathbf{n})} \right\rangle_{\varrho_{\mathcal{S}_\mathrm{p}}}, \tag{33}$$

with

$$\hat{N}_\mathcal{S}^{(\mathbf{n})} := \frac{\mathbb{P}(\boldsymbol{\phi}_\mathcal{A}|\varrho_\mathrm{t}) - 1 + \frac{1}{\mathbf{n}!} \langle \boldsymbol{\phi}|_\mathcal{A} \hat{N}^{(\mathbf{n})} |\boldsymbol{\phi}\rangle_\mathcal{A}}{\mathbb{P}(\boldsymbol{\phi}_\mathcal{A}|\varrho_\mathrm{t})}. \tag{34}$$

From this, the corresponding expressions for the classes $\mathcal{C}_\mathrm{LPSG}$ an $\mathcal{C}_\mathrm{LPSLO}$ follow, in turn, as the two particular cases $\mathbf{n} = \mathbf{0}$ and $\mathbf{n} = \mathbf{1}_n$ with $\hat{U}$ passive, respectively. See Section S2.A of the SI for details.

Non-Gaussian state nullifiers. It is instructive to mention that the operators

$$\hat{N}_j^{(0)} := \hat{\tilde{q}}_j^2 + \hat{\tilde{p}}_j^2 - 1/2, \tag{35}$$

for $j \in [m]$, correspond to the so-called *nullifiers* of the Gaussian states in $\mathcal{C}_\mathrm{G}$. The nullifiers are commuting operators that, despite originally introduced[48] as a tool to define Gaussian graph states, can be tailored to define any pure Gaussian state[49,50]: If a state is the simultaneous null-eigenvalue eigenstate of all m nullifiers of a given pure Gaussian state, then the former is necessarily equal to the latter. The bound $F^{(0)}$, given by equations (28) and (30), exploits the fact that if a preparation gives a sufficiently low expectation value for the sum $\hat{N}^{(0)} = \sum_{j=1}^m \hat{N}_j^{(0)}$ of all m nullifiers then its fidelity with the target state must be high. A similar intuition has been previously exploited[12,13] to experimentally check for multimode entanglement of ultra-large Gaussian cluster states. Here, we can not only certify entanglement but the quantum state itself.

Analogously, in the non-Gaussian case, from the derivation of equation (31), we can identify the operator

$$\hat{N}_j^{(\mathbf{n})} := \left(\hat{\tilde{q}}_j^2 + \hat{\tilde{p}}_j^2 - \frac{1+2n_j}{2}\right) \prod_{k=1}^{m} p_{n_k-1} \left(\hat{\tilde{q}}_k^2 + \hat{\tilde{p}}_k^2 - 1/2\right) \tag{36}$$

as the j-th nullifier of the m-mode non-Gaussian state ϱ_t of equation (15). Indeed, all m observables given by equation (36) for all $j \in [m]$ commute and have ϱ_t as their unique, simultaneous null-eigenvalue eigenstate. To end up with, due to the projection onto $|\boldsymbol{\phi}\rangle_\mathcal{A}$, the equivalent observables for post-selected target states do not in general commute. Nevertheless, their sum, given by $\hat{N}_\mathcal{S}^{(\mathbf{n})}$, still defines an observable with $\varrho_{\mathcal{S}_\mathrm{t}}$ as its unique null-eigenvalue eigenstate. These observables constitute, to our knowledge[33,49,50], the first examples of nullifiers for non-Gaussian states.

[1] V. Giovannetti, S. Lloyd, and L. Maccone, *Advances in quantum metrology*, Nat. Phot. **5**, 222 (2011).
[2] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 409 (2000).
[3] P. Schindler, J. T. Barreiro, T. Monz, V. Nebendahl, D. Nigg, M. Chwalla, M. Hennrich, and R. Blatt, *Experimental repetitive quantum error correction*, Science **332**, 1059 (2011).
[4] R. Barends *et al*, *Logic gates at the surface code threshold: Superconducting qubits poised for fault-tolerant quantum computing*, Nature **508**, 500 (2014).
[5] J. I. Cirac and P. Zoller, *Goals and opportunities in quantum simulation*, Nat. Phys. **8**, 264 (2012).
[6] A. Aspuru-Guzik and P. Walther, *Photonic quantum simulators*, Nat. Phys. **8**, 285 (2012).
[7] I. Bloch, J. Dalibard, and S. Nascimbéne, *Quantum simulations with ultra-cold quantum gases*, Nat. Phys. **8**, 267 (2012).
[8] R. Blatt and C. F. Roos, *Quantum simulations with trapped ions*, Nat. Phys. **8**, 277 (2012).
[9] A. A. Houk, H. E. Türeci, and J. Koch, *On-chip quantum simulation with superconducting circuits*, Nat. Phys. **8**, 292 (2012).
[10] X-C. Yao *et al*, *Observation of eight-photon entanglement*, Nat. Phot. **6**, 225 (2012).
[11] Y.-F. Huang *et al*, *Experimental generation of an eight-photon Greenberger-Horne-Zeilinger state*, Nat. Comm. **2**, 546 (2012).
[12] M. Chen, N. C. Menicucci, O. Pfister, *Experimental realisation of multipartite entanglement of 60 modes of the quantum optical frequency comb*, Phys. Rev. Lett. **112**, 120505 (2014).
[13] S. Yokoyama *et al*, *Optical generation of ultra-large-scale continuous-variable cluster states*, Nat. Phot. **7**, 982 (2013).
[14] J. B. Spring *et al*, *Boson sampling on a photonic chip*, Science **339**, 798 (2013).
[15] M. Tillmann *et al*, *Experimental boson sampling*, Nat. Phot. **7**, 540 (2013).
[16] M. A. Broome *et al*, *Photonic boson sampling in a tunable circuit*, Science **339**, 794 (2013).
[17] A. Crespi *et al*, *Integrated multimode interferometers with arbitrary designs for photonic boson sampling*, Nat. Phot. **7**, 545 (2013).

[18] A. Peruzzo *et al*., *Quantum walks of correlated photons*, Science **329**, 1500 (2010).
[19] A. Crespi *et al*, *Anderson localization of entangled photons in an integrated quantum walk*, Nat. Phot. **7**, 322 (2013).
[20] C. Gogolin, M. Kliesch, L. Aolita, and J. Eisert, *Boson sampling in the light of sample complexity*, arXiv:1306.3995 (2013).
[21] S. Aaronson and A. Arkhipov, *BosonSampling is far from uniform*, arXiv:1309.7460 (2013).
[22] N. Spagnolo *et al*, *Experimental validation of photonic boson sampling*, Nat. Phot. **10**, 1038 (2014).
[23] J. Carolan *et al*, *On the experimental verification of quantum complexity in linear optics* arXiv:1311.2913 (2013).
[24] M. C. Tichy, K. Mayer, A. Buchleitner, and K. Molmer, *Stringent and efficient assessment of Boson-Sampling devices*, Phys. Rev. Lett. **113**, 020502 (2014).
[25] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, *Quantum state tomography via compressed sensing*, Phys. Rev. Lett. **105**, 150401 (2010).
[26] D. Aharonov, M. Ben-Or, and E. Eban, *Interactive proofs for quantum computation,* arXiv:0810.5375 (2008).
[27] A. Broadbent, J. Fitzsimons, and E. Kashefi, *Universal blind quantum computation*, Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2009), 517 (2009); see also J. Fitzsimons and E. Kashefi, *Unconditionally verifiable blind computation*, arXiv:1203.5217 (2012).
[28] T. Kapourniotis, E. Kashefi, and A. Datta, *Verified delegated quantum computing with one pure qubit*, arXiv:1403.1438 (2014).
[29] G. Toth, W. Wieczorek, D. Gross, R. Krischek, C. Schwemmer, and H. Weinfurter, *Permutationally invariant quantum tomography*, Phys. Rev. Lett. **105**, 250403 (2010).
[30] S. T. Flammia and Y.-K. Liu, *Direct fidelity estimation from few Pauli measurements*, Phys. Rev. Lett. **106**, 230501 (2011).
[31] M. P. da Silva, O. Landon-Cardinal, and D. Poulin, *Practical characterisation of quantum devices without tomography*, Phys. Rev. Lett. **107**, 210404 (2011).
[32] S. T. Flammia, D. Gross, Y.-K. Liu, and J. Eisert, *Quantum tomography via compressed sensing: Error bounds, sample complexity, and efficient estimators*, New J. Phys. **14**, 095022 (2012).
[33] C. Weedbrook, S. Pirandola, R. Garcia-Patron, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, *Gaussian quantum information*, Rev. Mod. Phys. **84**, 621 (2012).
[34] J. Eisert and M. B. Plenio, *Introduction to the basics of entanglement theory in continuous-variable systems*, Int. J. Quant. Inf. **1**, 479 (2003).
[35] S. L. Braunstein, *Squeezing as an irreducible resource*, Phys. Rev. A **71**, 055801 (2005).
[36] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, *Direct characterisation of linear-optical networks*, Phys. Rev. Lett. **73**, 58 (1993).
[37] A. Mari and J. Eisert, *Positive Wigner functions render classical simulation of quantum computation efficient*, Phys. Rev. Lett. **109**, 230503 (2012).
[38] V. Veitch, C. Ferrie, D. Gross, J. Emerson, *Negative quasi-probability as a resource for quantum computation*, New J. Phys. **14**, 113011 (2012); V. Veitch, N. Wiebe, C. Ferrie, J. Emerson, *Efficient simulation scheme for a class of quantum optics experiments with non-negative Wigner representation*, New J. Phys. **15**, 013037 (2013).
[39] S. Aaronson and A. Arkhipov, *The computational complexity of linear optics*, Theory Comput. **9**, 143 (2013).
[40] E. Magesan, J. M. Gambetta, and J. Emerson, *Robust randomized benchmarking of quantum processes*, Phys. Rev. Lett. **106**, 180504 (2011).
[41] F. Dell'Anno, S. De Siena, L. Albano, and F. Illuminati, *Continuous-variable quantum teleportation with non-Gaussian resources*, Phys. Rev. A **76**, 022301 (2007).
[42] C. Navarrete-Benlloch, R. García-Patrón, J. H. Shapiro, and N. J. Cerf, *Enhancing quantum entanglement by photon addition and subtraction*, Phys. Rev. A **86**, 012328 (2012).
[43] F. Dell'Anno, D. Buono, G. Nocerino, A. Porzio, S. Solimeno, S. De Siena, and F. Illuminati, *Tunable non-Gaussian resources for continuous-variable quantum technologies*, Phys. Rev. A **88**, 043818 (2013).
[44] J. Eisert, D. E. Browne, S. Scheel, and M. B. Plenio, *Distillation of continuous-variable entanglement*, Ann. Phys. (NY) **311**, 431 (2004).
[45] E. Knill, R. Laflamme, and G. J. Milburn, *A scheme for efficient quantum computation with linear optics*, Nature **409**, 46 (2001).
[46] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn *Linear optical quantum computing with photonic qubits*, Rev. Mod. Phys. **79**, 135 (2007).
[47] N. C. Menicucci, P. van Loock, M. Gu, C. Weedbrook, T. C. Ralph, and M. A. Nielsen, *Universal quantum computation with continuous-variable cluster states*, Phys. Rev. Lett. **97**, 110501 (2006).
[48] M. Gu, C. Weedbrook, N. C. Menicucci, T. C. Ralph, and P. van Loock, *Quantum computing with continuous-variable clusters*, Phys. Rev. A **79**, 062318 (2009).
[49] L. Aolita, A. Roncaglia, A. Ferraro, and A. Acín, *Gapped two-body Hamiltonian for continuous-variable quantum computation*, Phys. Rev. Lett. **106**, 090501 (2010).
[50] N. C. Menicucci, S. T. Flammia, and P. van Loock, *Graphical calculus for Gaussian pure states*, Phys. Rev. A **83**, 042335 (2011).

Acknowledgements

We would like to thank F. G. S. L. Brandão and S. T. Flammia for discussions on certification of state preparation. We thank the EU (RAQUEL, SIQS, AQuS, REQS - Marie Curie IEF No 299141), the BMBF, the FQXI, the Studienstiftung des Deutschen Volkes, MPQ-ICFO, and FOQUS for support.

Author Contributions

All four authors participated in all the discussions and contributed with insights. LA conceived the fidelity bound and its estimation technique. LA, CG, and MK carried out all the calculations and worked out the details of the formalism.

Supplementary Information: Reliable quantum certification for photonic quantum technologies

Leandro Aolita[1], Christian Gogolin[1,2,3], Martin Kliesch[1], and Jens Eisert[1]

[1]Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany
[2]ICFO-Institut de Ciències Fotòniques, Mediterranean Technology Park, 08860 Castelldefels (Barcelona), Spain
[3]Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, 85748 Garching, Germany

In this Supplementary Information we present the technical details of the certification test and the proofs of the theorems. It is organised as follows: In Section S1 we provide a detailed description of the measurements schemes $\mathcal{M}_{\mathrm{G}}$ and $\mathcal{M}_{\mathrm{LO}}$ for the classes of target states $\mathcal{C}_{\mathrm{G}}$ and $\mathcal{C}_{\mathrm{LO}}$, respectively. In particular, in Boxes S1 and S2 in that section, a full specification of the necessary correlators to measure is given. In Section S2, we extend Theorems 2 and 3 to the classes of post-selected target states $\mathcal{C}_{\mathrm{LPSG}}$ and $\mathcal{C}_{\mathrm{LPSLO}}$. Section S3 contains the proofs of our three theorems, as well as of the Corollaries that follow for post-selected target states. In particular, in that section, equations (S55) and (S77), in Lemmas S7 and S10, respectively, give more precise expressions of the bounds (10) and (11) of Theorems 2 and 3 in the main text. In Section S4 we upper-bound the number of experimental settings necessary for our measurement schemes. In Section S5 we analyse the stability of our fidelity estimates under systematic errors. Finally, Section S6 contains some auxiliary mathematical relations necessary for our treatments. Equation and theorem numbers that do not start with an upper case S refer to the respective equations and theorems of the main text.

Contents

S1 The measurement scheme

In this section we elaborate on the fidelity bounds $F^{(0)}$ and $F^{(n)}$ of the fidelity bounds for the Gaussian and linear-optical case, respectively. To this end, it will be convenient to first specify some details of the symplectic matrix $\mathbf{S}$, which describes the optical network.

By virtue of the Euler decomposition[6,7], $\mathbf{S}$ can be decomposed as

$$\mathbf{S} = \mathbf{O}\,\mathbf{D}\,\mathbf{O}', \tag{S1}$$

where $\mathbf{D} \in \mathbb{R}^{2m\times 2m}$ is positive-definite and diagonal, with elements $D_{2j-1,2j-1} \coloneqq s_j \geq 1$ and $D_{2j,2j} \coloneqq s_j^{-1}$, for $j \in [m]$, and $\mathbf{O} \in \mathbb{R}^{2m\times 2m}$ and $\mathbf{O}' \in \mathbb{R}^{2m\times 2m}$ are orthogonal matrices. $\mathbf{D}$ describes m active single-mode squeezers in parallel, each one with squeezing parameter s_j along the position quadrature. The maximum single-mode squeezing is $s_{\max} \coloneqq \max_{1\leq j\leq m}\{s_j\}$. $\mathbf{O}$ and $\mathbf{O}'$, in turn, describe passive mode transformations that can be implemented by linear-optical networks of at most $m(m-1)/2$ beam-splitters and single-mode phase shifters[8]. In the two settings considered here, i.e., for any $\varrho_{\mathrm{t}} \in \mathcal{C}_{\mathrm{G}} \cup \mathcal{C}_{\mathrm{LO}}$, the unitary $\hat{U}$ in equations (4) and (5) is such that $\mathbf{O}'$ can be taken as the identity matrix. In the first setting, i.e., for $\varrho_{\mathrm{t}} \in \mathcal{C}_{\mathrm{G}}$, this holds because $\hat{U}$ acts on the vacuum state vector $|\mathbf{0}\rangle$ and any passive mode transformation maps the vacuum into itself. For the second setting, i.e., for $\varrho_{\mathrm{t}} \in \mathcal{C}_{\mathrm{LO}}$, this holds simply because there we assume that the total transformation itself is passive, i.e., in that case it holds also that $\mathbf{D} = \mathbb{1}$, so that $\mathbf{S} = \mathbf{O}$.

In both cases, coupling between different modes only takes place through the linear-optical network described by $\mathbf{O}$. A general circuit can couple all m modes with each other, meaning that the quadrature operators of each output mode are linear combinations of those of all m input modes. However, often, each mode is only coupled to at most $d \leq m$ other modes. In these situations, $\mathbf{O}$ is a sparse matrix with at most $4md$ non-zero elements. More precisely, the columns of $\mathbf{O}$ are given by $2m$ orthonormal vectors $(\mathbf{o}^{(k)})_{k\in[2m]}$ each having at most $2d$ non-zero entries. Furthermore, since the position and momentum of each mode is coupled to at most the $2d$ quadratures of the same d modes, each pair $\mathbf{o}^{(2j-1)}$ and $\mathbf{o}^{(2j)}$ shares the same *sparsity property*, i.e., $\mathbf{o}^{(2j-1)}$ and $\mathbf{o}^{(2j)}$ have at least $2(m-d)$ zero entries in common, for all $j \in [m]$.

S1.A Gaussian case

Using that in the Gaussian case $\mathbf{S} = \mathbf{O}\,\mathbf{D}$ and squaring equation (27) yields

$$\begin{aligned}\hat{\tilde{r}}^2 &= \hat{\mathbf{r}}^T\mathbf{O}\mathbf{D}^{-2}\mathbf{O}^{-1}\hat{\mathbf{r}} - 2\mathbf{x}^T\mathbf{O}\mathbf{D}^{-2}\mathbf{O}^{-1}\hat{\mathbf{r}} + \mathbf{x}^T\mathbf{O}\mathbf{D}^{-2}\mathbf{O}^{-1}\mathbf{x} \\ &= \mathrm{Tr}\left[\mathbf{O}\mathbf{D}^{-2}\mathbf{O}^T[\hat{\mathbf{r}}\hat{\mathbf{r}}^T - (2\hat{\mathbf{r}} - \mathbf{x})\mathbf{x}^T]\right],\end{aligned} \tag{S2}$$

where $\mathbf{O}^{-1} = \mathbf{O}^T$ has been used and the trace is taken not over the Hilbert space but over the $2m \times 2m$ matrix with operators as entries. Combining equations (S2), (28), and (30) yields

$$F^{(0)} = 1 - \mathrm{Tr}\left[\mathbf{O}\mathbf{D}^{-2}\mathbf{O}^T[\langle\hat{\mathbf{r}}\hat{\mathbf{r}}^T\rangle_{\varrho_{\mathrm{p}}} - (2\langle\hat{\mathbf{r}}\rangle_{\varrho_{\mathrm{p}}} - \mathbf{x})\mathbf{x}^T]\right] + \frac{m}{2}. \tag{S3}$$

Now we introduce the *first moment vector* $\boldsymbol{\gamma} \in \mathbb{R}^{2m}$ and the symmetric *second moment matrix* $\boldsymbol{\Gamma}^{(1)} \in \mathbb{R}^{2m\times 2m}$ of ϱ_{p}, with components

$$\gamma_l \coloneqq \langle\hat{r}_l\rangle_{\varrho_{\mathrm{p}}} \quad \text{and} \quad \Gamma^{(1)}_{l,l'} \coloneqq \left\langle \frac{\hat{r}_l\hat{r}_{l'} + \hat{r}_{l'}\hat{r}_l}{2} \right\rangle_{\varrho_{\mathrm{p}}}, \tag{S4}$$

respectively. Since the matrix $\mathbf{O}\mathbf{D}^{-2}\mathbf{O}^T$ is symmetric, it holds that

$$\mathrm{Tr}\left[\mathbf{O}\mathbf{D}^{-2}\mathbf{O}^T[\langle\hat{\mathbf{r}}\hat{\mathbf{r}}^T\rangle_{\varrho_{\mathrm{p}}}\right] = \mathrm{Tr}\left[\mathbf{O}\mathbf{D}^{-2}\mathbf{O}^T[\langle\hat{\mathbf{r}}\hat{\mathbf{r}}^T\rangle^T_{\varrho_{\mathrm{p}}}\right], \tag{S5}$$

so that we can rewrite equation (S3) in terms of the observables which Arthur has access to as

$$F^{(0)} = 1 - \mathrm{Tr}\left[\mathbf{O}\mathbf{D}^{-2}\mathbf{O}^{-1}[\boldsymbol{\Gamma}^{(1)} - (2\boldsymbol{\gamma} - \mathbf{x})\mathbf{x}^T]\right] + \frac{m}{2}. \tag{S6}$$

We will show later (see Lemma S5 in Section S3.C and the discussion immediately after its proof) that the bound (S6) actually depends on at most $2m\kappa$ out of the $4m^2$ entries of $\boldsymbol{\Gamma}^{(1)}$, with $\kappa = 2\min\{d^2, m\}$, as defined in equation (9). Thus, only the $2m\kappa$ corresponding observables, and the $2m$ observables necessary for $\boldsymbol{\gamma}$, as indicated in Box S1, need to be measured. All

Box S1 (Measurement scheme $\mathcal{M}_{\mathrm{G}}$)**.**
1) For each $1 \leq l \leq 2m$ Arthur uses C_1 copies of ϱ_{p}, with C_1 given by equation (S56a), to measure the observable $\hat{r}_l$, obtaining an estimate γ_l^* of the expectation value $\gamma_l = \langle \hat{r}_l \rangle_{\varrho_{\mathrm{p}}}$.
2) For each $1 \leq l \leq l' \leq 2m$ for which $(\mathbf{O}\mathbf{D}^{-2}\mathbf{O}^{-1})_{l,l'} = \sum_{k=1}^{2m} o_l^{(k)} D_{k,k}^{-2} o_{l'}^{(k)} \neq 0$, Arthur uses C_2 copies of ϱ_{p}, with C_2 given by equation (S56b), to measure the observable $\frac{1}{2}(\hat{r}_l\hat{r}_{l'} + \hat{r}_{l'}\hat{r}_l)$, obtaining an estimate $\Gamma^{(1)*}_{l,l'}$ of the expectation values $\Gamma^{(1)}_{l,l'} = \Gamma^{(1)}_{l',l}$ in equation (S4).
3) He obtains the estimate $F^{(0)*}$ of $F^{(0)}$ by replacing in equation (S6) the actual expectation values $\mathbf{\Gamma}^{(1)}$ and $\boldsymbol{\gamma}$ by the estimates $\Gamma^{(1)*}$ and $\boldsymbol{\gamma}^*$, respectively.

these observables can be measured by homodyne detection[6]. Furthermore, in Section S4.A we show that only $m+3$ different measurement settings are required. Finally, by classical post-processing, Arthur recombines his estimates according to the third step of Box S1 and obtains the fidelity estimate $F^{(0)*}$. This last step is also efficient in m.

S1.B Linear-optical case

For $\varrho_{\mathrm{t}} \in \mathcal{C}_{\mathrm{LO}}$ the unitary $\hat{U}$ is assumed to be passive. Hence, one has $\mathbf{x} = \mathbf{0}$ and $\mathbf{S} = \mathbf{O}$, and it follows that

$$\hat{\tilde{r}}^2 = \hat{r}^2 \,. \tag{S7}$$

The components of $\tilde{r}$ are

$$\hat{\tilde{q}}_j = \mathbf{o}^{(2j-1)^T}\hat{\mathbf{r}} \quad \text{and} \quad \hat{\tilde{p}}_j = \mathbf{o}^{(2j)^T}\hat{\mathbf{r}}, \tag{S8}$$

where $\mathbf{o}^{(k)}$ denotes the k-th column of $\mathbf{O}$. Defining

$$\mathbf{P}^{(j)} \coloneqq \mathbf{o}^{(2j-1)}\mathbf{o}^{(2j-1)^T} + \mathbf{o}^{(2j)}\mathbf{o}^{(2j)^T} \tag{S9}$$

as the projector onto the subspace spanned by the two vectors $\mathbf{o}^{(2j-1)}$ and $\mathbf{o}^{(2j)}$ and using equations (S8), (28), and (32), we obtain

$$F^{(n)} = 1 - \left\langle \left(\hat{r}^2 - \frac{m+2n}{2}\right) \prod_{j=1}^{n} \left(\hat{\mathbf{r}}^T\mathbf{P}^{(j)}\hat{\mathbf{r}} - \frac{1}{2}\right) \right\rangle_{\varrho_{\mathrm{p}}} . \tag{S10}$$

Next, we consider the $\binom{n}{j}$ subsets of $\{1, 2, \ldots, n\}$ of length j and define $\Omega_\mu^{(j)}$ as the μ-th of these subsets for some arbitrary ordering. With this, we expand the product inside (S10) as

$$\prod_{j=1}^{n} \left(\hat{\mathbf{r}}^T\mathbf{P}^{(j)}\hat{\mathbf{r}} - \frac{1}{2}\right) = \sum_{j=0}^{n} (-1/2)^{n-j} \sum_{\mu=1}^{\binom{n}{j}} \bigotimes_{i \in \Omega_\mu^{(j)}} \hat{\mathbf{r}}^T\mathbf{P}^{(i)}\hat{\mathbf{r}}. \tag{S11}$$

Using that a product of traces can be written as a trace over tensor products, equation (S10) can be written as

$$F^{(n)} = 1 - \left\langle \left(\hat{r}^2 - \frac{m+2n}{2}\right) \sum_{j=0}^{n} (-1/2)^{n-j} \sum_{\mu=1}^{\binom{n}{j}} \mathrm{Tr}\left[\left(\bigotimes_{i \in \Omega_\mu^{(j)}} \mathbf{P}^{(i)} \right) (\hat{\mathbf{r}}\hat{\mathbf{r}}^T)^{\otimes j} \right] \right\rangle_{\varrho_{\mathrm{p}}} , \tag{S12}$$

where $\bigotimes_{i \in \Omega_\mu^{(0)} = \emptyset} \mathbf{P}^{(i)} \coloneqq 1$ and the traces are again taken not over the Hilbert space but over tensors that have operators as components. For each $j \in [n+1]$, we introduce the *$2j$-th moment tensors* $\mathbf{\Gamma}^{(j)} \in (\mathbb{R}^{2m \times 2m})^{\otimes j}$ with components

$$\Gamma^{(j)}_{k_1,l_1,\ldots,k_j,l_j} \coloneqq \left\langle \frac{\hat{r}_{k_1}\hat{r}_{l_1} + \hat{r}_{l_1}\hat{r}_{k_1}}{2} \cdots \frac{\hat{r}_{k_j}\hat{r}_{l_j} + \hat{r}_{l_j}\hat{r}_{k_j}}{2} \right\rangle_{\varrho_{\mathrm{p}}} \tag{S13}$$

and define $\mathbf{\Gamma}^{(0)} \coloneqq 1$. Clearly, these tensors are invariant under the partial transposition with respect to any j'-th pair of subindices

Box S2 (Measurement scheme $\mathcal{M}_{\mathrm{LO}}$)**.**
1) For each $1 \leq j \leq n$*, each* $1 \leq \mu \leq \binom{n}{j}$*, and each* $1 \leq k_1, l_1, k_2, l_2, \ldots, k_j, l_j \leq 2m$*, for which*

$$\Big(\bigotimes_{i \in \Omega_\mu^{(j)}} \mathbf{P}^{(i)} \Big)_{k_1,l_1,k_2,l_2,\ldots,k_j,l_j} \neq 0, \tag{S16}$$

2) Arthur uses $C_{\leq 2(n+1)}$ *copies of* ϱ_{p}*, with* $C_{\leq 2(n+1)}$ *given by equation (S78), to measure the observable* $(\hat{r}_{k_1}\hat{r}_{l_1} + \hat{r}_{l_1}\hat{r}_{k_1})/2 \cdots (\hat{r}_{k_j}\hat{r}_{l_j} + \hat{r}_{l_j}\hat{r}_{k_j})/2$*, obtaining an estimate* ${\Gamma^{(j)*}}_{k_1,l_1,k_2,l_2,\ldots,k_j,l_j}$ *of the* $2j$*-th moment* $\Gamma^{(j)}_{k_1,l_1,k_2,l_2,\ldots,k_j,l_j}$*. For each* $1 \leq k_{j+1} \leq 2m$*, he uses* $C_{\leq 2(n+1)}$ *copies of* ϱ_{p} *to measure the observable* $((\hat{r}_{k_1}\hat{r}_{l_1} + \hat{r}_{l_1}\hat{r}_{k_1})/2) \cdots ((\hat{r}_{k_j}\hat{r}_{l_j} + \hat{r}_{l_j}\hat{r}_{k_j})/2)\hat{r}^2_{k_{j+1}}$*, obtaining an estimate* ${\Gamma^{(j+1)*}}_{k_1,l_1,k_2,l_2,\ldots,k_j,l_j,k_{j+1},k_{j+1}}$ *of the* $2(j+1)$*-th moment* $\Gamma^{(j+1)}_{k_1,l_1,k_2,l_2,\ldots,k_j,l_j,k_{j+1},k_{j+1}}$*.*
3) He obtains the estimate $F^{(\mathbf{n})*}$ *of* $F^{(n)}$ *by replacing in equation (S15) for all* $1 \leq j \leq n+1$ *the actual expectation values* $\mathbf{\Gamma}^{(j)}$ *by the estimates* $\mathbf{\Gamma}^{(j)*}$*.*

$k_{j'}$ and $l_{j'}$,

$$\Gamma^{(j)}_{k_1,l_1,\ldots,k_{j'},l_{j'},\ldots,k_j,l_j} = \Gamma^{(j)}_{k_1,l_1,\ldots,l_{j'},k_{j'},\ldots,k_j,l_j}. \tag{S14}$$

With the definition (S13) and the fact that each projector $\mathbf{P}^{(i)}$ is a symmetric matrix, equation (S12) finally becomes

$$F^{(n)} = 1 - \sum_{j=0}^{n} (-1/2)^{n-j} \sum_{\mu=1}^{\binom{n}{j}} \left\{ \mathrm{Tr}\left[\Big(\mathbb{1} \otimes \bigotimes_{i \in \Omega_\mu^{(j)}} \mathbf{P}^{(i)} \Big) \mathbf{\Gamma}^{(j+1)} \right] - \frac{m+2n}{2} \mathrm{Tr}\left[\Big(\bigotimes_{i \in \Omega_\mu^{(j)}} \mathbf{P}^{(i)} \Big) \mathbf{\Gamma}^{(j)} \rangle_{\varrho_{\mathrm{p}}} \right] \right\}. \tag{S15}$$

Note that this is an explicit expression for $F^{(n)}$ in terms of the correlators (S13) that Arthur can measure. Due to the sparsity of $\mathbf{O}$, each matrix $\mathbf{P}^{(i)}$ has at most $(2d)^2$ non-zero entries. Then, it follows (see Lemma S8 in Section S3.D for details) that the measurement of $\mathrm{O}\left(m\left(4d^2+1\right)^n\right)$ observables, those listed in Box S2, suffices for the estimation of (S15). As in the Gaussian case, all these observables can be measured by homodyne detection[6]. Furthermore, in Section S4.B we show that at most $\binom{m}{n} 2^{n+1} \leq (2m)^n/n!$ measurement settings are sufficient. Once again, by classical post-processing, Arthur recombines his estimates according to the third step of Box S2 and obtains the fidelity estimate $F^{(n)*}$. Provided that n is constant, this last step is also efficient in m.

S2 Quantum certification of locally post-selected target states

In this section, we extend our results to locally post-selected $(m-a)$-mode target states $\varrho_{\mathcal{S}_{\mathrm{t}}}$ in $\mathcal{C}_{\mathrm{LPSG}}$ or $\mathcal{C}_{\mathrm{LPSLO}}$. The entire treatment of the classes $\mathcal{C}_{\mathrm{LPSG}}$ or $\mathcal{C}_{\mathrm{LPSLO}}$ is similar to, and follows directly from, that already seen for the classes $\mathcal{C}_{\mathrm{G}}$ or $\mathcal{C}_{\mathrm{LO}}$. Therefore, instead of repeating all the details, we simply explain the specific differences.

S2.A The fidelity bound

The first step is to derive the fidelity bound $F_{\mathcal{S}}^{(\mathbf{n})}$ given by equation (33). We proceed in a similar fashion to the Methods Section in the main text. Due to equations (1) and (3), the facts that $\varrho_{\mathcal{S}_{\mathrm{t}}}$ and ϱ_{t} are pure, and the properties of the trace, it holds that

$$F_{\mathcal{S}} = F(\varrho_{\mathcal{S}_{\mathrm{t}}}, \varrho_{\mathcal{S}_{\mathrm{p}}}) = \mathrm{Tr}_{\mathcal{S}}\left[\mathrm{Tr}_{\mathcal{A}}\left[\frac{\varrho_{\mathrm{t}}(\mathbb{1}_{\mathcal{S}} \otimes |\boldsymbol{\phi}\rangle_{\mathcal{A}} \langle\boldsymbol{\phi}|_{\mathcal{A}})}{\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})} \right] \varrho_{\mathcal{S}_{\mathrm{p}}} \right] = \frac{\mathrm{Tr}\left[\varrho_{\mathrm{t}}(\varrho_{\mathcal{S}_{\mathrm{p}}} \otimes |\boldsymbol{\phi}\rangle_{\mathcal{A}} \langle\boldsymbol{\phi}|_{\mathcal{A}})\right]}{\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})} = \frac{F\left(\varrho_{\mathrm{t}}, \varrho_{\mathcal{S}_{\mathrm{p}}} \otimes |\boldsymbol{\phi}\rangle_{\mathcal{A}} \langle\boldsymbol{\phi}|_{\mathcal{A}}\right)}{\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})}, \tag{S17}$$

Box S3 (Certification test $\mathcal{T}_{\text{LPS}}$)**.**
1) Idem as in $\mathcal{T}$ from Box 1.
2) Arthur provides Merlin with the classical specification n, $\mathbf{S}$, $\mathbf{x}$, a, and $|\phi\rangle_{\mathcal{A}}$ of the target state $\varrho_{\mathcal{S}_{\text{t}}}$ and requests a sufficient number of copies of it.
3) If $n = 0$, Arthur measures $2m\kappa$ two-mode correlations and $2(m-a)$ single-mode expectation values specified by the measurement scheme $\mathcal{M}_{\text{LPSG}}$ (see Section S2.C), which can be done with $m-a+3$ single-mode homodyne settings. If $n > 0$, he measures $\mathrm{O}\left(m(4d^2+1)^n\right)$ multi-body correlators, each one involving between 1 and $2n+1$ modes, specified by the measurement scheme $\mathcal{M}_{\text{LPSLO}}$ (see Section S2.C), which can be done with at most $\binom{m}{n}2^{n+1}$ single-mode homodyne settings.
4) By classical post-processing, he obtains a fidelity estimate $F_{\mathcal{S}}^{(n)*}$ such that $F_{\mathcal{S}}^{(n)*} \in [F_{\mathcal{S}}^{(n)} - \varepsilon, F_{\mathcal{S}}^{(n)} + \varepsilon]$ with probability at least $1-\alpha$, where $F_{\mathcal{S}}^{(n)}$ is the lower bound to $F_{\mathcal{S}}$ given by equation (S20).
5) If $F_{\mathcal{S}}^{(n)*} < F_{\text{T}} + \varepsilon$, he rejects. Otherwise, he accepts.

where $\text{Tr}_{\mathcal{S}}$ indicates partial trace over the Fock space of the $m-a$ modes in $\mathcal{S}$. Now, due to equations (28) and (31), it holds that

$$F\left(\varrho_{\text{t}}, \varrho_{\mathcal{S}_{\text{p}}} \otimes |\boldsymbol{\phi}\rangle_{\mathcal{A}} \langle\boldsymbol{\phi}|_{\mathcal{A}}\right) \geq 1 - \text{Tr}\left[\hat{N}^{(\mathbf{n})}(\varrho_{\mathcal{S}_{\text{p}}} \otimes |\boldsymbol{\phi}\rangle_{\mathcal{A}} \langle\boldsymbol{\phi}|_{\mathcal{A}})\right] = 1 - \text{Tr}_{\mathcal{S}}\left[\langle\boldsymbol{\phi}|_{\mathcal{A}} \hat{N}^{(\mathbf{n})} |\boldsymbol{\phi}\rangle_{\mathcal{A}} \varrho_{\mathcal{S}_{\text{p}}}\right], \tag{S18}$$

with $\hat{N}^{(\mathbf{n})}$ the observable of equation (31). Using equations (S17) and (S18), we obtain the general fidelity bound $F_{\mathcal{S}}^{(\mathbf{n})}$ of equation (33), with $\hat{N}_{\mathcal{S}}^{(\mathbf{n})}$ given by (34).

In particular, setting $\mathbf{n} = \mathbf{0}$ in equations (33) and (34) yields the specialized fidelity bound $F_{\mathcal{S}}^{(0)} \geq 1 - \left\langle \hat{N}_{\mathcal{S}}^{(0)} \right\rangle_{\varrho_{\mathcal{S}_{\text{p}}}}$ for the case $\varrho_{\mathcal{S}_{\text{t}}} \in \mathcal{C}_{\text{LPSG}}$, with

$$\hat{N}_{\mathcal{S}}^{(0)} := \frac{\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\text{t}}) - 1 + \langle\boldsymbol{\phi}|_{\mathcal{A}} \hat{N}^{(0)} |\boldsymbol{\phi}\rangle_{\mathcal{A}}}{\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\text{t}})}, \tag{S19}$$

where $\hat{N}^{(0)}$ is the observable of equation (30) and ϱ_{t} is the m-mode state in $\mathcal{C}_{\text{G}}$ associated with $\varrho_{\mathcal{S}_{\text{t}}}$ through equation (3). Analogously, taking $\mathbf{n} = \mathbf{1}_n$ and $\hat{U}$ passive yields the corresponding fidelity bound $F_{\mathcal{S}}^{(n)} \geq 1 - \left\langle \hat{N}_{\mathcal{S}}^{(n)} \right\rangle_{\varrho_{\mathcal{S}_{\text{p}}}}$ for $\varrho_{\mathcal{S}_{\text{t}}} \in \mathcal{C}_{\text{LPSLO}}$, with

$$\hat{N}_{\mathcal{S}}^{(n)} := \frac{\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\text{t}}) - 1 + \langle\boldsymbol{\phi}|_{\mathcal{A}} \hat{N}^{(n)} |\boldsymbol{\phi}\rangle_{\mathcal{A}}}{\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\text{t}})}, \tag{S20}$$

where $\hat{N}^{(n)}$ is the observable from equation (31) and ϱ_{t} is the m-mode state in $\mathcal{C}_{\text{LO}}$ associated to $\varrho_{\mathcal{S}_{\text{t}}}$ through equation (3).

S2.B The certification test

Next, in Box S3, we present a test $\mathcal{T}_{\text{LPS}}$ that works for post-selected target states in $\mathcal{C}_{\text{LPSG}}$ or $\mathcal{C}_{\text{LPSLO}}$ and which is a slightly modified version of the test $\mathcal{T}$ from Box 1. It is, of course, possible to unify both tests so as to account for all four classes of target states in one single test. We have however opted for splitting the tests into the two cases with and without post-selection to avoid an excessive notational overhead in Box 1 of the main text.

S2.C The measurement scheme

The measurement schemes $\mathcal{M}_{\text{LPSG}}$ and $\mathcal{M}_{\text{LPSLO}}$ to estimate $F_{\mathcal{S}}^{(0)}$ and $F_{\mathcal{S}}^{(n)}$, respectively, are essentially replicas of the schemes $\mathcal{M}_{\text{G}}$ and $\mathcal{M}_{\text{LO}}$ to estimate $F^{(0)}$ and $F^{(n)}$, already described in detail in boxes S1 and S2. Thus, instead of repeating all the details of boxes S1 and S2, we simply outline the concrete differences between $\mathcal{M}_{\text{LPSG}}$ and $\mathcal{M}_{\text{G}}$, as well as between $\mathcal{M}_{\text{LO}}$ and $\mathcal{M}_{\text{LPSLO}}$. There are only three specific differences.

1. The moment vector and tensors are now defined with respect to $\varrho_{\mathcal{S}_{\text{p}}} \otimes |\boldsymbol{\phi}\rangle_{\mathcal{A}} \langle\boldsymbol{\phi}|_{\mathcal{A}}$ instead of ϱ_{p}. More precisely, we now

need to estimate the vector $\boldsymbol{\gamma}_{\mathcal{S}} \in \mathbb{R}^{2m}$ and tensors $\boldsymbol{\Gamma}^{(j)}_{\mathcal{S}} \in \left(\mathbb{R}^{2m\times 2m}\right)^{\otimes j}$, with elements

$$\gamma_{\mathcal{S}l} \coloneqq \langle \hat{r}_l \rangle_{\varrho_{\mathcal{S}_\mathrm{p}} \otimes |\boldsymbol{\phi}\rangle_{\mathcal{A}}\langle\boldsymbol{\phi}|_{\mathcal{A}}} = \langle \langle \boldsymbol{\phi}|_{\mathcal{A}}\, \hat{r}_l\, |\boldsymbol{\phi}\rangle_{\mathcal{A}} \rangle_{\varrho_{\mathcal{S}_\mathrm{p}}} = \begin{cases} \langle \phi_l|_{\mathcal{A}_l}\, \hat{r}_l\, |\phi_l\rangle_{\mathcal{A}_l} & , \text{ if } l \in \mathcal{A}, \\ \langle \hat{r}_l \rangle_{\varrho_{\mathcal{S}_\mathrm{p}}} & , \text{ if } l \notin \mathcal{A}, \end{cases} \tag{S21}$$

and

$$\begin{aligned} \Gamma^{(j)}_{\mathcal{S}\, k_1,l_1,\ldots,k_j,l_j} &\coloneqq \left\langle \frac{\hat{r}_{k_1}\hat{r}_{l_1} + \hat{r}_{l_1}\hat{r}_{k_1}}{2} \cdots \frac{\hat{r}_{k_j}\hat{r}_{l_j} + \hat{r}_{l_j}\hat{r}_{k_j}}{2} \right\rangle_{\varrho_{\mathcal{S}_\mathrm{p}} \otimes |\boldsymbol{\phi}\rangle_{\mathcal{A}}\langle\boldsymbol{\phi}|_{\mathcal{A}}} \\ &= \left\langle \langle\boldsymbol{\phi}|_{\mathcal{A}} \frac{\hat{r}_{k_1}\hat{r}_{l_1} + \hat{r}_{l_1}\hat{r}_{k_1}}{2} \cdots \frac{\hat{r}_{k_j}\hat{r}_{l_j} + \hat{r}_{l_j}\hat{r}_{k_j}}{2} |\boldsymbol{\phi}\rangle_{\mathcal{A}} \right\rangle_{\varrho_{\mathcal{S}_\mathrm{p}}}, \end{aligned} \tag{S22}$$

respectively.

2. $F^{(0)}_{\mathcal{S}}$ and $F^{(n)}_{\mathcal{S}}$ are obtained dividing the expressions on the right-hand sides of equations (S6) and (S15) by $\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_\mathrm{t})$, and with $\boldsymbol{\gamma}$ and $\boldsymbol{\Gamma}^{(j)}$ replaced by $\boldsymbol{\gamma}_{\mathcal{S}}$ and $\boldsymbol{\Gamma}^{(j)}_{\mathcal{S}}$, respectively.

3. The presence of the divisor $\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_\mathrm{t})$ in $F^{(0)}_{\mathcal{S}}$ and $F^{(n)}_{\mathcal{S}}$ is the reason for the third difference. As discussed in Lemmas S12 and S15 in Sections S3.E and S3.F, respectively, this divisor makes $F^{(0)}_{\mathcal{S}}$ and $F^{(n)}_{\mathcal{S}}$ $1/\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_\mathrm{t})$ times more unstable than $F^{(0)}$ and $F^{(n)}$. As a consequence, the number of copies of $\varrho_{\mathcal{S}_\mathrm{p}}$ required to estimate each relevant moment of $F^{(0)}_{\mathcal{S}}$ are $\frac{C_1}{\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_\mathrm{t})}$ and $\frac{C_2}{\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_\mathrm{t})}$, instead of C_1 and C_2. This is summarized in Lemma S13. Analogously, the number required for each relevant moment of $F^{(n)}_{\mathcal{S}}$ is $\frac{C_{\leq 2(n+1)}}{\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_\mathrm{t})}$, instead of $C_{\leq 2(n+1)}$. This is summarized in Lemma S16.

As is clear from equations (S21) and (S22), the estimation of $\boldsymbol{\gamma}_{\mathcal{S}}$ and $\boldsymbol{\Gamma}^{(j)}_{\mathcal{S}}$ requires only the measurement of multi-body correlators among the $(m-a)$ system modes in $\mathcal{S}$. This is due to the facts that after post selection the system is in a product state with respect to the bipartition in $\mathcal{S}$ and $\mathcal{A}$ and that the quadrature operators in equation (S22) can also be correspondingly grouped into two factors, one containing exclusively operators of modes in $\mathcal{S}$ and the other in $\mathcal{A}$. Furthermore, since $|\boldsymbol{\phi}\rangle_{\mathcal{A}}$ is a tensor-product state known to Arthur, he can efficiently calculate the expectation vale of any product of quadrature operators of modes in $\mathcal{A}$ with respect to $|\boldsymbol{\phi}\rangle_{\mathcal{A}}$. For instance, suppose that $k_1, l_1, k_2 \in \mathcal{A}$ and that $l_2, k_3, l_3 \ldots, k_j, l_j \notin \mathcal{A}$. Then, the corresponding $2j$-th moment decomposes as

$$\Gamma^{(j)}_{\mathcal{S}\, k_1,l_1,\ldots,k_j,l_j} = \langle\boldsymbol{\phi}|_{\mathcal{A}} \frac{\hat{r}_{k_1}\hat{r}_{l_1} + \hat{r}_{l_1}\hat{r}_{k_1}}{2} \hat{r}_{k_2} |\boldsymbol{\phi}\rangle_{\mathcal{A}} \left\langle \hat{r}_{l_2} \frac{\hat{r}_{k_3}\hat{r}_{l_3} + \hat{r}_{l_3}\hat{r}_{k_3}}{2} \cdots \frac{\hat{r}_{k_j}\hat{r}_{l_j} + \hat{r}_{l_j}\hat{r}_{k_j}}{2} \right\rangle_{\varrho_{\mathcal{S}_\mathrm{p}}}, \tag{S23}$$

and only the measurement of the $(2j-3)$-th moment given by the second factor of equation (S23) is required. As another example, consider the case where a given $|\phi_j\rangle_{\mathcal{A}_j}$ is a Fock-basis state. Then, all the moments containing an odd number of quadrature operators of the $\mathcal{A}_j$-th mode automatically vanish and need therefore not be measured at all.

In general, Arthur can always efficiently obtain $\boldsymbol{\gamma}_{\mathcal{S}}$ and the $\boldsymbol{\Gamma}^{(j)}_{\mathcal{S}}$'s as a product of an (a priori known) expectation value with respect to $|\boldsymbol{\phi}\rangle_{\mathcal{A}}$ of a multi-body product of quadrature operators of modes in $\mathcal{A}$ and a (measured) expectation value with respect to $\varrho_{\mathcal{S}_\mathrm{p}}$ of a multi-body product of quadrature operators of modes in $\mathcal{S}$, in a way analogous to the example of equation (S23).

S2.D Corollaries of Theorems 2 and 3

Since the moments to be estimated are now given, in equations (S21) and (S22), by expectation values with respect to $\varrho_{\mathcal{S}_\mathrm{p}} \otimes |\boldsymbol{\phi}\rangle_{\mathcal{A}}\langle\boldsymbol{\phi}|_{\mathcal{A}}$, instead of ϱ_p, a simple way to extend Theorems 2 and 3 to target states in $\mathcal{C}_{\mathrm{LPSG}}$ or $\mathcal{C}_{\mathrm{LPSLO}}$ is by redefining the variance upper bounds σ_i. More precisely, taking σ_i as an upper bound on the variances of any product of i phase space quadratures now in the state $\varrho_{\mathcal{S}_\mathrm{p}}$, we introduce the quantities

$$\varsigma_i \coloneqq \max_{j\in[a]\wedge k_1,k_2,\ldots k_j \in \mathcal{A}} \left\{ \langle\boldsymbol{\phi}|_{\mathcal{A}}\, \hat{r}_{k_1}\hat{r}_{k_2}\ldots\hat{r}_{k_j}\, |\boldsymbol{\phi}\rangle_{\mathcal{A}}\, \sigma_{i-j} \right\}, \tag{S24}$$

for $i \in [2(n+1)]$. In addition, we call $\varsigma_{\leq i} \coloneqq \max_{k\leq i}\{\varsigma_k\}$ the *maximal i-th generalised variance* of $\varrho_{\mathcal{S}_\mathrm{p}}$.

The parameters ς_i quantify the maximal variances of random variables defined by products of $i-j$ quadrature-measurement outcomes on $\varrho_{\mathcal{S}_\mathrm{p}}$ renormalised by the expectation value of products of j quadrature operators with respect to $|\boldsymbol{\phi}\rangle_{\mathcal{A}}$, therefore automatically accounting for factorisations of the type of equation (S23). They constitute very non-tight upper bounds to the real variances. In particular experimental situations, tighter bounds can be found. Here, we are simply interested in taking

advantage of the proofs of Theorems 2 and 3 without introducing too much extra notational overhead, for which the definition of equation (S24) is enough. Indeed, with these redefinitions, the following corollaries follow straightforwardly from Theorems 2 and 3.

Corollary S1 (Quantum certification of locally post-selected Gaussian states)**.** *Under the same conditions and for the same* ϱ_{t} *as in Theorem 2, test* $\mathcal{T}_{LPS}$ *from Box S3 is a certification test for* $\varrho_{\mathcal{S}_{\mathrm{t}}} \in \mathcal{C}_{\mathrm{LPSG}}$ *and requires at most*

$$\mathrm{O}\left(\frac{s_{\max}^4\left(2\varsigma_1^2\|\mathbf{x}\|_2^2 m^3+\varsigma_2^2\kappa^3 m^4\right)}{\left[\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})\,\varepsilon\right]^2\ln(1/(1-\alpha))}\right) \tag{S25}$$

copies of a preparation $\varrho_{\mathcal{S}_{\mathrm{p}}}$ *with first and second generalized variance bounds* $\varsigma_1 > 0$ *and* $\varsigma_2 > 0$*, respectively.*

Corollary S2 (Quantum certification of locally post-selected linear-optical network states)**.** *Under the same conditions and for the same* ϱ_{t} *as in Theorem 3, test* $\mathcal{T}_{LPS}$ *from Box S3 is a certification test for* $\varrho_{\mathcal{S}_{\mathrm{t}}} \in \mathcal{C}_{\mathrm{LPSLO}}$ *and requires at most*

$$\mathrm{O}\left(\frac{\varsigma_{\leq 2(n+1)}^2 m^4(\lambda\, d^6\, n\, m)^n}{\left[\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})\,\varepsilon\right]^2\ln(1/(1-\alpha))}\right) \tag{S26}$$

copies of a preparation $\varrho_{\mathcal{S}_{\mathrm{p}}}$ *with maximal* $2(n+1)$*-th generalised variance* $\varsigma_{\leq 2(n+1)}$*, where* $\lambda > 0$ *is the same absolute constant as in Theorem 3.*

Corollary S1 is proven in Section S3.E and Corollary S2 in Section S3.F. Equations (S25) and (S26) correspond to exactly the same expressions as in equations (10) and (11), respectively, with the replacements $\sigma \to \varsigma$ and $\varepsilon \to \mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})\,\varepsilon$. The rescaling of ε with the factor $\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})$ originates directly from the new expression for the fidelity given in equation (S17). As mentioned in Section S2.C, this makes the fidelity bounds $F_{\mathcal{S}}^{(0)}$ and $F_{\mathcal{S}}^{(n)}$ more unstable than $F^{(0)}$ and $F^{(n)}$, leading to the error rescaling discussed earlier. In most interesting cases, the post-selection success probability $\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})$ turns out to be exponentially small in a. Moreover, one can always come up with families of target states and post selection procedures for which $\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})$ decreases arbitrarily fast in m. In such cases, the scalings in equations (S25) and (S26) are not efficient in m, inheriting the inefficiency of the state preparation by local measurements and post selection. However, both bounds are efficient in $1/\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})$. That is, in every practical situation where state preparation via post selection is feasible, so is state certification. Interestingly, even for families of target states and post selection procedures for which $\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})$ decays exponentially in a, the overall scaling (of both bounds) with a is better than the scalings (of the bounds in equations (11) and (S26)) with n. Indeed, the bound in equation (S26) grows, just like the bound in equation (11), faster than exponentially in n. Finally, both bounds (S25) and (S26) scale polynomially with all the other relevant parameters, including $1/\varepsilon$. Thus, arbitrary m-mode target states from the classes $\mathcal{C}_{\mathrm{LPSG}}$ and $\mathcal{C}_{\mathrm{LPSLO}}$, with constant n, are certified by $\mathcal{T}_{\mathrm{LPS}}$ efficiently in m, $1/\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})$, and all the other relevant parameters.

S2.E Corollary of Theorem 5

Finally, it is possible to show that our certification test is robust also for the locally post-selected target states of the classes $\mathcal{C}_{\mathrm{LPSG}}$ or $\mathcal{C}_{\mathrm{LPSLO}}$. Writing $\varrho_{\mathcal{S}_{\mathrm{p}}}$ as

$$\varrho_{\mathcal{S}_{\mathrm{p}}} = F_{\mathcal{S}}\varrho_{\mathcal{S}_{\mathrm{t}}} + (1-F_{\mathcal{S}})\varrho_{\mathcal{S}_{\mathrm{t}}}^{\perp}, \tag{S27}$$

where $\varrho_{\mathcal{S}_{\mathrm{t}}}^{\perp}$ is such that $\mathrm{Tr}[\varrho_{\mathcal{S}_{\mathrm{t}}}\,\varrho_{\mathcal{S}_{\mathrm{t}}}^{\perp}] = 0$, and introducing the *generalised photon mismatch* $\tilde{n}_{\mathcal{S}}^{\perp}$ between $\varrho_{\mathcal{S}_{\mathrm{t}}}$ and $\varrho_{\mathcal{S}_{\mathrm{p}}}$ as

$$\tilde{n}_{\mathcal{S}}^{\perp} := \left\langle \frac{\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}}) - 1 + (\hat{n}-n)\prod_{j=1}^{n}\hat{n}_j}{\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})} \right\rangle_{\hat{U}^{\dagger}\varrho_{\mathcal{S}_{\mathrm{t}}}^{\perp}\otimes|\boldsymbol{\phi}\rangle_{\mathcal{A}}\langle\boldsymbol{\phi}|_{\mathcal{A}}\hat{U}} = \left\langle \hat{N}_{\mathcal{S}}^{(n)} \right\rangle_{\varrho_{\mathcal{S}_{\mathrm{t}}}^{\perp}}, \tag{S28}$$

where $\hat{N}_{\mathcal{S}}^{(n)}$ is the same observable as in (S20), the following holds true.

Corollary S3 (Robust quantum certification of locally post-selected states)**.** *Under the same conditions as in Corollaries S1 and S2, test* $\mathcal{T}$ *from Box 1 is a robust certification test for* $\varrho_{\mathcal{S}_{\mathrm{t}}} \in \mathcal{C}_{\mathrm{LPSG}} \cup \mathcal{C}_{\mathrm{LPSLO}}$ *with fidelity gap*

$$\Delta_{\mathcal{S}} := \max\left\{\frac{2\varepsilon + (1-F_{\mathrm{T}})(\tilde{n}_{\mathcal{S}}^{\perp}-1)}{\tilde{n}_{\mathcal{S}}^{\perp}}, 2\varepsilon\right\}, \tag{S29}$$

where $\tilde{n}_{\mathcal{S}}^{\perp}$ *is the generalised photon mismatch.*

The proof is identical to the proof of Theorem 5 presented in Section S3.G but with the replacements $F \to F_{\mathcal{S}}$, $F^{(n)} \to F_{\mathcal{S}}^{(n)}$, $F^{(n)*} \to F_{\mathcal{S}}^{(n)*}$, $\Delta \to \Delta_{\mathcal{S}}$, and $\tilde{n}^{\perp} \to \tilde{n}_{\mathcal{S}}^{\perp}$.

S3 Proofs of the theorems and corollaries

Before going to the proofs, we devote two sections to establish necessary notation, review some known facts, and prove a general lemma.

S3.A Norms

Here, we introduce some helpful notation used in the proofs and review a few facts about norms on finite dimensional vector spaces. The *max norm* $\|\cdot\|_{\max}$ of a tensor is the largest of the absolute values of its entries. For a matrix $\mathbf{A}$, for example, $\|\mathbf{A}\|_{\max} := \max_{k,l} |A_{k,l}|$. For $p \in [1,\infty]$, we denote the *vector p-norm* of a vector $\mathbf{a}$ by $\|\mathbf{a}\|_p$ and the *Schatten p-norm* of a matrix $\mathbf{A}$ by $\|\mathbf{A}\|_p$, which is defined to be the vector p-norm of the vector of its singular values. For any matrix $\mathbf{A}$, we define $\mathrm{vec}(\mathbf{A})$ to be a vector containing all the entries of $\mathbf{A}$ (in some order). Then one can see that

$$\|\mathbf{A}\|_2 = \|\operatorname{vec}(\mathbf{A})\|_2 \tag{S30}$$

and

$$\|\mathbf{A}\|_{\max} = \|\operatorname{vec}(\mathbf{A})\|_\infty. \tag{S31}$$

For the vector and Schatten p-norm of vectors with N elements and $N \times N$ matrices, respectively, the following inequalities hold

$$\|\cdot\|_1 \leq \sqrt{N}\|\cdot\|_2 \leq N\|\cdot\|_\infty. \tag{S32}$$

Because the Schatten ∞-norm is induced by the vector 2-norm, i.e.,

$$\|\mathbf{A}\|_\infty = \sup_{\mathbf{y}} \frac{\|\mathbf{A}\mathbf{y}\|_2}{\|\mathbf{y}\|_2}, \tag{S33}$$

it follows that for any two vectors $\boldsymbol{\epsilon}$ and $\mathbf{x}$

$$\|\boldsymbol{\epsilon}\mathbf{x}^T\|_\infty \leq \|\boldsymbol{\epsilon}\|_2 \|\mathbf{x}\|_2. \tag{S34}$$

S3.B Reliable estimation of expectation values from samples

We continue by proving a general large-deviation bound for estimates of expectation values from a finite number of measurements on independent copies, which we need for the proofs of Theorems 2 and 3.

Lemma S4 (Reliable estimation of multiple expectation values from samples)**.** *Let $\sigma > 0$, ρ be a state, and let $\hat{A}_1, \ldots, \hat{A}_N$ be observables with expectation values $A_i := \mathrm{Tr}[\rho \hat{A}_i]$ and variances bounded as $\mathrm{Tr}[\rho \hat{A}_i^2] - A_i^2 \leq \sigma^2$. For each $i \in [N]$ and χ, let $X_i^{(\chi)}$ be the random variable given by the measurement statistics of $\hat{A}_i$ on state ρ; such that, in particular, the $(X_i^{(\chi)})_{i,\chi}$ are independent random variables and the* finite sample average *over c measurements of $\hat{A}_i$ is the random variable*

$$A_i^* := \frac{1}{c} \sum_{\chi=1}^{c} X_i^{(\chi)}. \tag{S35}$$

Then, the $\{A_i^*\}_i$ are independent and, for every $\epsilon > 0$ and $\overline{\alpha} \in [1/2, 1)$, it holds that

$$\mathbb{P}\big[\forall i : \ |A_i^* - A_i| \leq \varepsilon\big] \geq \overline{\alpha} \tag{S36}$$

whenever

$$c \geq \frac{\sigma^2 (N+1)}{\varepsilon^2 \ln(1/\overline{\alpha})}. \tag{S37}$$

Proof. The sample averages $\{A_i^*\}_i$ are independent by definition. By Chebyshev's inequality it holds that

$$\forall i \in [c]: \quad \mathbb{P}\big[|A_i^* - A_i| > \varepsilon\big] < \frac{\sigma^2}{c\varepsilon^2}. \tag{S38}$$

Since the $\{A_i^*\}_i$ are independent random variables, this yields

$$\mathbb{P}\big[\forall i: \ |A_i^* - A_i| \leq \varepsilon\big] \geq \left(1 - \frac{\sigma^2}{c\varepsilon^2}\right)^N. \tag{S39}$$

Finally,

$$\left(1 - \frac{\sigma^2}{c\varepsilon^2}\right)^N \geq \overline{\alpha} \tag{S40}$$

is satisfied if

$$c \geq c_{\text{opt}} \coloneqq \left\lceil \frac{\sigma^2/\varepsilon^2}{1 - \overline{\alpha}^{1/N}} \right\rceil. \tag{S41}$$

To finish the proof we upper bound

$$c_{\text{opt}} = \left\lceil \frac{\sigma^2/\varepsilon^2}{1 - \mathrm{e}^{-\frac{\ln(1/\overline{\alpha})}{N}}} \right\rceil. \tag{S42}$$

Using that (see Section S6.B) for all $x \geq 0$

$$\frac{1}{1 - \mathrm{e}^{-1/x}} \leq x + \frac{1}{2 + 2x} + \frac{1}{2}, \tag{S43}$$

it follows that

$$c_{\text{opt}} \leq \frac{\sigma^2}{\varepsilon^2}\left(\frac{N}{\ln(1/\overline{\alpha})} + \frac{1}{2 + \frac{2N}{\ln(1/\overline{\alpha})}} + \frac{1}{2}\right). \tag{S44}$$

To simplify the right-hand side of this inequality, we use that, since $\overline{\alpha} \geq \frac{1}{2} \geq \mathrm{e}^{-1}$, it holds that $\ln(1/\overline{\alpha}) \leq 1$ and, therefore, $2 + \frac{2N}{\ln(1/\overline{\alpha})} \geq 4$. So, using again that $\ln(1/\overline{\alpha}) \leq 1$, we finally arrive at

$$c_{\text{opt}} \leq \frac{\sigma^2}{\varepsilon^2}\left(\frac{N}{\ln(1/\overline{\alpha})} + \frac{3}{4}\right) \leq \frac{\sigma^2(N+1)}{\varepsilon^2 \ln(1/\overline{\alpha})}. \tag{S45}$$

□

S3.C Proof of Theorem 2

Before the proof of Theorem 2, we present three auxiliary lemmas specific to the fidelity bound $F^{(0)}$ for the Gaussian case.

The first lemma upper-bounds the number of elements of $\mathbf{\Gamma}^{(1)}$ which the fidelity bound $F^{(0)}$ depends on.

Lemma S5 (Sparsity of the Gaussian fidelity bound)**.** *$F^{(0)}$ depends on at most $2m\kappa$ elements of $\mathbf{\Gamma}^{(1)}$. We call these the* relevant elements *of $\mathbf{\Gamma}^{(1)}$.*

Proof. Equation (S6) can be written as

$$F^{(0)} = 1 + \frac{m}{2} + \mathbf{x}^T\mathbf{O}\mathbf{D}^{-2}\mathbf{O}^T(2\boldsymbol{\gamma} - \mathbf{x}) - \mathrm{Tr}\big[\mathbf{O}\mathbf{D}^{-2}\mathbf{O}^T\mathbf{\Gamma}\big]. \tag{S46}$$

The last term can, in turn, be expressed as

$$\begin{aligned}\mathrm{Tr}\left[\mathbf{O}\mathbf{D}^{-2}\mathbf{O}^T\mathbf{\Gamma}^{(1)}\right] &= \sum_{k=1}^{2m} D_{k,k}^{-2}(\mathbf{o}^{(k)})^T\mathbf{\Gamma}^{(1)}\mathbf{o}^{(k)} \\ &= \mathrm{Tr}\left[\sum_{j=1}^{m}\left\{s_j^{-2}\mathbf{o}^{(2j-1)}(\mathbf{o}^{(2j-1)})^T + s_j^2\mathbf{o}^{(2j)}(\mathbf{o}^{(2j)})^T\right\}\mathbf{\Gamma}^{(1)}\right]. \end{aligned} \tag{S47}$$

Due to the sparsity of $\mathbf{O}$, as described in Section S1, each matrix $s_j^{-2}\mathbf{o}^{(2j-1)}(\mathbf{o}^{(2j-1)})^T + s_j^2\mathbf{o}^{(2j)}(\mathbf{o}^{(2j)})^T$ has at most $4d^2$ non-zero elements. Hence, summing over j, we see that $F^{(0)}$ depends on at most $4m\min\{d^2, m\} = 2\kappa m$ elements of $\mathbf{\Gamma}^{(1)}$. □

Note that the counting argument following equation (S47) does not take into account the fact that $\mathbf{\Gamma}^{(1)}$ is symmetric. Taking this fact into account, we see that, from the $4d^2$ relevant elements of $\mathbf{\Gamma}^{(1)}$ that appear in each term of the trace (S47), only $d(2d+1)$ are independent. Thus, even though $2m\kappa$ entries of $\mathbf{\Gamma}^{(1)}$ contribute to $F^{(0)}$, only $m\min\{d(2d+1), 4m\} \leq 2m\kappa$ of them must actually be measured.

The second auxiliary lemma bounds the deviation of $F^{(0)*}$ from $F^{(0)}$ in terms of the errors made in the estimation of the individual expectation values entering $F^{(0)}$.

Lemma S6 (Stability of the Gaussian fidelity bound)**.** *Let $F^{(0)*}$ be defined like $F^{(0)}$ in equation (S6) but with $\boldsymbol{\gamma}$ and $\mathbf{\Gamma}^{(1)}$ replaced by $\boldsymbol{\gamma}^*$ and $\mathbf{\Gamma}^{(1)*}$ and let $\epsilon_{\max} \coloneqq \|\boldsymbol{\gamma} - \boldsymbol{\gamma}^*\|_{\max}$ and $\varepsilon^{(1)}_{\max} \coloneqq \|\mathbf{\Gamma}^{(1)} - \mathbf{\Gamma}^{(1)*}\|_{\max}$. Then*

$$|F^{(0)} - F^{(0)*}| \leq 2s_{\max}^2\left(\varepsilon^{(1)}_{\max}\sqrt{\kappa}m + \epsilon_{\max}\|\mathbf{x}\|_2\sqrt{2m}\right). \tag{S48}$$

Proof. For convenience, we define the *error vector*

$$\boldsymbol{\epsilon} \coloneqq \boldsymbol{\gamma} - \boldsymbol{\gamma}^* \in \mathbb{R}^{2m} \tag{S49}$$

and the *error matrix*

$$\mathcal{E}^{(1)} \coloneqq \mathbf{\Gamma}^{(1)} - \mathbf{\Gamma}^{(1)*}. \tag{S50}$$

The fidelity estimation error can then be written as

$$F^{(0)} - F^{(0)*} = \mathrm{Tr}\left[\mathbf{O}\mathbf{D}^{-2}\mathbf{O}^T(\mathcal{E}^{(1)} + 2\boldsymbol{\epsilon}\mathbf{x}^T)\right]. \tag{S51}$$

Due to Hölder's inequality,

$$\begin{aligned}|F^{(0)} - F^{(0)*}| &\leq \|\mathbf{O}\mathbf{D}^{-2}\mathbf{O}^T\|_\infty\|\mathcal{E}^{(1)} + 2\boldsymbol{\epsilon}\mathbf{x}^T\|_1 \\ &\leq \|\mathbf{D}^{-2}\|_\infty\left(\|\mathcal{E}^{(1)}\|_1 + 2\|\boldsymbol{\epsilon}\|_2\|\mathbf{x}\|_2\right), \end{aligned} \tag{S52}$$

where in the last step we have used the bound (S34). The second inequality in equation (S32) implies that $\|\boldsymbol{\epsilon}\|_2 \leq \sqrt{2m}\|\boldsymbol{\epsilon}\|_\infty$. It remains to bound $\|\mathcal{E}\|_1$. To this end, we use the first inequality in equation (S32) and equation (S30) to arrive at

$$\|\mathcal{E}^{(1)}\|_1 \leq \sqrt{2m}\|\,\mathrm{vec}(\mathcal{E}^{(1)})\|_2. \tag{S53}$$

According to Lemma S5, $F^{(0)}$ depends on at most $2\kappa m$ entries of $\mathcal{E}^{(1)}$. Without loss of generality we can hence omit all other elements of $\mathcal{E}^{(1)}$ and thus take $\mathrm{vec}(\mathcal{E}^{(1)})$ as a vector with at most $2\kappa m$ elements. Using this fact and the second inequality in equation (S32) we obtain

$$\begin{aligned}\|\mathcal{E}^{(1)}\|_1 &\leq \sqrt{2m}\sqrt{2m\kappa}\|\,\mathrm{vec}(\mathcal{E}^{(1)})\|_\infty \\ &= 2m\sqrt{\kappa}\|\mathcal{E}^{(1)}\|_{\max}, \end{aligned} \tag{S54}$$

where we have used equation (S31) in the last equality. Finally, putting everything together and using that, by definition, $\|\mathbf{D}^{-2}\|_\infty = s_{\max}^2$, we arrive at the inequality (S48). □

The third auxiliary Lemma shows that the estimate of the fidelity lower-bound for target states $\varrho_t \in \mathcal{C}_G$ obtained with the measurement scheme $\mathcal{M}_G$ in Box S1 is reliable. This Lemma is potentially interesting in its own right in scenarios other than

certification.

Lemma S7 (Reliable estimation of the Gaussian fidelity bound)**.** *Let $\alpha \in (0, 1/2]$ and $\varepsilon > 0$. Let $F^{(0)*}$ be defined like $F^{(0)}$ in equation (S6) but with $\boldsymbol{\gamma}$ and $\boldsymbol{\Gamma}^{(1)}$ replaced by $\boldsymbol{\gamma}^*$ and $\boldsymbol{\Gamma}^{(1)*}$, where $\boldsymbol{\gamma}^*$ and $\boldsymbol{\Gamma}^{(1)*}$ are obtained as described by $\mathcal{M}_{\mathrm{G}}$ from*

$$C = 2mC_1 + 2\kappa m C_2 \tag{S55}$$

copies of ϱ_{p} , with C_1 and C_2 integers such that

$$C_1 \geq 2^6 \frac{\sigma_1^2 (2m+1)\, m\, s_{\max}^4\, \|\mathbf{x}\|_2^2}{\varepsilon^2 \ln\left(\frac{1}{1-\alpha}\right)} \tag{S56a}$$

and

$$C_2 \geq 2^5 \frac{\sigma_2^2 (2\kappa m+1)\, m^2\, s_{\max}^4\, \kappa}{\varepsilon^2 \ln\left(\frac{1}{1-\alpha}\right)}. \tag{S56b}$$

Then,

$$\mathbb{P}\left[|F^{(0)} - F^{(0)*}| \leq \varepsilon\right] \geq 1 - \alpha. \tag{S57}$$

Proof. Our proof strategy is to show that, with probability at least $1-\alpha$, the $2m$ elements of $\boldsymbol{\gamma}$ and the $2m\kappa$ relevant elements of $\boldsymbol{\Gamma}^{(1)}$ are estimated within additive errors bounded as

$$\epsilon_{\max} \leq \epsilon_{\max}^* := \frac{\varepsilon}{4 s_{\max}^2 \|\mathbf{x}\|_2 \sqrt{2m}} \tag{S58a}$$

and

$$\varepsilon_{\max}^{(1)} \leq \varepsilon_{\max}^{*(1)} := \frac{\varepsilon}{4 s_{\max}^2 \sqrt{\kappa m}}. \tag{S58b}$$

If the inequalities (S58) are fulfiled, then, due to Lemma S6, it holds that $|F^{(0)} - F^{(0)*}| \leq \varepsilon$.

Since all $2m$ estimates $\{\gamma_l^*\}_l$ are sample averages over independent copies of ϱ_{p}, the measurement outcomes to obtain the $\{\gamma_l^*\}_l$ are all independent random variables, for each l described by the same probability distribution. Furthermore, by assumption, the variances of these variables are all upper-bounded by σ_1. Analogously, the measurement outcomes to obtain all $2m\kappa$ relevant estimates $\{\Gamma_{l,l'}^{(1)*}\}_{l,l'}$ are independent random variables with variances upper-bounded by σ_2 and described, for each l and l', by the same probability distribution. Hence, according to Lemma S4, with the choice $\overline{\alpha} = \sqrt{1-\alpha}$, taking

$$C_1 \geq 2 \frac{\sigma_1^2 (2m+1)}{{\epsilon_{\max}^*}^2 \ln\left(\frac{1}{1-\alpha}\right)} \tag{S59a}$$

and

$$C_2 \geq 2 \frac{\sigma_2^2 (2\kappa m+1)}{{\varepsilon_{\max}^{*(1)}}^2 \ln\left(\frac{1}{1-\alpha}\right)}, \tag{S59b}$$

is sufficient for both

$$\mathbb{P}\left[\forall l:\ |\gamma_l^* - \gamma_l| \leq \epsilon_{\max}^*\right] \geq \sqrt{1-\alpha} \tag{S60a}$$

and

$$\mathbb{P}\left[\forall\, \Gamma_{l,l'}^{(1)} \text{ relevant}:\ |\Gamma_{l,l'}^{(1)*} - \Gamma_{l,l'}^{(1)}| \leq \varepsilon_{\max}^{*(1)}\right] \geq \sqrt{1-\alpha}. \tag{S60b}$$

Since the $\{\gamma_l^*\}_l$ and the $\{\Gamma_{l,l'}^{(1)*}\}_{l,l'}$ are independent random variables, equations (S60) imply that

$$\mathbb{P}\left[\begin{array}{rl} & \forall l: \ |\gamma_l^* - \gamma_l| \ \leq \ \epsilon_{\max}^* \\ \text{and } \forall\, \Gamma_{l,l'}^{(1)} \text{ relevant}: & |\Gamma_{l,l'}^{(1)*} - \Gamma_{l,l'}^{(1)}| \ \leq \ \varepsilon_{\max}^{*(1)} \end{array}\right] \geq 1-\alpha. \tag{S61}$$

Finally, inserting the definitions (S58) of $\epsilon_{\max}^*$ and $\varepsilon_{\max}^{*(1)}$ into equations (S59), we see that equations (S56) are equivalent to equations (S59). □

Now, we prove the theorem on quantum certification of Gaussian states.

Proof of Theorem 2. That the total number of copies of ϱ_{p} (see equation (S55)) needed for the certification test is asymptotically upper-bounded by equation (10) can be verified by straightforward calculation using equations (S56). It remains to show that (i) if $\varrho_{\mathrm{p}} = \varrho_{\mathrm{t}}$, then $\mathcal{T}$ accepts with probability at least $1-\alpha$, i.e.,

$$\mathbb{P}\left[F^{(0)*} \geq F_T + \varepsilon\right] \geq 1-\alpha, \tag{S62}$$

and (ii) if ϱ_{p} is such that $F < F_T$, then $\mathcal{T}$ rejects with probability at least $1-\alpha$, i.e.,

$$\mathbb{P}\left[F^{(0)*} < F_{\mathrm{T}} + \varepsilon\right] \geq 1-\alpha. \tag{S63}$$

To show (i), we first recall that, if $\varrho_{\mathrm{p}} = \varrho_{\mathrm{t}}$, $F^{(0)} = 1$. With this, equation (S57) in Lemma S7 implies that

$$\mathbb{P}\left[F^{(0)*} \geq 1-\varepsilon\right] \geq 1-\alpha. \tag{S64}$$

Since, by assumption of the theorem, the total estimation error is such that $\varepsilon \leq \frac{1-F_{\mathrm{T}}}{2}$, it holds that $1-\varepsilon \geq F_T + \varepsilon$. Substituting the latter inequality into equation (S64) yields equation (S62).

To show (ii), we first note that, since $F^{(0)} \leq F$ for all ϱ_{p}, if $F < F_T$, then

$$F^{(0)} < F_{\mathrm{T}}. \tag{S65}$$

On the other hand, equation (S57) implies also that

$$\mathbb{P}\left[F^{(0)*} \leq F^{(0)} + \varepsilon\right] \geq 1-\alpha. \tag{S66}$$

Inserting equation (S65) into equation (S66) yields equation (S63). □

S3.D Proof of Theorem 3

We proceed analogously to the last section and present three auxiliary lemmas specific to the fidelity bound $F^{(n)}$ for the linear-optical case before proving Theorem 3.

To state the first lemma in a compact form we introduce the shorthand $\mathbf{\Gamma} := \left(\mathbf{\Gamma}^{(i)}\right)_{i=1,\ldots,n+1}$ for the collection of all the moment tensors $\mathbf{\Gamma}^{(i)}$. Analogously, the collection of all the estimates $\mathbf{\Gamma}^{(i)*}$ of the moment tensors, defined in Box S2, is denoted by $\mathbf{\Gamma}^* := \left(\mathbf{\Gamma}^{(i)*}\right)_{i=1,\ldots,n+1}$.

Lemma S8 (Sparsity of the linear-optical fidelity bound)**.** *The fidelity bound $F^{(n)}$ defined in equation (S15) can be written as*

$$F^{(n)} = 1 - \sum_{j=0}^{n} (-1/2)^{n-j} f_j\left(\mathbf{\Gamma}^{(j)}, \mathbf{\Gamma}^{(j+1)}\right), \tag{S67}$$

where, for each $j \in \{0,\ldots,n\}$, f_j is a linear functional given by

$$f_j\left(\mathbf{\Gamma}^{(j)}, \mathbf{\Gamma}^{(j+1)}\right) := \sum_{\mu=1}^{\binom{n}{j}} \left\{ \mathrm{Tr}\left[\left(\mathbb{1} \otimes \bigotimes_{i\in\Omega_\mu^{(j)}} \mathbf{P}^{(i)}\right)\mathbf{\Gamma}^{(j+1)}\right] + \frac{m+2n}{2}\,\mathrm{Tr}\left[\left(\bigotimes_{i\in\Omega_\mu^{(j)}} \mathbf{P}^{(i)}\right)\mathbf{\Gamma}^{(j)}\right]\right\}. \tag{S68}$$

For each j, the functional f_j depends on at most $\binom{n}{j}(2d)^{2j}$ elements of $\mathbf{\Gamma}^{(j)}$ and on at most $\binom{n}{j}2m(2d)^{2j}$ elements of $\mathbf{\Gamma}^{(j+1)}$. We call these the relevant elements for f_j. *Moreover, $F^{(n)}$ depends on at most*

$$N_{\leq 2(n+1)} \coloneqq (1+2m)(4d^2+1)^n \in \mathrm{O}\left(m\left(4d^2+1\right)^n\right) \tag{S69}$$

elements of $\mathbf{\Gamma}$. We call these the relevant elements *of $\mathbf{\Gamma}$.*

The subindex "$\leq 2(n+1)$" in $N_{\leq 2(n+1)}$ makes reference to the fact that $2j$-th moments with $j \in [n+1]$ are taken into account.

Proof. Equations (S67) and (S68) can be checked by a straightforward calculation. We use again the sparsity of $\mathbf{O}$, i.e., the property that its columns $\mathbf{o}^{(2j-1)}$ and $\mathbf{o}^{(2j)}$ have at least $2(m-d)$ zero element in common. Hence, each of the symmetric matrices $\mathbf{P}^{(j)}$, defined in equation (S9), has at most $(2d)^2$ non-zero elements. Consequently, the projectors $\bigotimes_{i\in\Omega_\mu^{(j)}}\mathbf{P}^{(i)}$ and $\mathbb{1}\otimes\bigotimes_{i\in\Omega_\mu^{(j)}}\mathbf{P}^{(i)}$ in equation (S68) have at most $(2d)^{2j}$ and $2m(2d)^{2j}$ non-zero elements. This implies that the first trace inside the sum in equation (S68) depends on at most $2m(2d)^{2j}$ elements of $\mathbf{\Gamma}^{(j+1)}$ and the second trace inside the sum on at most $(2d)^{2j}$ elements of $\mathbf{\Gamma}^{(j)}$. Hence, each f_j depends on at most $\binom{n}{j}(2d)^{2j}$ elements of $\mathbf{\Gamma}^{(j)}$ and on at most $\binom{n}{j}2m(2d)^{2j}$ elements of $\mathbf{\Gamma}^{(j+1)}$. This proves the statements on the sparsity of the functionals f_i. From this, it follows that $F^{(n)}$ depends on at most

$$\sum_{i=0}^{n}\left(\binom{n}{i}(2d)^{2i}+\binom{n}{i}2m(2d)^{2i}\right) = (1+2m)(4d^2+1)^n \tag{S70}$$

elements of $\mathbf{\Gamma}$ in total, where in the last step we have used the binomial theorem. □

It is important to mention that, as in Lemma S5 for the Gaussian case, the symmetry (S14) of each $\mathbf{\Gamma}^{(j)}$ was not taken into account. Thus, even though the lemma gives the maximal total number of relevant elements that contribute to $F^{(n)}$, many of them are not independent and must therefore not be measured.

The second auxiliary lemma upper-bounds the deviation of $F^{(n)*}$ from $F^{(n)}$ in terms of the errors made in the estimation of the expectation values entering $F^{(n)}$.

Lemma S9 (Stability of the linear-optical fidelity bound)**.** *Let $F^{(n)*}$ be defined like $F^{(n)}$ in equation (S15) but with $\mathbf{\Gamma}$ replaced by $\mathbf{\Gamma}^*$ and let $\varepsilon_{\max} \coloneqq \|\mathbf{\Gamma}-\mathbf{\Gamma}^*\|_{\max}$. Then*

$$|F^{(n)}-F^{(n)*}| \leq \varepsilon_{\max}\,(n+5m/2)\left(1/2+2d\sqrt{2nm}\right)^n. \tag{S71}$$

Proof. For convenience, we define, for each $j\in[n]$, the *error tensor*

$$\mathcal{E}^{(j)} \coloneqq \Gamma^{(j)}-\Gamma^{(j)*} \in \left(\mathbb{R}^{2m\times 2m}\right)^{\otimes j}. \tag{S72}$$

Using equation (S67) and the fact that f_j is linear, we write the fidelity estimation error as

$$F^{(n)}-F^{(n)*} = \sum_{j=0}^{n}(-1/2)^{n-j}f_j\left(\mathcal{E}^{(j)},\mathcal{E}^{(j+1)}\right), \tag{S73}$$

Applying Hölder's inequality and using that the Schatten ∞-norm of a tensor product of projectors is bounded by 1 yields

$$|f_j\left(\mathcal{E}^{(j)},\mathcal{E}^{(j+1)}\right)| \leq \binom{n}{j}\left(\left\|\tilde{\mathcal{E}}^{(j+1)}\right\|_1+\frac{m+2n}{2}\left\|\tilde{\mathcal{E}}^{(j)}\right\|_1\right),$$

where the matrix $\tilde{\mathcal{E}}^{(j)}$ is defined element-wise by $\tilde{\mathcal{E}}^{(j)}_{\mathbf{k}^{(j)},\mathbf{l}^{(j)}} \coloneqq \mathcal{E}^{(j)}_{k_1,l_1,\ldots,k_j,l_j}$, where $\mathbf{k}^{(j)} \coloneqq (k_1,\ldots,k_j)$ and $\mathbf{l}^{(j)} \coloneqq (l_1,\ldots,l_j)$. Thanks to the first bound in equation (S32) and equation (S30), we arrive at

$$|f_j\left(\mathcal{E}^{(j)},\mathcal{E}^{(j+1)}\right)| \leq \binom{n}{j}(2m)^{j/2}\left(\sqrt{2m}\left\|\mathrm{vec}(\tilde{\mathcal{E}}^{(j+1)})\right\|_2+\frac{m+2n}{2}\left\|\mathrm{vec}(\tilde{\mathcal{E}}^{(j)})\right\|_2\right). \tag{S74}$$

According to Lemma S8, f_j depends on at most $\binom{n}{j}2m(2d)^{2j}$ elements of $\tilde{\mathcal{E}}^{(j+1)}$ and on at most $\binom{n}{j}(2d)^{2j}$ of $\tilde{\mathcal{E}}^{(j)}$. Without loss of generality we can hence omit, in equation (S74), all other elements in $\tilde{\mathcal{E}}^{(j)}$ and $\tilde{\mathcal{E}}^{(j+1)}$ and thus take $\mathrm{vec}(\tilde{\mathcal{E}}^{(j)})$ and $\mathrm{vec}(\tilde{\mathcal{E}}^{(j+1)})$

as vectors with at most $\binom{n}{j}(2d)^{2j}$ and $\binom{n}{j}2m(2d)^{2j}$ elements, respectively. Then the second bound in equation (S32) yields

$$|f_j\left(\mathcal{E}^{(j)},\mathcal{E}^{(j+1)}\right)| \leq \binom{n}{j}^{3/2}(2m)^{j/2}(2d)^j\left[2m\|\tilde{\mathcal{E}}^{(j+1)}\|_{\max}+\frac{m+2n}{2}\|\tilde{\mathcal{E}}^{(j)}\|_{\max}\right]. \tag{S75}$$

Next, from equation (S73), it follows that

$$|F^{(n)}-F^{(n)*}| \leq \varepsilon_{\max}\left[\sum_{j=0}^{n}\binom{n}{j}^{3/2}(1/2)^{n-j}\left(\sqrt{2m}2d\right)^j\times(5m/2+n)\right]. \tag{S76}$$

Finally, using $\binom{n}{j}^{1/2}\leq n^{j/2}$ and the binomial formula, we obtain the inequality (S71). □

The third auxiliary Lemma shows that the estimate of the fidelity lower-bound for target states $\varrho_{\mathrm{t}}\in\mathcal{C}_{\mathrm{LO}}$ obtained with the measurement scheme $\mathcal{M}_{\mathrm{LO}}$ in Box S2 is reliable. This Lemma is potentially interesting in its own right in scenarios other than certification.

Lemma S10 (Reliable estimation of the linear-optical fidelity bound)**.** *Let $\alpha\in(0,1/2]$ and $\varepsilon>0$. Let $F^{(n)*}$ be defined like $F^{(n)}$ in equation (S15) but with $\mathbf{\Gamma}$ replaced by $\mathbf{\Gamma}^*$, where $\mathbf{\Gamma}^*$ is obtained as described by $\mathcal{M}_{\mathrm{LO}}$ from*

$$C=N_{\leq 2(n+1)}C_{\leq 2(n+1)} \tag{S77}$$

copies of ϱ_{p}, with $N_{\leq 2(n+1)}$ an integer given by equation (S69) and $C_{\leq 2(n+1)}$ an integer given by

$$C_{\leq 2(n+1)}\geq\frac{\sigma^2_{\leq 2(n+1)}(N_{\leq 2(n+1)}+1)}{\varepsilon^2\ln(1/(1-\alpha))}\left(n+5m/2\right)^2\left(1/2+2d\sqrt{2nm}\right)^{2n}. \tag{S78}$$

Then,

$$\mathbb{P}\left[|F^{(n)}-F^{(n)*}|\leq\varepsilon\right]\geq 1-\alpha. \tag{S79}$$

Proof. Our proof strategy is similar to that of Lemma S7. That is, we show that, with probability at least $1-\alpha$, the $N_{\leq 2(n+1)}$ relevant elements of $\mathbf{\Gamma}$ are estimated within additive errors bounded as

$$\varepsilon_{\max}\leq\varepsilon^*_{\max}:=\frac{\varepsilon}{\left(n+5m/2\right)\left(1/2+2d\sqrt{2nm}\right)^n}. \tag{S80}$$

If this inequality is fulfiled, then, due to Lemma S9, it holds that $|F^{(n)}-F^{(n)*}|\leq\varepsilon$.

According to Lemma S4, with the choice $\overline{\alpha}=1-\alpha$, taking

$$C_{\leq 2(n+1)}\geq\frac{\sigma^2_{\leq 2(n+1)}(N_{\leq 2(n+1)}+1)}{{\varepsilon^*_{\max}}^2\ln(1/(1-\alpha))}. \tag{S81}$$

is sufficient to get

$$\mathbb{P}\left[\forall\,\Gamma^{(i)}_{k_1,l_1,\ldots,k_i,l_i}\text{ relevant}:|\Gamma^{(i)*}_{k_1,l_1,\ldots,k_i,l_i}-\Gamma^{(i)}_{k_1,l_1,\ldots,k_i,l_i}|\leq\varepsilon^*_{\max}\right]\geq 1-\alpha \tag{S82}$$

Finally, inserting the definition (S80) of $\varepsilon^*_{\max}$ into equation (S81), we see that equation (S79) is equivalent to equation (S81). □

Now, we prove the theorem on quantum certification of linear-optical network states.

Proof of Theorem 3. The proof is analogous to the proof of Theorem 2, but with equation (S77), equation (11), $F^{(n)}$, $F^{(n)*}$, Lemma S10, and equation (S79) playing respectively the roles of equation (S55), equation (10) , $F^{(0)}$, $F^{(0)*}$, Lemma S7 and equation (S57). □

S3.E Proof of Corollary S1

The proof relies on three auxiliary lemmas equivalent to Lemmas S5, S6, and S7.

Lemma S11 (Sparsity of the locally post-selected Gaussian fidelity bound)**.** *$F_{\mathcal{S}}^{(0)}$ depends on at most $2m\kappa$ elements of $\mathbf{\Gamma}_{\mathcal{S}}^{(1)}$. We call these the* relevant elements *of $\mathbf{\Gamma}_{\mathcal{S}}^{(1)}$.*

Proof. The proof of the lemma is analogous to that of Lemma S5. □

Lemma S12 (Stability of the locally post-selected Gaussian fidelity bound)**.** *Let $F_{\mathcal{S}}^{(0)*}$ be defined by the same expression to $F^{(0)}$ in equation (S6) but divided by $\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})$ and with $\boldsymbol{\gamma}$ and $\mathbf{\Gamma}^{(1)}$ replaced by $\boldsymbol{\gamma}_{\mathcal{S}}^{*}$ and $\mathbf{\Gamma}_{\mathcal{S}}^{(1)*}$, and let $\epsilon_{\max} \coloneqq \|\boldsymbol{\gamma}_{\mathcal{S}} - \boldsymbol{\gamma}_{\mathcal{S}}^{*}\|_{\max}$ and $\varepsilon_{\max}^{(1)} \coloneqq \|\mathbf{\Gamma}_{\mathcal{S}}^{(1)} - \mathbf{\Gamma}_{\mathcal{S}}^{(1)*}\|_{\max}$. Then*

$$|F_{\mathcal{S}}^{(0)} - F_{\mathcal{S}}^{(0)*}| \leq \frac{2s_{\max}^2}{\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})} \left(\varepsilon_{\max}^{(1)}\sqrt{\kappa}m + \epsilon_{\max}\|\mathbf{x}\|_2\sqrt{2m}\right). \tag{S83}$$

Proof. The proof of the lemma is similar to that of Lemma S6, with the differences already explained in Section S2.C. □

Lemma S13 (Reliable estimation of the locally post-selected Gaussian fidelity bound)**.** *Let $\alpha \in (0, 1/2]$ and $\varepsilon > 0$. Let $F_{\mathcal{S}}^{(0)*}$ be defined by the same expression to $F_{\mathcal{S}}^{(0)}$ in equation (S6) but divided by $\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})$ and with $\boldsymbol{\gamma}$ and $\mathbf{\Gamma}^{(1)}$ replaced by $\boldsymbol{\gamma}_{\mathcal{S}}^{*}$ and $\mathbf{\Gamma}_{\mathcal{S}}^{(1)*}$, where $\boldsymbol{\gamma}_{\mathcal{S}}^{*}$ and $\mathbf{\Gamma}_{\mathcal{S}}^{(1)*}$ are obtained as described in Section S2.C from*

$$C = 2mC_1 + 2\kappa mC_2 \tag{S84}$$

copies of $\varrho_{\mathcal{S}_{\mathrm{p}}}$, with C_1 and C_2 integers such that

$$C_1 \geq 2^6 \frac{\varsigma_1^2(2m+1)\, m\, s_{\max}^4\, \|\mathbf{x}\|_2^2}{[\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})\varepsilon]^2 \ln\left(\frac{1}{1-\alpha}\right)} \tag{S85a}$$

and

$$C_2 \geq 2^5 \frac{\varsigma_2^2(2\kappa m+1)\, m^2\, s_{\max}^4\, \kappa}{[\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})\varepsilon]^2 \ln\left(\frac{1}{1-\alpha}\right)}. \tag{S85b}$$

Then,

$$\mathbb{P}\left[|F_{\mathcal{S}}^{(0)} - F_{\mathcal{S}}^{(0)*}| \leq \varepsilon\right] \geq 1 - \alpha. \tag{S86}$$

Proof. The proof of the lemma is analogous to that of Lemma S7. □

Proof of Corollary S1. The proof is analogous to the proof of Theorem 2 but with Lemmas S11, S12, and S13 playing respectively the roles of Lemmas S5, S6, and S7. □

S3.F Proof of Corollary S2

As in the previous subsection, the proof relies on three auxiliaryary lemmas equivalent to Lemmas S8, S9, and S10. The proofs of the lemmas are analogous to, and follow immediately from, those of the latter.

Lemma S14 (Sparsity of the locally post-selected linear-optical fidelity bound)**.** *The fidelity bound $F_{\mathcal{S}}^{(n)}$, defined by the same expression as $F^{(n)}$ in equation (S15) but divided by $\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})$ and with $\mathbf{\Gamma}$ replaced by $\mathbf{\Gamma}_{\mathcal{S}}$, can be written as*

$$F_{\mathcal{S}}^{(n)} = \frac{1}{\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})}\left[1 - \sum_{j=0}^{n}(-1/2)^{n-j} f_j\left(\mathbf{\Gamma}_{\mathcal{S}}^{(j)}, \mathbf{\Gamma}_{\mathcal{S}}^{(j+1)}\right)\right], \tag{S87}$$

where, for each $j \in \{0, \ldots, n\}$, f_j is the same linear functional as in Lemma S8, defined by equation (S68). Moreover, $F_{\mathcal{S}}^{(n)}$ depends on at most $N_{\leq 2(n+1)}$ elements of $\mathbf{\Gamma}_{\mathcal{S}}$, with $N_{\leq 2(n+1)}$ the same as in Lemma S8 and given by equation (S69).

Proof. The proof of the lemma is analogous to that of Lemma S8. □

Lemma S15 (Stability of the locally post-selected linear-optical fidelity bound)**.** *Let $F_{\mathcal{S}}^{(n)*}$ be defined by the same expression as $F^{(n)}$ in equation (S15) but divided by $\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})$ and with $\boldsymbol{\Gamma}$ replaced by $\boldsymbol{\Gamma}_{\mathcal{S}}^{*}$, and let $\varepsilon_{\max} \coloneqq \|\boldsymbol{\Gamma}_{\mathcal{S}} - \boldsymbol{\Gamma}_{\mathcal{S}}^{*}\|_{\max}$. Then*

$$|F_{\mathcal{S}}^{(n)} - F_{\mathcal{S}}^{(n)*}| \leq \frac{\varepsilon_{\max}}{\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})} (n + 5m/2) \left(1/2 + 2d\sqrt{2nm}\right)^{n}. \tag{S88}$$

Proof. The proof of the lemma is similar to that of Lemma S9, with the differences already explained in Section S2.C. □

Lemma S16 (Reliable estimation of the locally post-selected linear-optical fidelity bound)**.** *Let $\alpha \in (0, 1/2]$ and $\varepsilon > 0$. Let $F_{\mathcal{S}}^{(n)*}$ be defined like $F^{(n)}$ in equation (S15) but with $\boldsymbol{\Gamma}$ replaced by $\boldsymbol{\Gamma}_{\mathcal{S}}^{*}$, where $\boldsymbol{\Gamma}_{\mathcal{S}}^{*}$ is obtained as described in Section S2.C from*

$$C = N_{\leq 2(n+1)} C_{\leq 2(n+1)} \tag{S89}$$

copies of $\varrho_{\mathcal{S}_{\mathrm{p}}}$, with $N_{\leq 2(n+1)}$ an integer given by equation (S69) and $C_{\leq 2(n+1)}$ an integer given by

$$C_{\leq 2(n+1)} \geq \frac{\varsigma_{\leq 2(n+1)}^{2} (N_{\leq 2(n+1)} + 1)}{\left[\mathbb{P}(\boldsymbol{\phi}_{\mathcal{A}}|\varrho_{\mathrm{t}})\, \varepsilon\right]^{2} \ln(1/(1-\alpha))} (n + 5m/2)^{2} \left(1/2 + 2d\sqrt{2nm}\right)^{2n}. \tag{S90}$$

Then,

$$\mathbb{P}\left[|F_{\mathcal{S}}^{(n)} - F_{\mathcal{S}}^{(n)*}| \leq \varepsilon\right] \geq 1 - \alpha. \tag{S91}$$

Proof. The proof of the lemma is analogous to that of Lemma S10. □

Proof of Corollary S2. The proof is analogous to the proof of Theorem 3 but with Lemmas S14, S15, and S16 playing respectively the roles of Lemmas S8, S9, and S10. □

S3.G Proof of Theorem 5

Crucial for the proof of this theorem is the expansion (12) of ϱ_{p} in terms of ϱ_{t} and $\varrho_{\mathrm{t}}^{\perp}$, which leads to the definition (13) of the photon mismatch $\tilde{n}^{\perp}$. Also, before the proof, we note that the fidelity gap cannot be smaller than $\Delta \geq 2\varepsilon$: The condition for acceptance of the test is $F^{(n)*} \geq F_{\mathrm{T}} + \Delta - \varepsilon$, whereas that for rejection is $F^{(n)*} < F_{\mathrm{T}} + \varepsilon$. So, the threshold of acceptance, $F = F_{\mathrm{T}} + \Delta - \varepsilon$, is not smaller than that of rejection, $F = F_{\mathrm{T}} + \varepsilon$, iff $\Delta \geq 2\varepsilon$.

Proof of Theorem 5. Theorems 2 and 3 imply that ϱ_{p} is rejected with probability at least $1 - \alpha$ whenever $F < F_{\mathrm{T}}$. Thus, it remains to show that if ϱ_{p} is such that $F \geq F_{\mathrm{T}} + \Delta$, with Δ given by equation (14), then ϱ_{p} is accepted with probability at least $1 - \alpha$, i.e., that

$$\mathbb{P}\left[F^{(n)*} \geq F_{\mathrm{T}} + \varepsilon\right] \geq 1 - \alpha. \tag{S92}$$

So, let $F \geq F_{\mathrm{T}} + \Delta$, with Δ given by (14). Using equations (8), (12), and (13), we write $F^{(n)}$ as

$$F^{(n)} = 1 - (1 - F)\tilde{n}^{\perp} \geq 1 - (1 - (F_{\mathrm{T}} + \Delta))\tilde{n}^{\perp}. \tag{S93}$$

Using that

$$\Delta \geq \frac{2\varepsilon + (1 - F_{\mathrm{T}})(\tilde{n}^{\perp} - 1)}{\tilde{n}^{\perp}} \tag{S94}$$

and inserting it into the inequality (S93), we obtain

$$F^{(n)} \geq F_{\mathrm{T}} + 2\varepsilon. \tag{S95}$$

Finally, using equations (S95) and (S79), we obtain equation (S92). □

S4 Number of measurement settings

In this section, we upper-bound the number of local measurement settings required for the estimation of our fidelity lower bounds. We do this explicitly only for the Gaussian and linear-optical network target states, the cases of the post-selected target states following immediately from them.

S4.A Gaussian case

Here, we show that the $2md$ single-quadrature and the $m\kappa$ two-quadrature observables listed in Box S1, required for the measurement scheme $\mathcal{M}_{\mathrm{G}}$, can all be measured using $m+3$ different experimental arrangements. We do this by explicitly describing a measurement strategy that features such a scaling.

The two-body observables $\hat{q}_j\hat{q}_k$, $\hat{q}_j\hat{p}_k$, and $\hat{p}_j\hat{p}_k$, for $j \neq k$, can be measured by simultaneously homodyning modes j and k. For all possible pairs of modes, this consumes $m+2$ different homodyne settings: A single setting $(\hat{q}_1, \hat{q}_2, \ldots, \hat{q}_m)$ for all the second moments of the form $\langle \hat{q}_j\hat{q}_k \rangle_{\varrho_{\mathrm{p}}}$; another single setting $(\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_m)$ for those of the form $\langle \hat{p}_j\hat{p}_k \rangle_{\varrho_{\mathrm{p}}}$; and the m settings $(\hat{p}_1, \hat{q}_2, \ldots, \hat{q}_m)$, $(\hat{q}_1, \hat{p}_2, \hat{q}_3, \ldots, \hat{q}_m)$, …, and $(\hat{q}_1, \ldots, \hat{q}_{m-1}, \hat{p}_m)$ for those of the form $\langle \hat{q}_j\hat{p}_k \rangle_{\varrho_{\mathrm{p}}}$ and $\langle \hat{p}_j\hat{q}_k \rangle_{\varrho_{\mathrm{p}}}$ with $j \neq k$. In addition, all the single-body observables $\hat{q}_j$, $\hat{p}_j$, $\hat{q}_j^2$, and $\hat{p}_j^2$, are measured also with these same settings. With this, we have accounted, so far, for all the first moments γ_l and all the second moments $\Gamma^{(1)}_{l,l'}$ with $(l,l') \neq (2j-1, 2j)$ for all $j \in [m]$.

The remaining second moments, $\Gamma^{(1)}_{2j-1,2j}$ with $j \in [m]$, correspond to the single-mode observables $(\hat{q}_j\hat{p}_j + \hat{p}_j\hat{q}_j)/2$. To measure these, Arthur can homodyne each mode j independently in the rotated quadrature $(\hat{q}_j + \hat{p}_j)/\sqrt{2}$. This requires a single setting: $\left[(\hat{q}_1 + \hat{p}_1)/\sqrt{2}, (\hat{q}_2 + \hat{p}_2)/\sqrt{2}, \ldots, (\hat{q}_1 + \hat{p}_1)/\sqrt{2}\right]$. In this setting, he can estimate all the moments of the form $\langle (\hat{q}_j + \hat{p}_j)^2/2 \rangle_{\varrho_{\mathrm{p}}}$. The latter estimates, upon subtraction of $\langle \hat{q}_j^2 \rangle_{\varrho_{\mathrm{p}}}/2$ and $\langle \hat{p}_j^2 \rangle_{\varrho_{\mathrm{p}}}/2$, whose settings have already been accounted for, finally make it possible to calculate an estimate of $\langle (\hat{q}_j\hat{p}_j + \hat{p}_j\hat{q}_j)/2 \rangle_{\varrho_{\mathrm{p}}}$, using the equation

$$\frac{1}{2}(\hat{q}_j\hat{p}_j + \hat{p}_j\hat{q}_j) = \left(\frac{\hat{q}_j + \hat{p}_j}{\sqrt{2}}\right)^2 - \frac{\hat{q}_j^2}{2} - \frac{\hat{p}_j^2}{2}. \tag{S96}$$

The last setting, plus the $m+2$ ones already accounted for in the previous paragraph, yields a total of $m+3$ different homodyne settings, as promised.

Finally, a comment on the error estimation is in order. In any measurement strategy where moments are estimated indirectly, their errors must be obtained from those of the directly measured quantities via error propagation. For instance, in the strategy just described, the error of each $\Gamma^{(1)}_{2j-1,2j}$ needs to be calculated from those of $\langle (\hat{q}_j + \hat{p}_j)^2/2 \rangle_{\varrho_{\mathrm{p}}}$, $\langle \hat{q}_j^2 \rangle_{\varrho_{\mathrm{p}}}$, and $\langle \hat{p}_j^2 \rangle_{\varrho_{\mathrm{p}}}$. This leads, for each indirectly estimated moment, to an increase in the number of copies of ϱ_{p} required to attain a given error. Nevertheless, this usually has no impact on the leading terms of the total resource scaling of the protocol. For example, in the described strategy the global scaling given in equation (10) remains unaltered.

S4.B Linear-optical case

Here, we show that the $N_{\leq 2(n+1)} \in \mathrm{O}\left(m\left(4d^2+1\right)^n\right)$ observables listed in Box S2, required for the measurement scheme $\mathcal{M}_{\mathrm{LO}}$, can all be measured using at most $\binom{m}{n}2^{n+1}$ different experimental arrangements. As in the previous section, we do this by explicitly describing a measurement strategy that features the promised scaling.

The scheme $\mathcal{M}_{\mathrm{LO}}$ requires the measurement of products of an even number between 2 and $2(n+1)$ quadrature operators. We describe the measurement strategy as follows. First, we upper-bound the number of homodyne settings required for the measurement of all possible products of $2n$ quadrature operators, necessary for estimating all n-th moments $\mathbf{\Gamma}^{(n)}$. The measurement of products of fewer quadrature operators can clearly be carried out with the same settings. Then, we show that the particular products of $2(n+1)$ quadrature operators that appear in $\mathcal{M}_{\mathrm{LO}}$, corresponding to the relevant elements of $\mathbf{\Gamma}^{(n+1)}$, do not require extra settings either.

Consider all products of $2n$ quadrature operators. Among these, we focus first on those containing exclusively either $\hat{q}_j$ or $\hat{p}_j$ (or powers thereof) for each j-th mode but exclude observables such as $(\hat{q}_j\hat{p}_j + \hat{p}_j\hat{q}_j)/2$, which we address in the next paragraph. Let us divide this family into two subfamilies: (i) those for which the number of operators $\hat{q}_j$ is smaller or equal than that of the operators $\hat{p}_j$ and (ii) those for which the number of operators $\hat{q}_j$ is greater than that of the operators $\hat{p}_j$. All correlators in the subfamily (i) can be measured with homodyne settings where n modes are detected in the position quadrature $\hat{q}_j$ and the remaining $m-n$ ones in the momentum quadrature $\hat{p}_j$. All those in the subfamily (ii) can be measured with homodyne settings where n modes are detected in momentum and the remaining $m-n$ ones in position. Taking the two subfamilies into account, there are at most $2\binom{m}{n}$ different such settings.

Let us now focus on the products of $2n$ quadrature operators that include different quadrature operators on a same mode, such as $(\hat{q}_j\hat{p}_j + \hat{p}_j\hat{q}_j)/2$ (or powers thereof). At most, n factors as $(\hat{q}_j\hat{p}_j + \hat{p}_j\hat{q}_j)/2$ can appear in each product of $2n$ quadrature operators. From equation (S96), we know that by replacing in each of the settings for the subfamily (i) above a quadrature $\hat{q}_j$ with the rotated quadrature $(\hat{q}_j + \hat{p}_j)/\sqrt{2}$, Arthur can indirectly estimate the expectation values of all the $2n$-quadrature products of the form:

$$\begin{aligned}&\frac{1}{2}(\hat{q}_j\hat{p}_j + \hat{p}_j\hat{q}_j)\\ &\times \text{ up to } n-1 \text{ position operators}\\ &\times \text{ at least } n \text{ momentum operators.}\end{aligned} \tag{S97}$$

In turn, by replacing, in each of the resulting settings, a further quadrature $\hat{q}_{j'}$ with $(\hat{q}_{j'} + \hat{p}_{j'})/\sqrt{2}$, he can measure all the observables of the form

$$\begin{aligned}&\frac{1}{2}(\hat{q}_j\hat{p}_j + \hat{p}_j\hat{q}_j) \times \frac{1}{2}(\hat{q}_{j'}\hat{p}_{j'} + \hat{p}_{j'}\hat{q}_{j'})\\ &\times \text{ up to } n-2 \text{ position operators}\\ &\times \text{ at least } n \text{ momentum operators.}\end{aligned} \tag{S98}$$

Concatenating this procedure, he can measure all the $2n$-quadrature products where each mode contributes with either $\hat{q}_j\hat{p}_j + \hat{p}_j\hat{q}_j$, $\hat{q}_j$, or $\hat{p}_j$, and the number of operators $\hat{q}_j$ is smaller or equal than that of the operators $\hat{p}_j$. Equivalently, by proceeding analogously with the subfamily (ii) and the quadratures $\hat{p}_j$, he can measure all $2n$-quadrature products where each mode contributes with either $\hat{q}_j\hat{p}_j + \hat{p}_j\hat{q}_j$, $\hat{q}_j$, or $\hat{p}_j$, and the number of operators $\hat{q}_j$ is greater than that of the operators $\hat{p}_j$. This is enough to indirectly estimate the expectation values of all $2n$-quadrature products. For each setting of the two subfamilies, n modes can be rotated, giving rise to 2^n setting ramifications. Hence, taking into account all the settings of the two subfamilies and their ramifications, we count a total of at most $2\binom{m}{n}2^n = \binom{m}{n}2^{n+1}$ different settings. This counting clearly over-counts the necessary settings but is enough for our purposes.

Finally, we consider the products of $2(n+1)$ quadrature operators appearing in the relevant elements of $\mathbf{\Gamma}^{(n+1)}$. The $(n+1)$-th moment tensor $\mathbf{\Gamma}^{(n+1)}$ is special in that, in contrast to the lower-moment tensors, it appears in just the first of the two traces in equations (S15) and (S68). In particular, according to Box S2, $\Gamma^{(n+1)}_{k_1,l_1,\ldots,k_n,l_n,k_{n+1},l_{n+1}}$ is a relevant element of $\mathbf{\Gamma}^{(n+1)}$ if, and only if, $k_{n+1} = l_{n+1}$. This implies that the observables containing the factor $(\hat{q}_{n+1}\hat{p}_{n+1} + \hat{p}_{n+1}\hat{q}_{n+1})$ do not contribute to the relevant elements of $\mathbf{\Gamma}^{(n+1)}$, only those containing either $\hat{q}^2_{n+1}$ or $\hat{p}^2_{n+1}$ are relevant. Hence, the relevant $2(n+1)$ quadrature products are those composed of the $2n$ quadrature products relevant for $\mathbf{\Gamma}^{(n)}$ times either $\hat{q}^2_{n+1}$ or $\hat{p}^2_{n+1}$. Now, in each setting of the two subfamilies of the previous paragraph, $2n$ modes are used to measure a $2n$-quadrature observable relevant for $\mathbf{\Gamma}^{(n)}$ and the other $m-n$ modes, which are all set either to position or momentum, are ignored. Thus, each relevant element of $\mathbf{\Gamma}^{(n+1)}$ can be estimated by not ignoring one out of the latter $m-n$ modes. That is, the settings to estimate the $2n$-moments $\mathbf{\Gamma}^{(n)}$ already cover also the estimation of $2(n+1)$-moments $\mathbf{\Gamma}^{(n+1)}$. So, the total number of settings used throughout is at most $\binom{m}{n}2^{n+1}$.

As in the end of the previous section, we make a final remark on the error estimation. Also here, the errors of the indirectly estimated moments must be obtained via error propagation, which leads again to an increase in the total number of copies of ϱ_{p}. Nevertheless, their global scaling with n remains of the same order as that given in equation (11) .

S5 Stability against systematic errors

Apart from statistical errors, Arthur's measurement procedure could also have systematic errors. That is, if the characterisation of his single-mode measurement channels is erroneous, he could actually be measuring different observables from the ones he thinks he does. Theorems 2 and 3, as well as their Corollaries S1 and S2, consider only statistical errors, i.e., those that can be decreased by increasing the number of measurement repetitions (and, hence, the number of copies of ϱ_{p}). Since systematic errors cannot be decreased by accumulating statistics, no certification method based exclusively on the measurement statistics can rule them out. However, the stability analyses of Lemmas S6, S9, S12 and S15 hold regardless of the nature of errors. Thus, the experimental estimates $F^{(0)*}$, $F^{(n)*}$, $F^{(0)*}_{\mathcal{S}}$, and $F^{(n)*}_{\mathcal{S}}$ (and, therefore, also the certification tests) turn out to be robust also against small systematic errors: The total fidelity deviations due to systematic errors scales linearly with the magnitude of the largest systematic error and polynomially in all the other relevant parameters as given in equations (S48), (S71), (S83), and (S88).

Still, it is illustrative to consider a physically relevant example. A typical systematic error is non-unit quantum efficiency of the detectors used for homodyning. In that case, the probability density function $\tilde{\mathcal{P}}$ of measurement outcomes r of a quadrature $\hat{r}$ equals the ideal one $\mathcal{P}$ convolutioned with the normal distribution $\mathcal{N}$ of mean zero and squared variance $(1-\eta)/4\eta$, where η

is the quantum efficiency of the detectors[9]. That is, $\tilde{\mathcal{P}}(r) = (\mathcal{P} * \mathcal{N})(r) \coloneqq \int dr' \mathcal{P}(r')\mathcal{N}(r - r')$. Using that the first and second *non-central moments* of $\mathcal{N}$ satisfy

$$\langle r \rangle_{\mathcal{N}} \coloneqq \int dr r \mathcal{N}(r - r') = r' \tag{S99a}$$

and

$$\langle r^2 \rangle_{\mathcal{N}} \coloneqq \int dr r^2 \mathcal{N}(r - r') = r'^2 + \frac{1-\eta}{4\eta}, \tag{S99b}$$

respectively, one obtains that

$$\langle r \rangle_{\tilde{\mathcal{P}}} = \langle r \rangle_{\mathcal{P}} \tag{S100a}$$

and

$$\langle r^2 \rangle_{\tilde{\mathcal{P}}} = \langle r^2 \rangle_{\mathcal{P}} + \frac{1-\eta}{4\eta}. \tag{S100b}$$

That is, the expectation value of $\hat{r}$ is not affected by this type of systematic errors and that of $\hat{r}^2$ deviates from the ideal one by $(1-\eta)/(4\eta)$. Furthermore, the expectation values of products of quadrature operators acting on different modes are also not affected, as this type of systematic error acts independently on different modes.

In the absence of statistical errors, this leads to an error vector $\boldsymbol{\epsilon} = \mathbf{0}$ and an error matrix $\mathcal{E}^{(1)}$ that is diagonal and such that $\|\mathcal{E}^{(1)}\|_{\max} \leq (1-\eta)/(4\eta)$, so that $\|\mathcal{E}^{(1)}\|_1 \leq m(1-\eta)/(2\eta)$. Inserting this into equation (S52), we see for instance that, for Gaussian targets, the contribution to the deviation of the fidelity estimate due to non-ideal detector efficiency in the homodyne detectors is smaller than $s_{\max}^2 m \frac{1-\eta}{2\eta}$. This, in turn, is smaller or equal than a desired constant maximal error ε if

$$\eta \geq \frac{s_{\max}^2 m}{2\varepsilon + s_{\max}^2 m} \approx 1 - \frac{2\varepsilon}{s_{\max}^2 m}, \tag{S101}$$

where the approximation holds whenever $s_{\max}^2 m \gg 2\varepsilon$. The scaling given by the bound (S101) is experimentally convenient in that, in particular, it implies that the detector inefficiency $1-\eta$ needs to decrease only inversely proportional with the number of modes m.

Another typical systematic error is the limited power of the local oscillator field used for the homodyne detection: The homodyne (photocurrent difference) statistics, i.e., the distribution of homodyne measurement outcomes, match exactly the statistics of the corresponding quadrature only in the limit of an intense local-oscillator beam[10]. The most obvious difference is that the homodyne statistics is discrete whereas the quadrature statistics is continuous, with the former approximating the latter increasingly better as the local-oscillator power increases. However, we emphasise that our method relies on the estimation of only the expectation values of quadratures and not their full statistics. It can be seen that, provided that the local oscillator is in a coherent state, the effect of limited power is just to increase the variance of the effective quadrature without changing its expectation value with respect to the ideal case. Furthermore, in the multi-mode scenario, if the different modes are homodyned with independent local oscillators, the latter is also true for products of quadratures, as the ones considered in this work. Therefore, the effect of systematic errors due to limited homodyne local-oscillator power in our fidelity estimates is expected not to be critical either.

S6 Auxiliary mathematical relations

S6.A Derivation of the properties of the operator valued Pochhammer-Symbol

We begin with equation (21a). The general relationship

$$(a_j^\dagger)^t \hat{n}_j (a_j)^t = p_t(\hat{n}_j), \tag{S102}$$

for $t \in \mathbb{N}$, can be shown by induction starting from $p_0(\hat{n}_j) = \hat{n}_j$ and noting that, for all $t \geq -1$,

$$\begin{aligned} a_j^\dagger p_t(\hat{n}) a_j &= a_j^\dagger \hat{n}_j(\hat{n}_j - 1)(\hat{n}_j - 2)\cdots(\hat{n}_j - t)a_j \\ &= a_j^\dagger \hat{n}_j(\hat{n}_j - 1)(\hat{n}_j - 2)\cdots(\hat{n}_j - (t-1))a_j(\hat{n}_j - (t+1)) \\ &= p_t(\hat{n}_j)(\hat{n}_j - (t+1)) \\ &= p_{t+1}(\hat{n}_j), \end{aligned} \tag{S103}$$

as can be verified using the commutation relations between a_j and $a_j^\dagger$. Setting $t = n_j$ gives equation (21a) .

In turn, equation (21b) can be shown by noting that

$$(a_j^\dagger)^{n_j}(a_j)^{n_j} = (a_j^\dagger)^{n_j - 1}\hat{n}_j(a_j)^{n_j - 1} \tag{S104}$$

and applying equation (S102), for $t = n_j - 1$, to the right-hand side of (S104).

S6.B Proof of the bound (S43)

Note that for $x = 0$ both sides of equation (S43) yield 1 and hence the bound holds in that case. We make the substitution $y = 1/x$ and show that the bound (S43) holds for all $x > 0$ by proving the following:

$$\frac{1}{1 - \mathrm{e}^{-y}} \leq \frac{1}{y} + \frac{1}{2(1 + 1/y)} + \frac{1}{2} \quad \forall y \geq 0. \tag{S105}$$

But this is equivalent to

$$2y^2 + 3y + 2 \leq \mathrm{e}^y(2 + y). \tag{S106}$$

A straight forward calculation shows that both sides and also the first derivatives of both sides coincide at $y = 0$, while the second derivative of the right hand side is always larger than the second derivative of the left hand side. This proves equation (S105) and hence finishes the proof of the bound (S43).

---

[1] F. Dell'Anno, D. Buono, G. Nocerino, A. Porzio, S. Solimeno, S. De Siena, and F. Illuminati, *Tunable non-Gaussian resources for continuous-variable quantum technologies*, Phys. Rev. A **88**, 043818 (2013).

[2] F. Dell'Anno, S. De Siena, L. Albano, and F. Illuminati, *Continuous-variable quantum teleportation with non-Gaussian resources*, Phys. Rev. A **76**, 022301 (2007).

[3] C. Navarrete-Benlloch, R. García-Patrón, J. H. Shapiro, and N. J. Cerf, *Enhancing quantum entanglement by photon addition and subtraction*, Phys. Rev. A **86**, 012328 (2012).

[4] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn *Linear optical quantum computing with photonic qubits*, Rev. Mod. Phys. **79**, 135 (2007).

[5] E. Knill, R. Laflamme, and G. J. Milburn, *A scheme for efficient quantum computation with linear optics*, Nature **409**, 46 (2001).

[6] C. Weedbrook, S. Pirandola, R. Garcia-Patron, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, *Gaussian quantum information*, Rev. Mod. Phys. **84**, 621 (2012).

[7] S. L. Braunstein, *Squeezing as an irreducible resource*, Phys. Rev. A **71**, 055801 (2005).

[8] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, *Direct characterization of linear-optical networks*, Phys. Rev. Lett. **73**, 58 (1993).

[9] A. Ferraro, S. Olivares, and M. G. A. Paris, *Gaussian states in continuous variable quantum information*, ISBN 88-7088-483-X (Bibliopolis, Napoli, 2005), arXiv: quant-ph/0503237.

[10] S. L. Braunstein, *Homodyne statistics*, Phys. Rev. A **42**, 474 (1990).

C Back matter

C.1 Acknowledgements

First of all, I would like to say that I enjoyed very much being a Ph.D. student in the research group of Jens Eisert. In particular, I very much thank him for his enthusiasm, support, advice, and trust. I also thank all other current and former members of the group who all contributed to the friendly atmosphere and made the group a perfect environment for studying and researching. In particular, I would like to thank Christian Gogolin, Arnau Riera, Mathis Friesdorf, Matthias Ohliger, Albert Werner, Leandro Aolita, Christian Krumnow, and Henrik Wilming, as well as all my co-authors for fruitful and enjoyable discussions.

I also thank my family for their encouragement, help, and welcome distractions.

I acknowledge funding from the Studienstiftung des Deutschen Volkes.

C.2 Abstract

Simulations play a crucial role in the investigation of complex quantum systems. In this thesis, locality structures of quantum systems are exploited to obtain complexity theoretic results with various physical implications. More specifically, rigorous mathematical tools are used and developed further, to investigate open quantum systems and thermal states. Moreover, important advances in photonic quantum simulations are discussed.

For open quantum spin lattice systems new simulation schemes are provided. It is shown that Markovian dynamics can be simulated efficiently in the unitary circuit model, which can be seen as a dissipative Church-Turing type theorem. Moreover, Markovian dynamics is quasi-local and can locally be simulated on classical computers with a cost scaling polynomially in the system size. These results generalize standard tools from the investigation of Hamiltonian systems to open quantum systems. In particular, they provide a rigorous basis for their numerical simulation. However, also a major roadblock for making such simulations reliable is identified: Testing positivity of certain common approximations to mixed quantum states, called matrix product operators, is shown to be NP-hard in the system size and undecidable in the thermodynamic limit. Also more state space structures, originating from the spatial locality structure are discussed: Most states in state space cannot be generated efficiently with local Liouvillian dynamics and pure states generated in real-space renormalization schemes turn out to have local corrections in spatial dimensions larger than one.

For thermal states on spin and fermionic lattice systems, a perturbation formula is provided and exponential clustering of correlations at high enough temperature is proven. This has various consequences: It leads to the extension of the concept of intensive temperature to interacting quantum systems, allows for efficient classical local simulations at high enough temperature, provides an upper bound on phase transition temperatures, and implies stability of thermal states against Hamiltonian perturbations.

For photonic quantum simulations, sample complexity lower bounds for the verification of Boson-Sampling simulations are explained, which are applicable to a restricted setting. These bounds rely on a lower bound on the min-entropy of the output distribution of Boson-Sampling. This indicates that Boson-Sampling cannot be verified efficiently classically. Complementary to that, a reliable verification scheme for photonic state preparations is discussed, which uses single mode measurements as a simple quantum resource. The verification scheme is efficient for a large class of photonic simulations, including Boson-Sampling experiments with constantly many photons and state preparations necessary for measurement based quantum computing.

C.3 Zusammenfassung

Simulationen spielen eine wichtige Rolle in der Untersuchung von komplexen Quantensystemen. In dieser Arbeit werden unter Ausnutzung von Lokalitätsstrukturen koplexitätstheoretische Ergebnisse erzielt, die vielseitige Implikationen haben. Aufbauend auf rigorosen mathematischen Methoden und deren Weiterentwicklung werden offene Quantensysteme, thermische Zustände und photonische Quantensimulationen untersucht.

Es werden neue Simulationsmethoden für offene Spingittersysteme eingeführt. Für Markovsche Zeitentwicklung von Spingittersystemen wird gezeigt, dass sie effizient im unitären Quantengattermodell simuliert werden können, was als eine Church-Turing-Aussage interpretiert werden kann. Außerdem ist Markovsche Dynamik quasilokal und mit klassischen Computern lokal simulierbar, wobei der Simulationsaufwand polynomiell in der Systemgröße beschränkt ist. Diese Ergebnisse bilden eine rigorose Grundlage für die Simulation von offenen Quantensystemen. Allerdings wird auch ein großes Hindernis für die Verbesserung der Zuverlässigkeit solcher Simulationen identifiziert: Das Testen von Positivität von bestimmten üblichen Approximationen an gemischte Quantenzustände, die Matrixproduktoperatoren genannt werden, ist NP-hart in der Systemgröße und im thermodynamischen Limes unentscheidbar. Zusätzlich werden auch weitere Zustandsraumstrukturen, welche durch räumliche Lokalität bedingt sind, diskutiert: Die allermeisten Quantenzustände können nicht effizient mittels lokaler Liouvillscher Dynamik generiert werden.

Für thermische Zustände auf Spin- und Fermionischen Gittersystemen wird eine Störungsformel eingeführt und exponentielles Clustering von Korrelationen bei ausreichend hohen Temperaturen bewiesen. Dies hat eine Reihe von Konsequenzen: Es führt zu einer Erweiterung des Konzepts der Intensivität von Temperatur auf wechselwirkende Quantensysteme, zu effizienteren klassischen Simulationen von Systemen bei ausreichend hoher Temperatur, es bietet eine obere Schranke an Phasenübergangstemperaturen und impliziert Stabilität von thermischen Zuständen gegenüber Störungen des Hamiltonoperators.

Für photonische Quantensimulationen werden untere Schranken an die Sample-Komplexität der Verifizierung von Boson-Sampling-Simulationen in einem eingeschränkten Rahmen erklärt. Eine untere Schranke an die min-Entropie der Ausgabeverteilung von Boson-Sampling stellt das technische Hauptresultat dieser Untersuchung dar. Diese deutet darauf hin, dass Boson-Sampling klassisch nicht effizient verifiziert werden kann. Ergänzend zu diesem Resultat wird ein zuverlässiges Verifizierungsschema für photonische Zustandspräparationen vorgestellt, in welchem Einzelmodenmessungen als Quantenressource verwendet werden. Dieses Verifizierungsschema ist für eine große Klasse von photonischen Simulationen effizient, die Boson-Sampling-Simulationen mit konstanter Photonenzahl und bestimmten Zustandspräparation, die für messbasiertes Quantenrechnen notwendig sind, mit einschließt.

C.4 Liste der Publikationen des Verfassers

Diese kumulative Dissertation basiert auf den folgenden Erstautorenschaften:

[Kli+11a] s. S. 32ff — M. Kliesch, T. Barthel, C. Gogolin, M. Kastoryano und J. Eisert, "*Dissipative Quantum Church-Turing Theorem*", Phys. Rev. Lett. **107**, 120501 (2011), DOI: 10.1103/PhysRevLett.107.120501

[KGE14b] s. S. 13ff — M. Kliesch, C. Gogolin und J. Eisert, "*Lieb-Robinson Bounds and the Simulation of Time-Evolution of Local Observables in Lattice Systems*", in: Many-Electron Approaches in Physics, Chemistry and Mathematics, hrsg. von V. Bach und L. Delle Site, Mathematical Physics Studies, Springer International Publishing, 2014, S. 301–318, DOI: 10.1007/978-3-319-06379-9_17, arXiv:1306.0716

[Kli+14] s. S. 47ff — M. Kliesch, C. Gogolin, M. J. Kastoryano, A. Riera und J. Eisert, "*Locality of Temperature*", Phys. Rev. X **4**, 031019 (2014), DOI: 10.1103/PhysRevX.4.031019

[KGE14a] s. S. 69ff — M. Kliesch, D. Gross und J. Eisert, "*Matrix-Product Operators and States: NP-Hardness and Undecidability*", Phys. Rev. Lett. **113**, 160503 (2014), DOI: 10.1103/PhysRevLett.113.160503

Aus der Promotion sind die folgende weitere Veröffentlichungen hervorgegangen:

[BKE10] s. S. 101ff — T. Barthel, M. Kliesch und J. Eisert, "*Real-Space Renormalization Yields Finite Correlations*", Phys. Rev. Lett. **105**, 010502 (2010), DOI: 10.1103/PhysRevLett.105.010502

[BK12] s. S. 108ff — T. Barthel und M. Kliesch, "*Quasilocality and Efficient Simulation of Markovian Quantum Dynamics*", Phys. Rev. Lett. **108**, 230504 (2012), DOI: 10.1103/PhysRevLett.108.230504

[GKAE13] s. S. 117ff — C. Gogolin, M. Kliesch, L. Aolita und J. Eisert, "*Boson-Sampling in the light of sample complexity*", (2013), arXiv:quant-ph/1306.3995

[AGKE14] s. S. 139ff — L. Aolita, C. Gogolin, M. Kliesch und J. Eisert, "*Quantum certification of photonic quantum simulations*", submitted to Nat. Comm. (2014), arXiv:quant-ph/1407.4817

C.5 Anteil des Autors bei Konzeption, Durchführung und Berichtsabfassung

Im Folgenden wird für jede der obigen Publikationen der Anteil des Autors bei Konzeption, Durchführung und Berichtsabfassung aufgezählt.

[Kli+11a] Der Autor war federführend bei diesem Projekt und hat insbesondere wichtige Teile der im Haupttext und Supplemental Material präsentierten Beweise entwickelt und ausformuliert.

[KGE14b] Der Autor war federführend und hat wichtige Beiträge zur Konzeption dieses Buchkapitels beigetragen. Substantielle Teile des Textes sowie die Abbildungen wurden vom Autor erstellt.

[Kli+14] Der Autor war federführend bei diesem Projekt und hat wichtige Teile der im Haupttext und Anhang präsentierten Beweise entwickelt und ausformuliert. Erhebliche Teile des gesamten Textes und der Abbildungen stammen ebenfalls vom Autor.

[KGE14a] Der Autor war federführend bei diesem Projekt. Er hat klassische Methoden für Hidden Markov Modelle auf den Quantenfall angepasst und die entsprechenden technischen Abschnitte mit Definitionen, Theoremen und Beweisen ausformuliert. Die Abbildungen wurden ebenfalls von dem Autor entworfen und erstellt. Außerdem hat er wichtige Teile der Einleitung sowie vom Review-Teil im Haupttext sowie im Anhang ausformuliert.

[BKE10] Der Autor hat einen großen Anteil zur Entwicklung der Notation, der Formulierung des technischen Rahmens, der Beweisführung, der “refinements” und Beispiele im Supplemental Material und ebenfalls zu deren Niederschriften beigetragen. Ein zentraler Bestandteil dieser Veröffentlichung sind die Abbildungen, welche zu großen Teilen vom Autor stammen.

[BK12] Der Autor hat wichtige Beiträge zur Konzeption der Arbeit und zu der Entwicklung der Beweise beigetragen und war auch an der Ausformulierung des Haupttexts sowie des Supplemental Materials beteiligt.

[GKAE13] Der Autor hat wichtige Beiträge zur Ausarbeitung der Problemstellung, welche einen wichtigen Bestandteil der Arbeit bildet, sowie zu der Konzeption der Arbeit beigetragen. Außerdem hat der Autor bei Entwicklung und Ausformulierung der Beweise einen wichtigen Anteil geleistet und auch bei der Ausformulierung der gesamten Arbeit geholfen.

[AGKE14] Der Autor hat zu der Konzeption der Arbeit, zu der Ausarbeitung der Theoreme und deren Beweise, zur Erstellung der Abbildungen und zur Auformulierung der Arbeit wichtige Beiträge geleistet.

C.6 Eigenständigkeitserklärung

Hiermit bestätige ich, dass ich die vorliegende Arbeit selbstständig und nur mit Hilfe der angegebenen Hilfsmittel angefertigt habe. Alle Stellen der Arbeit, die wörtlich oder sinngemäß aus Veröffentlichungen oder aus anderweitigen fremden Quellen entnommen wurden, sind als solche kenntlich gemacht. Ich habe die Arbeit noch nicht in einem früheren Promotionsverfahren eingereicht.

Martin Kliesch