Universität Potsdam

Mathematisch-Naturwissenschaftliche Fakultät

Institut für Physik und Astronomie

Characterizing and Measuring Properties of Continuous-Variable Quantum States

Kumulative Dissertation

Kandidat: **Matthias Ohliger**

Hauptbetreuer: **Prof. Jens Eisert**

Potsdam, Juni 2012

Abstract

We investigate properties of quantum mechanical systems in the light of quantum information theory. We put an emphasize on systems with infinite-dimensional Hilbert spaces, so-called "continuous-variable systems", which are needed to describe quantum optics beyond the single photon regime and other Bosonic quantum systems. We present methods to obtain a description of such systems from a series of measurements in an efficient manner and demonstrate the performance in realistic situations by means of numerical simulations. We consider both unconditional quantum state tomography, which is applicable to arbitrary systems, and tomography of matrix product states. The latter allows for the tomography of many-body systems because the necessary number of measurements scales merely polynomially with the particle number, compared to an exponential scaling in the generic case. We also present a method to realize such a tomography scheme for a system of ultra-cold atoms in optical lattices.

Furthermore, we discuss in detail the possibilities and limitations of using continuous-variable systems for measurement-based quantum computing. We will see that the distinction between Gaussian and non-Gaussian quantum states and measurements plays a crucial role. We also provide an algorithm to solve the large and interesting class of naturally occurring Hamiltonians, namely frustration free ones, efficiently and use this insight to obtain a simple approximation method for slightly frustrated systems. To achieve this goals, we make use of, among various other techniques, the well developed theory of matrix product states, tensor networks, semi-definite programming, and matrix analysis.

Contents

List of publications

- **Limitations of quantum computing with Gaussian cluster states**,
  M. Ohliger, K. Kieling, and J. Eisert,
  Phys. Rev. A **82**, 042336 (2010).

- **Solving frustration-free spin systems**,
  N. de Beaudrap, M. Ohliger, T. J. Osborne, and J. Eisert,
  Phys. Rev. Lett. **105**, 060504 (2010).

- **Continuous-variable quantum compressed sensing**,
  M. Ohliger, V. Nesme, D. Gross, Y.-K. Liu, and J. Eisert,
  arXiv:1111.0853.

- **Efficient measurement-based quantum computing with continuous-variable systems**,
  M. Ohliger and J. Eisert,
  Phys. Rev. A **85**, 062318 (2012).

- **Efficient and feasible state tomography of quantum many-body systems**,
  M. Ohliger, V. Nesme and J. Eisert, arXiv:1204.5735.

Introduction

The advent of quantum information science brought a new way of looking at quantum mechanics. A property of particular importance is entanglement. In a nutshell, a state is entangled if it posses correlations which are not possible within the rules of classical physics. Entanglement was first viewed as a strange and rather paradox feature of quantum mechanics [1]. In his seminal paper, Bell introduced an inequality, later named after him, which sets a limit to the correlations between two systems possible in local, realistic theories [2]. Quantum mechanics, which is non-local, on the other hand, predicts a violation of Bell's inequality. Quantum-optical systems were the first to allow for an experimental test of this fundamental inequality and to confirm the prediction of quantum mechanics [3]. Even though, over the course of the years, many loop-holes in the experiments have been closed, Bell's and related inequalities remain of crucial importance in the investigation of the foundations of quantum mechanics.

Quantum effects as resources for information processing

In more recent years, the focus of attention has shifted, and entanglement is not anymore seen as a mere curiosity but as a resource for protocols of information processing. One central aim of the field of quantum information theory is to gain insight into the nature of such resource and to learn how to use them. One of the most remarkable things made possible by quantum mechanics is the quantum computer. By making use of the specific features of quantum mechanics, namely superpositions of states and the already mentioned entanglement, it is possible to perform computational tasks efficiently for which no efficient algorithm running on a classical computer exists. The most important problem in this class is the task of decomposing numbers into their prime factors for which Shor provided his famous algorithm [4]. Effi-

ciency here means that the time required for the computation scales at most polynomially in the number of digits. The ability to factorize large numbers would bring oneself into the position of breaking the public-key encryption ubiquitously used on the internet. However, yet another class of quantum protocols can be used to perform distribution of secret keys which is secure even against attackers which have access to a quantum computer. Quantum cryptography using light as a carrier of information has been demonstrated over large distances and has been even used in real world applications [5].

Measurement-based quantum computing

In the most common paradigm, often called the circuit model, the working of a quantum computer is rather similar to the functioning of a classical one: First, the register, consisting of a number of qubits, is initialized in some standard state. Second, the computation is performed by applying a section of single-qubit gates and two-qubit gates, the quantum analogue of gates like NOT and XOR used in classical computers. Finally, the qubits of the register are measured. The implementation of the gates requires very precise control of the qubits and is, therefore, very challenging. However, the technique of measurement-based quantum computing (MBQC) as pioneered by Briegel and Raussendorff [6] and further developed by Gross et al. [7,8], allows to circumvent the necessity of performing gates at all. Instead, the computation is achieved by first preparing a suitable entangled quantum state, the resource, followed by a sequence of single-qubit measurements in various bases. The resource state is universal, i.e., it does not depend on the algorithm one wants to perform. Only the measurement sequence and bases are determined by the desired algorithm and input. Therefore, the difficult step, i.e., the preparation of the resource state can be performed off-line while the actual computation requires only presumably easier single qubit measurements. In this thesis, we investigate, among other things, what states are resources for MBQC and what kind of measurements on those states are needed to perform universal quantum computing. Not completely unexpectedly, there is a "conservation of difficulties" at work: If one requires the resource to belong to some class of states which are easy to prepare, one needs more complicated measurements while if one allows for more elaborate resource states, simpler measurements suffice.

Quantum state tomography

If one wants to use a quantum state to perform some task of information processing as quantum computing or quantum key distribution, it is highly desirable to know what state is prepared by the actual experimental apparatus. To full extend, this question is solved by quantum state tomography which gives a complete description in the most general setting without any prior assumption at the expense of needing the measurement of a huge number of observables and a considerable amount of post-processing. The recently developed method of compressed sensing (CS) notably reduce the required number of measurements in the important situations where the state are of low rank or belong to some particular class like ground states of local Hamiltonians [9, 10, 12]. We improve those techniques substantially such that they can be applied in a much larger number situations including quantum optical systems and ultra-cold atoms in optical lattices.

Physical realizations

Such optical lattices are periodic structures which are produced by superimposing counter-propagating laser beams in such a fashion that standing waves ares formed. Ultra-cold atoms placed in such structures can be used to study quantum many body physics in a much cleaner and more controllable way than possible in solid-state systems. They allow to realize quantum phase transitions and can function as a quantum simulator of other systems which are inaccessible to direct experimentation [11, 13]. We show how to use experimentally feasible techniques to perform full quantum state tomography of ultra-cold Bosons in optical lattices by employing our above mentioned newly developed theory of compressed sensing. To this aim, we also have a closer look at the commonly used single band Bose-Hubbard model and investigate under which conditions it constitutes a valid approximation.

Even though optical lattices have the potential to be used as a scalable quantum computer, the systems used in most of the experiments demonstrating quantum information processing so far are optical systems. This is mainly due to the very weak interaction of photons with the environment keeping the detrimental effects of decoherence low. Continuous-variable states, i.e., quantum states beyond the single photon regime, have desirable features concerning production and detection but their theoretical treatment is slightly hindered because the respective Hilbert spaces are infinite dimensional. We

provide methods to perform efficient tomography on such systems, to asses their applicability as resources for MBQC, and to quantify their non-classical nature.

The dimension of the Hilbert space of a quantum many-body systems, for reasons of clarity to be assumed to consist of spins, increases exponentially with the number of particles which is the reason that their straightforward treatment by means of exact diagonalization is limited to very small systems. The quantum states occurring in nature, on the other hand, are often described by only polynomially many parameters. For example, the ground states of one-dimensional systems containing only local interactions are, generically, described by matrix product states with slowly growing bond dimension [14], a fact we use in Chapter 6 to provide a tomography scheme which requires only polynomially (in the system size) many measurements. In most situations, this description even efficient to find. For higher-dimensional systems, the situation is less clear. However, if a nearest-neighbor Hamiltonian is frustration free, i.e., its ground state is also the ground state of all of the contributing interaction terms, it can be efficiently calculated. In Chapter 7, we provide a method to do this and show that it can be also used as an approximative method for systems where the frustration is merely weak.

Structure of the thesis

This cumulative thesis is organized as follows. In the first two chapters, we give an introduction to the main topics of this work. We explain the experimental and theoretical aspects of continuous-variable quantum optics and the physics of ultra-cold atoms in optical lattices. We also provide an introduction to the paradigm of measurement-based quantum computing (MBQC) including the derivation of a formalism which allows for the description of almost all MBQC protocols. After this, we shortly discuss quantum state tomography and give an overview over various methods while focusing especially on compressed sensing.

The main part of the thesis is formed by five articles which all are either already printed in or submitted to peer reviewed journals. We also include two appendices of unpublished material, one introducing and discussing various methods to quantify the non-Gaussianity and non-classicality of continuous-variable quantum states and a second one investigating the validity of the single-band Bose-Hubbard model which is used to describe ultra-cold Bosons in optical lattices. We end with a summary and a conclusion of the various

papers and provide some closing remarks concerning this works as a whole as well as a short outlook.

Chapter 1

Quantum systems for information processing

1.1 Continuous-variable quantum states

The building blocks of classical computers are bits which can be either in the state "0" or "1". The most straightforward generalization to the quantum case are systems with two-dimensional Hilbert space $\mathcal{H} = \mathbb{C}^2$ which can be realized, for example, by the polarization degree of freedom of a single photon. As single photons are difficult to produce, detect, and manipulate, other classes like Gaussian states have gained interest.

A single light mode is described by a harmonic oscillator with Hilbert space $\mathcal{H} = \mathbb{C}^\infty$ and energy eigenstates, also called Fock states, $|n\rangle$ for $n = 0, 1, \ldots$. The photon creation and annihilation operators $\hat{a}^\dagger$ and $\hat{a}$ act as $\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n\rangle$ and $\hat{a}|n\rangle = \sqrt{n}|n-1\rangle$ while the number operator $\hat{n} = \hat{a}^\dagger\hat{a}$ is diagonal in the Fock basis, i.e., $\hat{n}|n\rangle = n|n\rangle$. With their help, we can define the Hermitian position and momentum operators, sometimes also called quadrature operators,

$$\hat{x} = \frac{1}{\sqrt{2}}(\hat{a} + \hat{a}^\dagger), \quad \hat{p} = \frac{-i}{\sqrt{2}}(\hat{a} - \hat{a}^\dagger), \tag{1.1}$$

whose expectation values can be determined by the technique of homodyne detection which is the quantum optical measurement possessing the highest achievable accuracy [15].

1.1.1   Wigner function

A single-mode quantum state $\rho \in \mathcal{B}(\mathcal{H})$ can be represented by the real Wigner function depending on two real variables [16]:

$$W_\rho(x,p) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}\xi \exp(-ip\xi)\langle q + \tfrac{1}{2}\xi|\rho|q - \tfrac{1}{2}\xi\rangle \tag{1.2}$$

which fulfills $\int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}p\, W(x,p) = 1$. The probability densities obtained when measuring position or momentum operators (1.1) are given by the marginals of the Wigner function, i.e.,

$$P_{\hat{x}}(\xi) = \int_{-\infty}^{\infty} \mathrm{d}p\, W(x,p), \tag{1.3}$$

$$P_{\hat{p}}(\xi) = \int_{-\infty}^{\infty} \mathrm{d}x\, W(x,p). \tag{1.4}$$

Thus, the Wigner function shares features with a probability distribution but can have regions where it takes negative values. If the Wigner function is positive everywhere, it constitutes an actual probability distribution and the measurements of the quadratures are explainable by a classical model. Therefore, negative Wigner functions are a signature of quantum behavior, and such states have been produced in experiments which can be certified from the measurement of $\hat{x}$ and $\hat{p}$ [17]. The definition of the Wigner function (1.2) can be extended to systems consisting of N Bosonic modes yielding a function depending on the vector $\mathbf{q} = (x_1, p_1, \ldots, x_N, p_N)$.

1.1.2   Gaussian states and operations

A class of states of both theoretical and experimental interest is provided by states for which the corresponding Wigner function is a Gaussian function. For a state $\rho \in \mathcal{B}(\mathcal{H}^{\otimes N})$, we define the vector of operators $\hat{\mathbf{O}} = (\hat{x}_1, \hat{p}_1, \ldots, \hat{x}_1, \hat{p}_1)$ which allows us to get the first moments $d_j = \mathrm{Tr}(\hat{O}_j \rho)$ and the covariance matrix

$$\gamma_{j,k} = 2\Re \mathrm{Tr}\left[(\hat{O}_j - d_j)(\hat{O}_k - d_k)\rho\right] \tag{1.5}$$

which collects the second moments. Gaussian states are fully determined by their first and second moments. Gaussian unitary operations, i.e., those mapping Gaussian states to Gaussian states are represented on the level of covariance matrices as $\gamma \mapsto S\gamma S^T$ with $S \in \mathrm{Sp}(2N, \mathbb{R})$, where $\mathrm{Sp}(2N, \mathbb{R})$ is the $2N$-dimensional real symplectic group. The corresponding unitary operations in state space are of the form $U = \exp(iH(\hat{\mathbf{O}}))$ where H is a quadratic polynomial in the quadrature operators or, equivalently, in the creation and

annihilation operators. As the entanglement properties are independent of the first moments, we assume all our states to be shifted such that $\mathbf{d} = 0$. The Gaussian operations fall into two classes: Passive operations do not change the total mean number of photons and fulfill $S^T S = \mathbb{1}$. They can be decomposed into a number of single mode optical phase shifters with $U_\phi(-i\phi\hat{a}^\dagger\hat{a})$, where ϕ is the angle of rotation, and beam splitters which couple two modes by the operation

$$B_\theta = \exp\left[\frac{\theta}{2}(\hat{a}_1^\dagger\hat{a}_2 - \hat{a}_2^\dagger\hat{a}_1)\right], \tag{1.6}$$

where $\cos(\theta/2)$ is the transmittivity of the beam splitter. In order to be able to perform any Gaussian unitary, we have to add one active operation, i.e., one that changes the mean photon number. This is provided by the single mode squeezer which can be realized by coupling the mode in question with a strong laser in a non-linear medium and which transforms the state by

$$S(\xi) = \exp\left[\frac{r}{2}(\hat{a}^2 - \hat{a}^{\dagger 2})\right], \tag{1.7}$$

where r parametrizes the strength of the squeezing. Measurements corresponding to projections on Gaussian states are called Gaussian measurements. Note that measurements of quadrature operators belong to this class as these states can be view as displaced, squeezed vacuum states in the limit of infinite squeezing. The covariance matrix of a state after a Gaussian measurement can be obtained by means of a Schur complement as detailed in Chapter 3.

1.1.3   Limitations of Gaussian states

As N-mode Gaussian states are described by $O(N^2)$ parameters, it is intuitive that they can represent only a tiny fraction of possible states. In Ref. [18], it has been shown that any protocol which starts with a Gaussian state and employs only Gaussian unitary operations and Gaussian measurements can be simulated efficiently, i.e., with polynomial effort in the number of modes N, on a classical computer. Thus, a quantum computer relying solely on Gaussian states, operations, and measurements can not allow for an exponential speed-up.

There are additional tasks which are impossible when restricting oneself to the Gaussian world with entanglement distillation being one of the most important ones: Assume two parties, called Alice and Bob, share a number of weakly entangled two-mode states. A protocol which transforms those states, not necessary deterministically, to a fewer number of pairs with higher entanglement, is called entanglement distillation. Such protocols are of vital

importance for distributing entanglement over large distances. Due to the effect of noise, the entanglement after the transmission of one part of the pair might be too little for the desired task, e.g. quantum key distribution, and distillation must be used. Ref. [19–21] showed that this is impossible when acting with Gaussian operations and measurements on Gaussian states. However, when the initial states are allowed to be non-Gaussian, the same class of protocolls indeed allow for distillation [19, 22]. We will see later on in Chapters 3 and 4 that both measurement-based quantum computing with Gaussian states and *arbitrary* measurements on one hand and with arbitrary states and Gaussian measurements is impossible, thus substantially extending the known no-go results.

1.1.4   Non-Gaussian states

Because quantum information with only Gaussian states and operations suffers from severe limitations, as discussed above, non-Gaussian states recently became one focus of continuous-variable quantum optics and quantum information. One of the easiest way to prepare a non-Gaussian state, though still quite challenging experimentally, is to subtract a photon from a squeezed vacuum [23]. To this aim, one combines a squeezed state with a vacuum mode on an imbalanced beam splitter and performs a photon counting measurement on one of the output ports. Success of the protocol is heralded by the detection of a single photon and the corresponding state reads

$$|\psi_{\theta,\xi}\rangle \propto \langle 1|_1 B_\theta \left(S(\xi)|0\rangle_1 \otimes |0\rangle_2 \right). \tag{1.8}$$

In the limit of vanishing reflectivity, i.e., $\theta \to 0$, one gets $|\psi_{\theta,0}\rangle \propto \hat{a}|S(\xi)|0\rangle$. However, in this limit, also the success probability goes to zero.

Other classes of non-Gaussian states include photon-added thermal states and squeezed single photons [24,25]. The latter is of particular interest because it can be used to approximate the Schroedinger cat, or kitten, states which are a building block of some proposed schemes of quantum computing [26].

1.2   Ultra-cold atoms in optical lattices

The principal drawback of using photons as carriers of quantum information is their weak interaction with each other. Even though continuous-variable states allow for the engineering of stronger interaction than states in the single photon regime, as we see later on, this still motivates to consider systems where the occurring interactions are stronger. Ultra-cold atoms have this feature. Their

interaction is short-ranged and by confining the atoms in space, it can be made considerably strong.

1.2.1   Realization and application of optical lattices

Optical lattices are standing waves created by counter-propagating laser beams. Due to the Stark effect, ultra-cold atoms, which are brought into this light field, experience a potential which is proportional to its intensity. Depending on the sign of the laser's detuning with respect to the closest transition in the atom, the potential minimum is either in the minima or the maxima of the intensity. Assuming the lattice to be translationally invariant and isotropic, the corresponding effective potential reads

$$V(\mathbf{r}) = V_0 \sum_{\nu=1}^{D} \sin^2(kr_\nu), \tag{1.9}$$

where V_0 is called the lattice depth, k the lattice wave number, and D denotes the number of spatial dimensions. By changing the laser's intensity, one can control both the mobility and the effective interaction of the atoms [11], by using Feshbach resonances, one can also control the atom-atom interaction strength directly and even change its sign [27]. Those possibilities, unparalleled in any condensed matter system, make optical lattices an ideal testbed for quantum many-body theory for Fermions, Bosons and mixtures of them [28]. In addition to realizing models which are believed to represent systems of interest occurring in nature, e.g. high-temperature superconductors, optical lattices are also used to engineer quantum states for information processing by using super-lattices, i.e., a second laser with a wave-length which is twice as large, it is possible to periodically vary V_0 [29]. This allows to perform quantum gates between neighboring lattice sites and has the potential of preparing resource states for MBQC.

1.2.2   Bose-Hubbard model

The interaction between spin-polarized, i.e., effectively spinless, ultra-cold atoms of alkali metals like sodium or potassium, is well approximated by a contact pseudo-potential. Using the language of second quantization, the total Hamiltonian can be written as $\hat{H} = \hat{H}_0 + \hat{H}_I$ with single particle part

$$\hat{H}_0 = \int d\mathbf{r}\, \hat{\Psi}^\dagger(\mathbf{r}) \left(-\frac{\hbar^2 \nabla^2}{2m} + V(\mathbf{r}) \right) \hat{\Psi}(\mathbf{r}), \tag{1.10}$$

and an interaction part

$$\hat{H}_\mathrm{I} = \frac{g}{2} \int \mathrm{d}\mathbf{r}\, \hat{\Psi}(\mathbf{r})^\dagger \hat{\Psi}(\mathbf{r})^\dagger \hat{\Psi}(\mathbf{r}) \hat{\Psi}(\mathbf{r}). \tag{1.11}$$

where g is the interaction strength which is positive for a repulsive and negative for an attractive interaction. One can now expand the field operators as

$$\hat{\psi}(\mathbf{r}) = \sum_{\mathbf{i}} \sum_{n=0}^{\infty} w_{\mathbf{i},n}(\mathbf{r}) \hat{b}_{\mathbf{i},\mathbf{n}} \tag{1.12}$$

where $[\hat{b}_{\mathbf{i},\mathbf{n}}, \hat{b}_{\mathbf{j},\mathbf{m}}^\dagger] = \delta_{\mathbf{i},\mathbf{j}} \delta_{\mathbf{n},\mathbf{m}}$ and $w_{\mathbf{i},\mathbf{n}}$ is the Wannier function of the $\mathbf{n}$th band ($\mathbf{n}$ is an D-dimensional vector) centered around lattice site $\mathbf{j}$. An often made approximation, whose applicability and accuracy will be discussed in Appendix B, is to approximate the sum over n in (1.12) by the zeroth term (in this case, one can omit the band index) and, in addition, neglecting all interactions but on-site interactions and all hopping processes but those between neighboring sites. In this limit, one obtains the famous single-band Bose-Hubbard model [30]

$$\hat{H}_\mathrm{BH} = -J \sum_{<\mathbf{i},\mathbf{j}>} \hat{b}_{\mathbf{i}}^\dagger \hat{b}_{\mathbf{j}} + \sum_{\mathbf{i}} \frac{U}{2} \hat{n}_{\mathbf{i}}(\hat{n}_{\mathbf{i}} - 1) \tag{1.13}$$

where $< \mathbf{i},\mathbf{j} >$ denotes the summation over nearest-neighbor pairs and $\hat{n}_{\mathbf{i}} = \hat{b}_{\mathbf{i}}^\dagger \hat{b}_{\mathbf{i}}$. The hopping and interaction parameters J and U are determined by

$$J = -\int \mathrm{d}\mathbf{r}\, w_{\mathbf{i}}^*(\mathbf{r}) \left(-\frac{\hbar^2 \nabla^2}{2m} + V(\mathbf{r}) \right), \tag{1.14}$$

$$U = g \int \mathrm{d}\mathbf{r}\, |w_{\mathbf{i}}|^4, \tag{1.15}$$

where $\mathbf{i}$ and $\mathbf{j}$ are arbitrary neighboring sites.

Even though the Bose-Hubbard Hamiltonian (1.13) is arguably one of the simplest non-trivial Bosonic model, its physics is very rich. By changing the ratio between J and U, which can be done by adjusting the lattice depth V_0, one can realize a quantum phase transition between a delocalized superfluid (for large J/U) and a localized Mott insulator (for small J/U) which has been observed in a seminal experiment in 2002 [11]. Due to particle number conservation, the total atom operator $\hat{N} = \sum_{\mathbf{j}} \hat{n}_{\mathbf{j}}$ commutes with $\hat{H}_\mathrm{BH}$ and a super-selection rule present for massive Bosons forbids the superposition of states belonging to different eigenvalues of $\hat{N}$. Therefore, assuming the Bose-Hubbard model to be a sufficiently accurate description of the system present in the experiment under question, the Hilbert space for a given total atom number is finite dimensional, making tomography at least conceivable.

Furthermore, for a repulsive interaction, there is an energy penalty for too many atoms being on the same lattice site which allows to truncate the *local* Hilbert space at some maximal occupation number n_{max} while still getting a good approximation.

The Bose-Hubbard model can be generalized in many directions: Interactions beyond nearest neighbors can result in the formation of exotic phases like super-solids, using atoms with spin degrees of freedom allows to study effects of magnetism, and adding Fermions can suppress or enhanced super-fluidity due to interaction effects [28, 31, 32]. In Chapter 6, we give a detailed description how to perform tomography in a, yet to be defined, efficient fashion while in Appendix B, the validity of the single-band approximation in the Bose-Hubbard model is discussed in detail, especially for the situation where additional Fermions are present in the system.

Chapter 2

Important protocols in quantum information

As already mentioned in the introduction, an important part of quantum information theory is to explore how the fundamental effects of quantum mechanics can be turned into protocols for information processing. One example for such a protocol, which has the largest practical importance at the present time, is quantum key distribution where two parties want to agree on a common secret key without a potential adversary being able to obtain knowledge about this. This key can then be used to encrypt a message between them. We come back to the question of key distribution in the appendix and focus on two other protocols in this chapter: Measurement-based quantum computing which allows for the execution of quantum algorithms without controlled interactions during the computation and quantum state tomography which deals with the problem of obtaining an accurate description of a quantum state.

2.1 Measurement-based quantum computing

As already discussed in the introduction, measurement-based quantum computing (MBQC) has the potential to make the construction of a scalable quantum computer notably easier compared to the commonly used circuit model. The universal resource state discovered first was the cluster state described by Briegel and Raussendorff [6]. On any graph, such a cluster state can be prepared by placing a qubit in the state $|+\rangle = (1/\sqrt{2})(|0\rangle + |1\rangle)$ on every vertex of the cluster and applying controlled-Z gates, i.e., $C_Z = \mathrm{diag}\,(1, 1, 1, -1)$, to all pairs of vertexes (also called sites), which are connected by an edge (also called bond). Gross et al. have developed a formalism based on matrix

product states (MPS), which allows to describe MBQC using this and other resources. In the present section, we give a very short introduction into the formalism while in Chapter 4 we provide a way to generalize this formalism to make it applicable to continuous-variable systems.

We only consider one-dimensional systems, called quantum wires [33], which are used to process a single qubit and discuss the issue of coupling those wires to universal resources in later chapters. Consider a system of L sites with local Hilbert space dimension d_p. Assume that to any $i = 1, \ldots, d_p$, a $D \times D$ complex matrix $A[i]$ is associated, and they fulfill the normalization relation $\sum_{i=1}^{D} A[i]^\dagger A[i] = \mathbb{1}$. Furthermore, let $|L\rangle$, $|R\rangle \in \mathbb{C}^D$ be normalized state vectors describing the boundary conditions. This defines a translationally invariant matrix product state

$$|\Psi\rangle = \sum_{i_L, \ldots, i_1 = 0}^{d_p - 1} \langle L|A[i_L] \ldots A[i_2]A[i_1]|R\rangle|i_L r., \ldots, i_1\rangle. \tag{2.1}$$

We now want to know what happens if we perform a projective measurement on the first lattice site and obtain a result corresponding to a projection to the state $|\psi_i\rangle$. Defining $B[i] := \sum_{j=0}^{d_p - 1} \langle \psi_i | j \rangle A[j]$, the post-measurement state reads

$$|\Phi\rangle \propto \sum_{i_L, \ldots, i_2 = 0}^{d_p - 1} \langle L|A[i_L] \ldots A[i_2]B[i_1]|R\rangle|i_L, \ldots i_2\rangle, \tag{2.2}$$

which can be interpreted as an application of $B[i_1]$ to the state vector $|R\rangle$, which is often called the correlation system. To turn this insight into a protocol for single qubit processing, we need three things: First, by a clever choice of the measurement bases we want to perform, up to arbitrary accuracy, any single-qubit gate. Second, we need a basis which initializes the correlation system and third, we need a way to perform some measurement of this, purely virtual, correlation system.

As the requirements for this to be possible are extensively discussed later in this thesis, we only give an example for the most transparent case of the cluster state for which $D = d_p = 2$ and the MPS-matrices read $A[0] = |+\rangle\langle 0|$ and $A[1] = |-\rangle\langle 1|$. Here, we already see that that a measurement in the computational basis leaves the correlation system in the state $|+\rangle$ (if the outcome is "0") or $|-\rangle$ (if the outcome is "1"). Furthermore, if the correlation systems is, prior to the measurement, in the state $|0\rangle$ ($|1\rangle$), one always gets the outcome "0" ("1"). Thus, the second and the third requirement are already fulfilled. To see that one can perform any single qubit gate, we define for any

$\phi \in [0, 2\pi)$, a basis consisting of

$$|\phi_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\phi}|1\rangle), \quad |\phi_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - e^{i\phi}|1\rangle). \tag{2.3}$$

Let H be the Hadamard gate, Z the Z-gate and $S(\phi) = \mathrm{diag}\,(1, \exp(i\phi))$ be the phase gate. Then, the matrices corresponding to a measurement in the basis (2.3) are

$$B_\phi[0] \propto HS(\phi), \quad B_\phi[1] \propto HZS(\phi). \tag{2.4}$$

As H and $S(\phi)$ form a universal gate set, any $U \in U(2)$ is reachable up to an irrelevant global phase factor. As one can not control the measurement outcomes, undesired additional Z-gates, sometimes called by-products, might occur. However, this can be undone by additional measurements for $\phi = 0$. The probabilistic nature of quantum measurements results in a varying length of the computation which is no problem if it is on average still efficient as it is the case here.

2.2   Quantum state tomography

Quantum state tomography is the task of getting an accurate description of a quantum state, i.e., the density matrix or, especially when dealing with continuous-variable systems, the Wigner function. To this aim, one has to perform measurements, record the frequencies of the various outcomes, and do some form of post-processing of the obtained data. There are three major families of tomography protocols: Methods based on direct linear inversion, maximum likelihood techniques, and ideas based on compressed sensing. We give a very short introduction while focussing on continuous variable systems which are most relevant for the present work.

2.2.1   Linear inversion

As discussed above, the Wigner function is a faithful description of the state of a single harmonic oscillator. Homodyne detection is performed by combining the mode in question in an interferometer with a strong coherent state, called the local oscillator, as a phase reference, measure the photon current on both output ports and subtract them. A parameter θ can be chosen by shifting the relative phase between the signal mode and the local oscillator. The obtained probability distribution of the differences is

$$p_\theta(x) = \int_{-\infty}^{\infty} \mathrm{d}p\, W(q\cos\theta - p\sin\theta, q\sin\theta + p\cos\theta). \tag{2.5}$$

As detailed in Ref. [15], Eq. (2.5) can be inverted as

$$W(q,p) = \frac{1}{2\pi^2} \int_0^\pi \mathrm{d}\theta \int_{-\infty}^\infty \mathrm{d}x\, p_\theta(x,\theta) K(q\cos\theta + p\sin\theta - x) \qquad (2.6)$$

where

$$K(x) = \frac{1}{2} \int_{-\infty}^\infty \mathrm{d}\xi\, |\xi| \exp(i\xi x) \qquad (2.7)$$

is the kernel of integration. Unfortunately, K does only exist as a distribution and not as a proper function. However, it can be regularized and be used to reconstruct the Wigner function from recorded data.

Any physical quantum state ρ must only posses a finite amount of energy $E_{\mathrm{mean}} = \mathrm{Tr}(\rho\hat{n})$. This implies, that the matrix elements of ρ in the Fock basis, i.e., $\langle m|\rho|n\rangle$ must decay to zero for growing m and n. This fact, which is made quantitative and generalized to the multi-mode case in Chapter 5, allows us to truncate the Hilbert space at some Fock level N. The matrix elements can be calculated as

$$\langle m|\rho|n\rangle = \frac{1}{2\pi} \int_{-\pi}^\pi \mathrm{d}\theta \int_{-\infty}^\infty \mathrm{d}x\, p_\theta(x) = f_{m,n}(x)\exp(i(m-n)\theta), \qquad (2.8)$$

where $f_{m,n}$ are called pattern functions which are well behaved and can be calculated by some recursion relation. We also note that, if the Hilbert space is truncated at N, the integral over the phase angle θ can be replaced by an average over $N+1$ equidistant value of θ.

2.2.2   Maximum likelihood techniques

The probability distributions p_θ are never known exactly. First, any experiment suffers from decoherence and other inaccuracies. Second, as one can only perform a finite number of repetitions of the experiment, there is always statistical noise present. Therefore, the reconstructed density matrix does, in general contain negative eigenvalues, rendering it unphysical. Maximum likelihood methods, which allow to tackle this problem, are one of the standard methods of obtaining a probability distribution from samples. The key idea is to find the probability distribution which maximizes the probability to obtain the recorded measurement result. We shortly sketch how this idea can be used to perform continuous-variable quantum state tomography [34].

The probability distribution (2.5) can, for some state ρ, also be written as $p_\rho(\theta,x) \propto \mathrm{Tr}(\hat{\Pi}_\theta(x)\rho)$ where the operator $\hat{\Pi}_\theta(x)$ reads in the Fock basis

$$\langle m|\hat{\Pi}_\theta(x)|n\rangle = e^{i(m-n)\theta}\psi_m(x)\psi_n(x), \qquad (2.9)$$

where ψ_n is the nth energy eigenfunction of the harmonic oscillator. Assume that we have taken m measurements in total and observed data points (θ_i, x_i) for $i = 1, \ldots, m$. Then, the likelihood, i.e., the probability that ρ produces the given data points is

$$\mathcal{L} = \prod_i p_\rho(\theta, x). \tag{2.10}$$

An iterative method to find the state ρ which maximizes $\mathcal{L}$, or, more accurately, a finite-dimensional truncation of ρ, is the following: Start with an initial state $\rho^{(0)}$, e.g. the maximally mixed state. Then apply the iteration $\rho^{(k+1)} = \mathcal{N} R(\rho)^{(k)} \rho^{(k)} R(\rho(k))$, where

$$R(\rho) = \sum_{i=1}^{m} \frac{\hat{\Pi}(\theta_i, x_i)}{p_\rho(\theta, x)} \tag{2.11}$$

and $\mathcal{N}$ is chosen such that the state stays normalized. Iterating this procedure, one converges to the desired result.

2.2.3 Compressed sensing

In many situations where quantum state tomography is performed, the state in question is not completely arbitrary but close to a state with low rank. Compressed sensing allows to both reduce the number of measurements needed to reconstruct the state notably and to make the post-processing much more efficient [9, 10, 35, 36]. Here, we only present the general idea while developing a theory of compressed sensing applicable to the most general situation in Chapter 5.

Let $\mathcal{H} = \mathbb{C}^d$. The bounded operators (or Hermitian matrices) acting on $\mathcal{H}$, denoted by $\mathcal{B}(\mathcal{H})$, form a d^2 dimensional Hilbert space with the scalar product, sometimes called the Hilbert-Schmidt scalar product, $(A, B) = \text{Tr}(AB)$. This gives a straightforward method to reconstruct an unknown quantum state $\rho \in \mathcal{B}(\mathcal{H})$: Just chose observables $\{w_1, \ldots, w_{d^2}\}$ being an orthonormal basis of $\mathcal{B}$ and measure the expectation values of all of them to obtain ρ in the basis of the w_i. For a fully general state, one has to determine the expectation value of $d^2 - 1$ observables. If, on the other hand, rank $\rho = r$, the state is described by only $(2d - 1)r - 1$ real parameters, which is a great reduction if $r \ll d$.

However, how to find the lowest rank state compatible with the measurement results and to check whether the reconstruction is unique is highly non-trivial even though the algorithm of compressed sensing is quite simple: Let $\|\cdot\|_1$ denote the trace-norm, i.e., the sum over all singular values. From the d^2 observable in the basis chose $m = Cr\kappa d \log^2 d$ ones, $w_{i_1}, \ldots, w_{i_m}$ uniformly at

random where C is a constant and the purpose of κ is explained below. Then measure their expectation values and calculate

$$\hat{\sigma} = \text{argmin}_{\sigma \in \mathcal{B}} \|\sigma\|_1 \text{ subject to } (w_{i_k}, \sigma) = (w_{i_k}, \rho) \, \forall k = 1, \ldots, m. \qquad (2.12)$$

This optimization problem can be written as a semi-definite program and, therefore, be solved efficiently in d. If the observables fulfill a certain incoherence property as provided in Chapter 5, the probability that the reconstruction fails, i.e., that $\hat{\sigma} \neq \rho$ is smaller than $2^{-\kappa}$. This is true for any state ρ. Thus, compressed sensing allows to reduce the number of necessary measurements from $O(d^2)$ to the almost optimal $O(d \log^2 d)$. Note that for certain classes of observables, one can even adapt the algorithm in such a way that it becomes deterministic [36].

If the mentioned condition is not fulfilled, there are quantum states which will still need on the order of d^2 measurement settings to be reconstructed. However, as we show in Chapter 5, the described algorithm is useful even in this case as one can certify its success from the obtained data only. Thus, one can just try to reconstruct the state from few measurements and just perform additional measurements when the certification fails.

Chapter 3

Limitations of quantum computing with Gaussian cluster states

Limitations of quantum computing with Gaussian cluster states

M. Ohliger,[1] K. Kieling,[1] and J. Eisert[1,2,*]

[1]*Institute of Physics and Astronomy, University of Potsdam, D-14476 Potsdam, Germany*
[2]*Institute for Advanced Study Berlin, D-14193 Berlin, Germany*
(Received 18 May 2010; published 28 October 2010)

We discuss the potential and limitations of Gaussian cluster states for measurement-based quantum computing. Using a framework of Gaussian-projected entangled pair states, we show that no matter what Gaussian local measurements are performed on systems distributed on a general graph, transport and processing of quantum information are not possible beyond a certain influence region, except for exponentially suppressed corrections. We also demonstrate that even under arbitrary non-Gaussian local measurements, slabs of Gaussian cluster states of a finite width cannot carry logical quantum information, even if sophisticated encodings of qubits in continuous-variable systems are allowed for. This is proven by suitably contracting tensor networks representing infinite-dimensional quantum systems. The result can be seen as sharpening the requirements for quantum error correction and fault tolerance for Gaussian cluster states and points toward the necessity of non-Gaussian resource states for measurement-based quantum computing. The results can equally be viewed as referring to Gaussian quantum repeater networks.

I. INTRODUCTION

Optical systems offer a highly promising route to quantum information processing and quantum computing. The seminal work in Ref. [1] showed that even with linear optical gate arrays alone and appropriate photon counting measurements, efficient linear optical computing is possible. The resource overhead of this proof-of-principle architecture for quantum computing was reduced, indeed by orders of magnitude, by directly making use of the idea of measurement-based quantum computing with cluster states [2–4]. Such an approach is appealing for many reasons; the reduction of resource overhead is one, and the clear-cut distinction between the creation of entanglement as a resource and its consumption in computation is another. This idea was further developed into the continuous-variable (CV) version thereof [5–8], which aims at avoiding limitations related to efficiencies of creation and detection of single photons. In this context, Gaussian states play a quite distinguished role, as they can be created by passive optics, optical squeezers, and coherent states, i.e., the states produced by a usual laser [9–13]: Indeed, Gaussian cluster states are a promising resource for instances of quantum computing with light. Such a CV scheme allows for deterministic preparation of resource states, while schemes based on linear optics with single photons require preparation methods which are intrinsically probabilistic.

In this work, however, we highlight and flesh out some limitations of such an approach. We do so to clarify the exact requirements that any scheme for CV quantum computing based on Gaussian cluster states eventually will have to fulfill and what quantum error correction and fault-tolerant approaches eventually have to deliver. Specifically, we show that Gaussian local measurements alone will not suffice to transport quantum information across the lattice, even on complicated lattices described by an arbitrary graph of finite dimension: Any influence of local measurements is confined

to a local region, except from exponentially suppressed corrections. This can be viewed as an impossibility of Gaussian error correction in the measurement-based setting. What is more, even under non-Gaussian measurements, this obstacle cannot be overcome, to transport or process quantum information along slabs of a finite width: Any influence of local measurements will again exponentially decay with the distance. This observation suggests that—although the initial state is perfectly known and pure—finite squeezing has to be tackled with a full machinery of quantum error correction and fault tolerance [14–16], yet developed for this type of system and, presumably, giving rise to a massive overhead. No local measurements or suitable sophisticated encodings of qubits in finite slabs—reminding, e.g. of encodings of the type of Ref. [16]—can uplift the initial state to an almost perfect universal resource. To arrive at this conclusion, in some ways, we explore ideas of measurement-based computing beyond the one-way model [2] as introduced in Ref. [17] and further developed in Refs. [18–22]. We highlight the technical results as "observations" and discuss implications of these results in the text. While these findings do not constitute a "no-go" argument for Gaussian cluster states, they do seem to require a very challenging prescription for quantum error correction and further highlight the need to identify alternative schemes for CV quantum computing, specifically schemes based on *non-Gaussian CV states*. *Small-scale implementations* of *Gaussian cluster-state computing*, as we will see, are also *not* affected by these limitations.

The structure of this article is as follows: In Sec. III we discuss the concept of Gaussian projected entangled pair states (GPEPSs), forming a family of states including the physical CV Gaussian cluster state. In Sec. IV we discuss the impact of Gaussian measurements on GPEPSs and show that under this restriction the localizable entanglement in every GPEPS decays exponentially with the distance between any two points on an arbitrary lattice. This also has implications for Gaussian quantum repeaters, which we investigate in detail. Then we leave the strictly Gaussian stage in Sec. V and present our main result, showing that under more general measurements

———————
*jense@qipc.org

of GPEPSs, quantum information processing in finite slabs is still not possible. We discuss requirements for error correction, before presenting concluding remarks.

II. PRELIMINARIES

A. Gaussian states

Before we turn to measurement-based quantum computing (MBQC) on CV states, we briefly review some basic elements of the theory of Gaussian states and operations which are needed in this article [9–12]. Readers familiar with these concepts can safely skip this section. Although the statements made in this work apply to all physical systems described by *quadratures* or *canonical coordinates*, including, for example, micromechanical oscillators, we have a *quantum optical system* in mind and often use language from this field as well. Any system of N bosonic degrees of freedom, for example, N light modes, can be described by canonical coordinates $x_n = (a_n + a_n^\dagger)/2^{1/2}$ and $p_n = -i(a_n - a_n^\dagger)/2^{1/2}$, $n = 1, \ldots, N$, where a_n ($a_n^\dagger$) annihilates (creates) a photon in the respective mode. When we collect these $2N$ canonical coordinates in a vector $O = (x_1, p_1, \ldots, x_N, p_N)$, we can write the commutation relations as $[O_j, O_k] = i\sigma_{j,k}$, where the *symplectic matrix* σ is given by

$$\sigma = \bigoplus_{j=1}^{N} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{1}$$

Gaussian states are fully characterized by their first and second moments alone. The *first moments* form a vector d with entries $d_j = \text{tr}(O_j \rho)$, while the second moments, which capture the fluctuations, can be collected in a $2N \times 2N$ matrix γ, the so-called *covariance matrix* (CM), with entries

$$\gamma_{j,k} = 2\text{Re} \, \text{tr} \, [\rho(O_j - d_j)(O_k - d_k)]. \tag{2}$$

Hence, Gaussian states are complete characterized by d and γ. *Gaussian unitaries*, that is, unitary transformations acting in Hilbert space preserving the Gaussian character of the state correspond to symplectic transformations on the CM. They in turn correspond to maps $\gamma \mapsto S\gamma S^T$ with $S\sigma S^T = \sigma$. The set of such *symplectic transformations* forms the group $Sp(2N, \mathbb{R})$. A set of particularly important example Gaussian states are the *coherent states*, for which the state vectors read, in the photon number basis,

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \tag{3}$$

and are described by $d = \sqrt{2}(\text{Re} \, \alpha, \text{Im} \, \alpha)$ and $\gamma = \text{diag}(1, 1)$. Single-mode *squeezed states* are characterized by lower fluctuations in one phase-space coordinate. The CM can, in a suitable basis, then be written as $\gamma = \text{diag}(x, 1/x)$ with $x \neq 0$.

B. MBQC on Gaussian cluster states

The first proposal for MBQC on CV states has been based on so-called Gaussian cluster states and works in almost-complete analogy to the qubit case [5–8]. As such, the formulation is based on "infinitely squeezed" and hence unphysical states using infinite energy in preparation: It can be created by initializing every mode in the $p = 0$ "eigenstate" of p (formally an improper eigenstate of momentum, a concept that can be made rigorous, for example, in an algebraic formulation [23]). This is the CV analog to the state vector $|+\rangle = (|0\rangle + |1\rangle)/2^{1/2}$ in the qubit case. Then the operation $e^{ix \otimes x}$, the analog to the C_Z gate, is applied between all adjacent modes. This state allows universal MBQC to be performed with Gaussian and one non-Gaussian measurement. The state as such is not physical and not contained in Hilbert space. The argument, however, is that it should be expected that a finitely squeezed version inherits essentially the same properties. Replacing them by finitely squeezed ones, we obtain a state which we call a *physical Gaussian cluster state*.

III. GPEPS

Projected entangled pair states or *tensor product states* have been used for qubits to generalize matrix product states or *finitely correlated states* [24,25] from one-dimensional (1D) chains to arbitrary graphs [26–28]. One suitable way to define them is via a valence-bond construction: One can create a state by placing entangled pairs—constituting "virtual systems"—on every bond of the lattice and then applying a suitable projection to a single mode at every lattice site. These projections, often taken to be equal, together with the specification of the initial entangled states, then serve as a description of the resulting state. *Matrix product states for Gaussian states* (MPSGs) have been studied to obtain correlation functions and entanglement scaling in 1D chains [29].

In this work we focus on GPEPSs which can be obtained from non–perfectly entangled pairs. The bonds we consider are *two-mode squeezed states* (TMSSs), the state vectors of which have the photon number representation

$$|\psi_\lambda\rangle = (1 - \lambda^2)^{1/2} \sum_{n=0}^{\infty} \lambda^n |n, n\rangle, \tag{4}$$

where $\lambda \in (0, 1)$ is the *squeezing parameter*. We denote the corresponding density matrix ρ_λ. For $\lambda \to 1$ the state becomes "maximally entangled," but this limit is not physical because it is not normalizable and has infinite energy as already mentioned. We, therefore, carefully analyze the effects stemming from the fact that $\lambda < 1$. The CM of this state reads

$$\gamma_\lambda = \begin{bmatrix} \cosh(2r) & 0 & \sinh(2r) & 0 \\ 0 & \cosh(2r) & 0 & -\sinh(2r) \\ \sinh(2r) & 0 & \cosh(2r) & 0 \\ 0 & -\sinh(2r) & 0 & \cosh(2r) \end{bmatrix}, \tag{5}$$

where $\tanh(r/2) = \lambda$. This number r is also referred to as the *squeezing parameter* when there is no risk of mistaking one for the other. It is also known that any pure bipartite multimode Gaussian state can be brought into the tensor product of a TMSS [10,30] by means of local unitary Gaussian operations, each having a CM in the form of Eq. (5). Then the largest r in the vector of the resulting TMSS is referred to as its squeezing parameter.
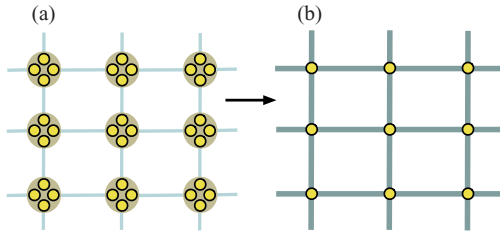
FIG. 1. (Color online) GPEPS on an arbitrary graph, here one representing a cubic lattice. (a) Connected dots represent two-mode squeezed states; circles denote vertices where Gaussian projections are being performed. (b) The resulting GPEPS after local Gaussian projections have been performed on the virtual systems. Any Gaussian cluster state can be prepared in this fashion.

We also discuss GPEPSs on general graphs $G = (V, E)$, as shown in Fig. 1. Vertices G here correspond to physical systems, and edges E to connections of neighborhood. In any such graph, $d(.,.)$ is the natural graph-theoretical distance between two vertices. As we often consider the system of bonds before the projection operation is performed, we employ the following notation: We speak of operations on *virtual systems* when referring to collective operations on modes before the projection is applied and often emphasize this when speaking of a single physical system with Hilbert space $\mathcal{H} = \mathcal{L}^2(\mathbb{R})$. Note that we also allow for more than one edge between two vertices in a graph.

When a particular vertex has N adjacent bonds, the projection map is a Gaussian operation of the form

$$V : \mathcal{H}^{\otimes N} \to \mathcal{H}. \tag{6}$$

This operation can always be made trace preserving [9,12,31,32], in quite sharp contrast to the situation in the finite-dimensional setting. This operation is also referred to as GPEPS projection. This operation can always be realized by mixing single-mode squeezed states on a suitably tuned beam splitter, which means that *inline* squeezers are not necessary [33]. Note that any such state can also be used as a variational state to describe ground states of many-body systems and, by construction, satisfies an *entanglement area law* [34].

IV. GAUSSIAN OPERATIONS ON A GPEPS

In this section, we discuss Gaussian operations on a GPEPS and derive some statements on entanglement swapping, the localizable entanglement, and the usefulness as a resource for MBQC. Since all measurements are assumed to be Gaussian as well, this is, as such, not yet a full statement on universality, but already shows that the natural operation for transport of logical information in such a Gaussian cluster state does not work with such local measurements.

A. Localizable entanglement

The *localizable entanglement* between two sites A and B in the graph $G = (V, E)$ is defined by the maximal entanglement obtainable on average when performing projective measurements at all sites but A and B [35]. When we require both the

initial state and the measurements to be Gaussian [36,37], the situation is simplified, as the entanglement properties do not depend on the measurement outcomes [9,12,31,32]. Thus, we do not need to average, but only to find the best measurement strategy. To be specific, we measure the entanglement in terms of the *logarithmic negativity*, which can be defined as [38–40]

$$E(\rho) = \log_2 \|\rho^{T_A}\|_1, \tag{7}$$

where T_A denotes the *partial transpose* with respect to subsystem A and $\| \cdot \|_1$ the trace-norm, and we use the natural logarithm. For a TMSS, E coincides with the squeezing parameter as $E(\rho_\lambda) = r$. It is important to note, however, that this choice has only been made for notational convenience: In our statements on asymptotic degradation of entanglement, any other measure of entanglement would also do, specifically the entropy of entanglement for pure Gaussian states and the *distillable entanglement* or the *entanglement cost* for mixed states.

We mostly focus on two variants of the concept of localizable entanglement: Whenever we allow only for Gaussian local measurements, we refer to this quantity as *Gaussian localizable entanglement*, abbreviated E_G. Then we consider the situation where we ask for fixed subspaces S_A and S_B in the Hilbert spaces associated with sites A and B to become entangled by means of local measurements. We then refer to *subspace localizable entanglement* E_S. Both concepts directly relate to transport in MBQC.

B. Entanglement swapping

The task of localizing entanglement in a PEPS is closely related to that of *entanglement swapping* [41]. In this situation we have three parties, A, B, and C, where both A and B and B and C share an entangled pair each. Then B, consisting of B_1 and B_2, is allowed to perform an arbitrary Gaussian operation on its parts of the two pairs, followed by a measurement. The task is to choose the operation in such a way that the resulting entanglement between A and B is maximum.

Lemma 1. Optimality of Gaussian Bell measurement for entanglement swapping of TMSSs. For two pairs of entangled TMSSs shared between A and B_1 and between B_1 and C, the supremum of maximum achievable negativity between A and C by a local Gaussian measurement in B_1, B_2 is approximated by the measurement that best approximates a Gaussian Bell measurement.

We consider the situation of having a TMSS (5)

$$|\psi\rangle_{A,B_1} = |\psi_{\lambda_1}\rangle_{A,B_1}, \quad |\psi\rangle_{B_2,C} = |\psi_{\lambda_2}\rangle_{B_2,C} \tag{8}$$

with some $\lambda_1, \lambda_2 > 0$ and restricting the operation on B to be Gaussian. Furthermore, we allow for operations which do not succeed with unit probability. We have to allow for general local Gaussian operations and, also, for arbitrary local additional Gaussian resources, with CM γ_B on mode B_3, on an arbitrary number of modes. The initial CM of the system hence reads

$$\gamma = \gamma_{\lambda_1} \oplus \gamma_{\lambda_2} \oplus \gamma_{B_3}. \tag{9}$$

Without loss of generality, one can assume that one performs a single projection onto a pure Gaussian state on all modes

referring to B. Ordering modes to A,C,B_1,B_2,B_3, one can write the CM in block form as

$$\gamma = \begin{bmatrix} U & V & 0 \\ V^T & W & 0 \\ 0 & 0 & \gamma_{B_3} \end{bmatrix}, \qquad (10)$$

with U referring to A, C and V referring to B_1,B_2. When we project the modes B_1, B_2, and B_3 onto a pure Gaussian state with CM Γ, the CM of the resulting state of A and C, postselected for that outcome, is given by the Schur complement [9,31,32],

$$\gamma_{A,C} = \begin{bmatrix} U & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} V & 0 \end{bmatrix} \left(\begin{bmatrix} W & 0 \\ 0 & \gamma_{B_3} \end{bmatrix} + \Gamma \right)^{-1} \begin{bmatrix} V^T \\ 0 \end{bmatrix}. \qquad (11)$$

Any symplectic operation S applied to B before the measurement can, of course, also just be absorbed into the choice of the CM Γ. Writing

$$\begin{bmatrix} W & 0 \\ 0 & \gamma_{B_3} \end{bmatrix} + \Gamma = \begin{bmatrix} X & Y \\ Y^T & Z \end{bmatrix}, \qquad (12)$$

one finds that the upper-left principal submatrix of the inverse can be written as

$$\begin{bmatrix} X & Y \\ Y^T & Z \end{bmatrix}^{-1} \Bigg|_{B_1,B_2} = (X - YZ^{-1}Y^T)^{-1}, \qquad (13)$$

again, in terms of a Schur complement expression. Since $\gamma_{B_3} + i\sigma \geqslant 0$ and the same holds for the subblock on B_3 of Γ, these matrices are clearly positive. Using operator monotonicity of the inverse function, one finds that

$$(X - YZ^{-1}Y^T)^{-1} \geqslant 0 \qquad (14)$$

holds, since $YZ^{-1}Y^T \geqslant 0$. Therefore,

$$\gamma_{A,C} = \gamma'_{A,C} + P, \qquad (15)$$

with a matrix $P \geqslant 0$. Here $\gamma'_{A,C}$ is the CM following the same protocol, but where Γ is replaced by an identical CM, but with $Y = 0$. To arrive at such a CM is always possible and still gives rise to a valid CM by virtue of the pinching inequality. This is still merely the CM of the Gaussian state, subjected to additional classically correlated Gaussian noise. In other words, it is always optimal to treat B_3 as an innocent bystander and not to perform an entangling measurement between B_1 and B_2, on the one hand, and B_3, on the other hand: quite consistent with what one could have intuitively assumed. We can hence focus on the situation where B_3 is absent and we merely project onto a pure Gaussian state in B_1 and B_2.

It is then easy to see that there is no optimal choice, but the supremum can be better and better approximated by considering more and more squeezed TMSSs (or "infinitely squeezed states" in the first place), that is, on $|\psi_\lambda\rangle$ in the limit of $\lambda \to 1$, which is the CV analog to the Bell state for qudits. This measurement can be realized by mixing B_1 and B_2 on a beam splitter with reflectivity $R = 1/2$ and performing homodyne measurement on both modes afterward (i.e., a projection on an infinitely squeezed single-mode state being an improper eigenstate of the position operator). From Eqs. (5) and (11),

with $\Gamma = \gamma_\lambda$ and performing the limit $\lambda \to 1$, we can calculate the CM of the resulting state. It has the form of (5), with

$$r = f(r_1,r_2) = \frac{1}{2}\text{arcosh}\frac{1 + \cosh 2r_1 \cosh 2r_2}{\cosh 2r_1 + \cosh 2r_2}. \qquad (16)$$

We note that f is symmetric in its arguments and fulfills $f(r_1,r_2) < \min\{r_1,r_2\}$ and $\lim_{r_1 \to \infty} f(r_1,r_2) = r_2$. This means that arbitrarily faithful entanglement swapping is possible exactly in the limit of infinite entanglement. Otherwise, the entanglement necessarily deteriorates [41].

To show that this measurement is indeed optimal, we set

$$\Gamma = S\gamma_\lambda S^T, \qquad (17)$$

where $S \in Sp(4,\mathbb{R})$. Calculating the resulting degree of entanglement, a direct and straightforward inspection reveals that $E(\rho_{A,C})$ can only decrease whenever we choose $S \neq \mathbb{1}$.

C. 1D chain

We now turn to a one-dimensional GPEPS, not allowing multiple bonds in the valence-bond construction, and are in the position to show the following observation.

Observation 1. Exponential decay of Gaussian localizable entanglement in a 1D chain. Let G be a 1D GPEPS, and A and B two sites. Then

$$E_G(A,B) \leqslant c_1 e^{-d(A,B)/\xi_1}, \qquad (18)$$

where $c_1, \xi_1 > 0$ are constants. The best performance is reachable by passive optics and homodyning only.

To prove this, we interpret the preparation projection (6) and the following measurements of the localizable entanglement protocol as a sequence of instances of entanglement swapping. Clearly, to allow for general Gaussian projections is more general than using (i) the specific Gaussian projection of the PEPS, followed by a (ii) suitable Gaussian projection onto a single mode; hence every bound shown for this setting will also give rise to a bound to the actual 1D Gaussian chain. If $d(A,B)$ is again the graph-theoretical distance between A and B, we have to swap $k = d(A,B) - 1$ times. Defining $g(r) = f(r,r_I)$, where r_I is the initial strength of all bonds, and iterating the argument, we obtain

$$r_{A,B} = (g^{\circ k})(r_I) = F(k). \qquad (19)$$

As the negativity is up to a simple rescaling equal to this two-mode squeezing parameter, the only task left is to show that $F(k)$ decays exponentially. To do this, we need $\text{arcosh}(x) = \log_2[x + (x^2 - 1)^{1/2}]$ and the following relations which hold for $x \geqslant 0$: $\cosh(x) \geqslant e^x/2$ and $\cosh(x) \leqslant e^x$. With the help of these, we can conclude that

$$F(k + 1)/F(k) < Q < 1 \qquad (20)$$

for a Q depending only on r_I. Thus, $F(k)$ decays exponentially, which proves Observation 1. Note that to maximize the entanglement between A and B, we have chosen the supremum of the maps better and better approximating the projection onto an infinitely entangled TMSS. Thus, for a specific GPEPS which is characterized by a fixed map V, the E_G is generally lower.
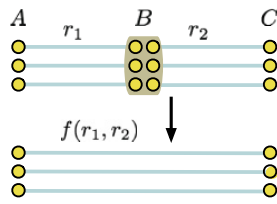
FIG. 2. (Color online) The situation referred to in Lemma 2. The strongest bonds before the projection are r_1 and r_2. The most significantly entangled bond has the strength $f(r_1, r_2)$.

This result has a remarkable consequence for Gaussian *quantum repeater lines*: It is not possible to build a 1D quantum repeater relying on Gaussian states, if only local measurements and no distillation steps are being used. We show in Sec. V that even non-Gaussian measurements cannot improve the performance. If one sticks to the Gaussian setting, also relying on complex networks does not remedy the exponential decay, as we see. Of course, non-Gaussian distillation schemes can be used to realize CV quantum repeater networks.

D. General graphs in arbitrary dimensions

One should suspect that the exponential decay of E_G is a special feature of the 1D situation and that higher dimensional graphs would eventually allow localization of a constant amount of entanglement. In this section we show that this is not the case. We first need a lemma which follows directly from our discussion of entanglement swapping.

Lemma 2. Collective operations on pure Gaussian states. Let ρ_{A,B_1} be a pure Gaussian state on $\mathcal{H}^{\otimes 2n}$ of n modes, and $\rho_{B_2,C}$ a pure Gaussian $\mathcal{H}^{\otimes 2m}$ state, where one part of each is held by A, B, and C, respectively (see Fig. 2). Let the maximum two-mode squeezing parameter be r_1 between A and B and r_2 between B and C. Then the maximum two-mode squeezing parameter achievable with a Gaussian projection in B between A and C is $f(r_1, r_2)$.

To prove this, we again use the fact that any two-party multimode pure Gaussian state can be transformed by local unitary Gaussian operations on both parties into a product of the TMSS [10,30]. This is nothing but the Gaussian version of the *Schmidt decomposition*. It hence does not restrict generality to start from this situation. As already noted, the best strategy for entanglement swapping between two pairs is a Gaussian Bell measurement, where the squeezing parameter changes according to f.

We now allow for global Gaussian operations on all subsystems belonging to B. We relax this situation to the following, where we allow for even more general operations: namely, a local Gaussian operation onto all modes of B, as well as onto all modes of A and C that are not the two modes that share the largest r. Clearly, this is a more general map than is actually considered in the physical situation. This, however, is exactly the situation already considered: an entanglement swapping scheme with an unentangled bystander. Hence, we again find that to project each pair onto a two-mode pure Gaussian state is optimal. For that, the sequence of projections better and better approximating an infinitely squeezed TMSS gives rise to the supremum. Hence, we have shown the
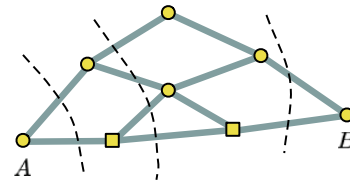
preceding result. Now we can prove a central result of this work.

Observation 2. Exponential decay of Gaussian localizable entanglement of a GPEPS in a general graph. Consider a GPEPS in a general graph with finite dimension and let A and B be two vertices of this graph. Then there exist constants $c_2, \xi_2 > 0$ such that

$$E_G(A, B) \leqslant c_2 e^{-d(A,B)/\xi_2}. \tag{21}$$

We take the shortest path between A and B—achieving the graph-theoretical distance $d(A, B)$—and denote its vertices $A, v_1, \ldots, v_{d(A,B)-1}, B$. We partition the graph in such a way that the boundaries do not intersect or touch each other and every vertex on the shortest path from A and B is contained in one region, which is called R_v (see Fig. 3). Again, we consider the situation of having TMSSs distributed in the graph between vertices sharing an edge—a general local Gaussian measurement on a GPEPS—so the GPEPS projection, now on several modes, followed by a specific single-mode Gaussian measurement, can only be less general than a general collective Gaussian measurement; thus, we again arrive at a bound to the localizable entanglement in the GPEPS.

Now we face exactly the situation to which Lemma 2 applies. In fact, in each step in each of the parts A, B, and C, we will have a collection of TMSSs, shared across the cut of the three regions. If r_{Av_1} is the strongest bond, in terms of the two-mode squeezing parameter, between R_A and R_{v_1}, and $r_{v_1v_2}$ is the strongest bond between R_{v_1} and R_{v_2}, then the strongest bond between R_A and R_{v_2} is given according to Lemma 2 by $f(r_{Av_1}, r_{Av_2})$. Now we can proceed exactly as in the proof of Theorem 1—and again, any uncorrelated bystanders will not help to improve the degree of entanglement—and thus show Theorem 2. This again has a consequence for quantum repeaters: Even when an arbitrary number of parties can share arbitrary many Gaussian entangled bonds, it is not possible to teleport quantum information over an arbitrary distance, as shown here.

In fact, using this statement, one can show that any impact of measurements in terms of a measurable signal is confined to a finite region in the graph, with I now being a subset of the graph, except from exponentially suppressed corrections. This region could be a poly-sized region in which the input to the computation is encoded. The readout of the quantum computation is then estimated from measurements in some region O, giving rise to a bit that is the result of the original decision problem to be solved by the quantum computation. From the decay of localizable entanglement, it is not difficult to show that the probability distribution
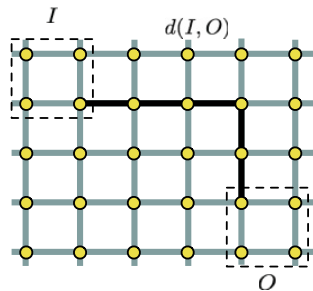

FIG. 3. (Color online) Partitioning of the graph according to the shortest path as described in the text. Sites drawn as squares are the those which lie on the shortest path connecting A and B.

FIG. 4. (Color online) Exponential decay of any influence of any measurements in region I on statistics of measurement outcomes in region O in the graph-theoretical distance $d(I,O)$ between the regions.

of this bit is unchanged by measurements in I, except from corrections that are exponentially decaying with $d(I,O)$ (see Fig. 4).

Note that concerning small-scale "proof-of-principle" applications, the arguments presented do *not* impose a fundamental restriction, as they apply only to the situation where entanglement distribution over an arbitrary number of modes (or repeater stations) is required. For any *finite* distance $d(A,B)$ and required entanglement $E(A,B)$, there exists a *finite* minimal squeezing $\lambda_{\min}$ which allows performance of the task. Only asymptotically will one necessarily encounter this situation. The result can equally be viewed as the impossibility of Gaussian quantum error correction in a measurement-based setting, complementing the results in Ref. [42].

E. Remarks on Gaussian repeater networks

These results of course also apply to general *quantum repeater networks*, where the aim is to end up with a highly entangled pair between any two points in the repeater network (see, e.g., Ref. [43] for a qubit version thereof). That is, in Gaussian repeater networks, one will also need non-Gaussian operations to make the network work, quite consistent with the findings in Refs. [9,31,32].

F. MBQC

The impossibility of encountering a localizable entanglement that is not exponentially decaying directly leads to a statement on the impossibility of using a GPEPS as a quantum wire. Such a wire should be able to perform the following task [17]: Assume that a single mode holds an unknown qubit in an arbitrary encoding; that is,

$$|\phi_{\text{in}}\rangle = \alpha|0_L\rangle + \beta|1_L\rangle. \tag{22}$$

This system is then coupled to a defined site A, the first site of the wire, of a GPEPS by a fixed in-coupling unitary operation which can in general be non-Gaussian. To complete the in-coupling operation, the input mode is measured in an arbitrary basis, where we also allow for probabilistic protocols; that is, the operation does not have to succeed for all measurement outcomes. Then one performs local Gaussian measurements on each of the modes. Then, at the end, one

expects the mode at a single site B to be in the state vector $|\phi_{\text{out}}\rangle = U|\phi_{\text{in}}\rangle$ (or at least arbitrarily close in trace norm) for any chosen $U \in \text{SU}(2)$. Note that the length of the computation, and therefore the position of output mode B, may vary and that the computational subspace can be left during the measurement. We want to stress that it is also possible to consider quantum wires which process qudits or even CV quantum information, where even on the logical level, information is encoded continuously. However, the capability of processing a qubit is clearly the weakest requirement. Thus, we address this situation only because the corresponding statements for other quantum wires immediately follow. With this clarification we can state the following lemma.

Observation 3. Impossibility of using Gaussian operations on arbitrary GPEPSs in general graphs for quantum wires. No GPEPS on any graph together with Gaussian measurements can serve as a perfect quantum wire for even a single qubit.

This is obvious from the previous considerations, as the measurements for the localizable entanglement and the incoupling operation commute, and clearly, the procedure is especially not possible for $U = \mathbb{1}$. The same argument, of course, also holds true in general graphs: No wire can be constructed from local Gaussian measurements in this sense, again for an exponential decay of the localizable entanglement. This observation is related to the decay of fidelity when performing CV quantum teleportation with squeezed vacuum states, as discussed in Ref. [44]. As mentioned, this statement can also be refined to having up to exponential corrections of finite-influence regions altogether.

V. NON-GAUSSIAN OPERATIONS

We now turn to our second main result, namely, that—under rather general assumptions which we detail below—Gaussian states defined on slabs of a finite width cannot be used as perfect primitives for resources for MBQC, even if non-Gaussian measurements are allowed for: Any influence of local measurements will again exponentially decay with distance.

More specifically, we first show that a 1D GPEPS cannot constitute a quantum wire in the sense of the definition in Sec. IV F extended to arbitrary measurements. This already covers all kinds of sophisticated encodings that can be carried by a single quantum wire, including ideas of "encoding qubits in oscillators" [16]. We then discuss the situation where an entire cubic slab of constant width is being used to encode a single quantum logical degree of freedom and find that the fidelity of transport will still decay exponentially. Not even using many modes and coupled quantum wires, possibly employing ideas of distillation, can this obstacle be overcome with local measurements alone. That is, we show that Gaussian states cannot be uplifted to serve as perfect universal resource states by measurements on finite slabs alone: Frankly, the finite squeezing present in the initial resources—although the state is pure and known—must be treated as a faulty state, and some full machinery of *fault tolerance* [14,15], which has yet to be developed for this kind of system, necessarily has to be applied even in the absence of errors. This contrasts quite severely with other limitations known for Gaussian quantum states. For example, while the distillation of entanglement is
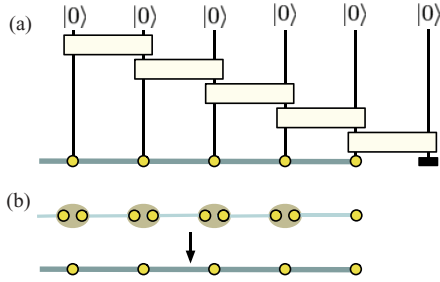
FIG. 5. (Color online) (a) Sequential preparation of a GMPS state: Each line represents a mode of a unitary tensor network, whereas each box stands for a Gaussian unitary. For a suitable choice of Gaussian unitaries, the resulting state is a Gaussian cluster state being prepared in the valence-bond construction (b).

not possible using Gaussian operations alone, non-Gaussian operations help to accomplish this task [45].

A. Sequential preparation of 1D Gaussian quantum wires

To make the statement, we first have to introduce another equivalent way of defining a GPEPS— or, specifically, a GMPS—in one dimension: It is easy to see that a GMPS with state vector $|\psi\rangle$ of N modes can be prepared as

$$|\psi\rangle = \langle\omega|_{N+1} \prod_{j=1}^{N} U^{(j,j+1)}|0\rangle^{\otimes(N+1)}, \qquad (23)$$

with identical Gaussian unitaries $U^{(j,j+1)}$ supported on modes $j, j+1$, depicted as gray bars in Fig. 5. This follows immediately from the original construction in Ref. [24] (see also Ref. [25]), translated into the Gaussian setting. A detailed study of sequentially preparable infinite-dimensional quantum systems with an infinite or finite bond dimension will be presented elsewhere.

B. Impossibility of transport by non-Gaussian measurements in one dimension: General considerations

We start by stating the main observation here: Frankly, even under general non-Gaussian measurements, transport along a 1D chain is not possible. We refer to the notions of both localizable entanglement and the *probability of transport*: This is the average maximum probability of recovering an unknown input state in a fixed subspace S of dimension at least $\dim(S) \geqslant 2$ which has been transported through the wire: Specifically, one asks for the maximum average success probability of a positive operator-valued measure applied to the output of the wire that leads to the identity channel up to a constant, where the average is taken with respect to all possible outcomes when performing local measurements transporting along the wire. We see that this probability decays exponentially with the distance between the input and the output site.

This decay follows regardless of the encoding chosen. Note that by no means do we require logical information to be contained in a certain fixed logical subspace along the

computation: Only in the first and last steps—when initially encoding quantum information or coupling to another logical qubit—do we ask for a fixed subspace. This logical subspace is even allowed to stochastically fluctuate along the computation dependent on measurement outcomes that are obtained in earlier steps of the computation.

Observation 4. Impossibility of using Gaussian 1D chains as quantum wires under general measurements. Let G be a one-dimensional GPEPS. Let S be either $S = \mathcal{H}$ or a subspace thereof. Then the probability of transport between any two sites A and B of the wire satisfies

$$p \leqslant c_3 e^{-d(A,B)/\xi_3} \qquad (24)$$

for suitable constants $c_3, \xi_3 > 0$. This implies that for any subsets of sites E_A and E_B and for fixed local subspaces, the entanglement between E_A and E_B that can be achieved by arbitrary local measurements of all sites except those contained in E_A and E_B is necessarily exponentially decaying in $d(E_A, E_B)$. This also means that, for any two sites A and B,

$$E_S(A,B) \leqslant c_4 e^{-d(A,B)/\xi_3} \qquad (25)$$

for some $c_4 > 0$ are constants, even if arbitrary local measurements are taken into account.

We now proceed in two steps. First, it is shown that there exists no subspace $S \in \mathcal{H}$ of dimension at least $\dim(S) \geqslant 2$ such that V_j can be chosen to be unitary, for all j for which $p_j > 0$ and

$$\langle\eta_j|U|\psi\rangle|0\rangle = p_j^{1/2} V_j|\psi\rangle \qquad (26)$$

for all $|\psi\rangle \in S$, where all U is the Gaussian unitary of the sequential preparation in Eq. (23), where the index of the mode, and also any label of tensor factors, is suppressed (see Fig. 6). $\{|\eta_j\rangle\}$ is an orthonormal basis of $\mathcal{H}$, with j labeling the respective outcome of the local measurement, possibly a continuous function. Because the computational subspace S is allowed to vary during the processing but must be invariant for the computation as a whole, we have to consider all N steps in the sequential preparation and all measurements together. For reasons of simplicity, we present the argument for a wire consisting of just two sites first and extend it afterward. We
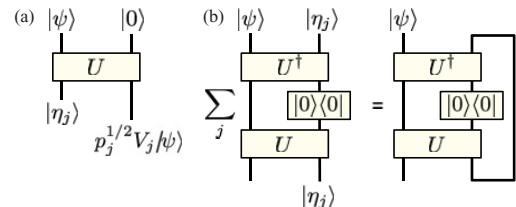

FIG. 6. (Color online) (a) Network representing a single step of a sequential preparation of a GMPS; (b) tensor network representation of $\langle\psi|\langle 0|U^\dagger(\mathbb{1} \otimes |0\rangle\langle 0|)U|0\rangle|\psi\rangle$.

define the operator

$$M = U^\dagger(\mathbb{1} \otimes |0\rangle\langle 0|)U \qquad (27)$$

and formulate the subsequent lemma.

Lemma 3. Conditions for nondecaying transport fidelity. A necessary condition for Eq. (26) to be satisfied is that

$$\langle\psi|\langle\eta_j|M|\psi\rangle|\eta_j\rangle = p_j \qquad (28)$$

for all j and all $|\psi\rangle \in S$, with $\sum_j p_j = 1$ and $\{|\eta_j\rangle\}$ forming a complete orthonormal basis of $\mathcal{H}$.

To see this, note that the fact that Eq. (26) holds true for each j for any $|\psi\rangle \in S$ means that

$$P_S\langle\eta_j|U|0\rangle P_S = p_j^{1/2}P_S, \qquad (29)$$

where P_S denotes the projection onto S. Using completeness of $\{|\eta_j\rangle\}$,

$$\sum_j |\eta_j\rangle\langle\eta_j| = \mathbb{1}. \qquad (30)$$

A moment of thought reveals that for any $|\phi\rangle \in S^\perp$, the latter denoting the orthogonal complement of S, one has that

$$P_S\langle\eta_j|U|\phi\rangle|0\rangle = 0. \qquad (31)$$

What is more,

$$\langle\phi|\langle\eta_j|U|0\rangle P_S = 0, \qquad (32)$$

again, for all $|\phi\rangle \in S^\perp$. This further means that (see Fig. 6)

$$\langle\psi|\langle\eta_j|U^\dagger(\mathbb{1} \otimes |0\rangle\langle 0|)U|\psi\rangle|\eta_j\rangle = \langle\psi|\langle\eta_j|M|\psi\rangle|\eta_j\rangle = p_j, \qquad (33)$$

which proves Lemma 3. Now summing over all measurement outcomes j in Eq. (33), which is the same as performing the partial trace (see Fig. 6) with respect to the second mode, we obtain

$$\langle\psi|\mathrm{tr}_2(U^\dagger(\mathbb{1} \otimes |0\rangle\langle 0|)U)|\psi\rangle = 1, \qquad (34)$$

which in turn implies, together with the preceding, that

$$P_S\mathrm{tr}_2[U^\dagger(\mathbb{1} \otimes |0\rangle\langle 0|)U]P_S = P_S. \qquad (35)$$

But this in turn means that the Gaussian operator $\mathrm{tr}_2[U^\dagger(\mathbb{1} \otimes |0\rangle\langle 0|)U]$ has at least two spectral values that are identical. Now it is only possible for a Gaussian operator to have two equal, nonzero spectral values if the spectrum is flat and corresponds to an operator that is not of trace class (related to "infinite squeezing" and "infinite energy," which was excluded due to the restriction to proper quantum states with finite energy).

We now extend the argument to a wire of arbitrary length. Toward this aim we denote the measurement basis on the kth site $\{|\eta_j^{(k)}\rangle\}$ and the corresponding probabilities $p_j^{(k)}$. Definition (27) is generalized to

$$M = [U^\dagger(\mathbb{1} \otimes |0\rangle)]^N[(\langle 0| \otimes \mathbb{1})U]^N. \qquad (36)$$

Condition (26) becomes

$$(\otimes_k \langle\eta_j^{(k)}|)U^{\otimes N}|\psi\rangle|0\rangle^{\otimes N} = \prod_k (p_j^{(k)})^{1/2}V_j^{(k)}|\psi\rangle, \qquad (37)$$

where $\prod_k V_j^{(k)}$ is unitary for all sequences of measurement outcomes and, furthermore, acts trivially on $S^\perp$. Modifying also Eqs. (32), (33), and (35) in a similar manner and using the completeness of the N measurement bases $\{|\eta_j^{(k)}\rangle\}$, we find that for Eq. (37) to hold, the Gaussian operator $O = \mathrm{tr}^N(M)$, where tr^N denotes the N-fold partial trace (or suitable tensor contraction), has two equal spectral values, which is not possible, as already mentioned, and thus, the first step in the proof is complete.

C. Impossibility of transport by non-Gaussian measurements in one dimension: Proving a gap

In the second step we now show that Observation 4 holds if Eq. (26) is not fulfilled. The problem of recovering an unknown state after propagation through the wire is equivalent to that of undoing a nonunitary operation. Obviously, it is a fundamental feature of quantum mechanics that it is not possible to implement a nonunitary linear transformation in a deterministic fashion. Since one does not have to correct for a nonunitary operation in each step, however, the technicality of the argument is related to the fact that we only have to undo an entire sequence of nonunitary Kraus operators once.

Assume that we aim to use our wire for the transport of a single pure qubit. After N steps of transport it will still be pure but, in general, distorted, due to the application of some nonunitary operator

$$V_J = V_{j_N}^{(N)} \cdots V_{j_1}^{(1)}, \qquad (38)$$

where $J = (j_1, \ldots, j_N)$ is an index reflecting the entire sequence of measurement outcomes on the N lattice sites. To recover the initial state, one has to apply an X_J such that

$$X_J V_J = c_J\mathbb{1}, \qquad (39)$$

with $c_j \in \mathbb{C}$. The success probability of this recovery operation, averaged over all measurement outcomes, is nothing but the *probability of transport*. It will decay exponentially in N whenever, for any k, at least a single $V_{j_k}^{(k)}$ is not unitary. The maximal average probability to undo random sequence V_J of Kraus operators is found to be

$$p_N = \max \mathrm{tr}(X_J V_J \rho V_J^\dagger X_J^\dagger), \qquad (40)$$

subject to

$$X_J^\dagger X_J = \mathbb{1}, \qquad (41)$$

$$X_J V_J = c_J\mathbb{1}. \qquad (42)$$

A moment of thought reveals that this probability of transport is then found to be

$$p_N = \sum_J \lambda_1[(V_J^\dagger V_J)^{-1}]^{-1} = \sum_J \lambda_n(V_J^\dagger V_J), \qquad (43)$$

where λ_1 (λ_n) denotes the largest (smallest) eigenvalue.

To show that Observation 4 is true if V_k is not proportional to a unitary matrix for at least one k can be shown by induction. Denoting, again, the operators applied by the measurements of the first N sites by V_J and the corresponding operators for

site $N + 1$ by $\{W_j\}$, we get from Eq. (43) that

$$p_{N+1} = \sum_{J,j} \lambda_n(V_J^\dagger W_j^\dagger W_j V_J). \tag{44}$$

Before we proceed, we note that it is possible to assume that all W_j and V_j are effective 2×2 matrices, corresponding to the situation where the computational subspace S does not change. If this is not the case, one can account for the fluctuation of the computational subspace by replacing $V_j \mapsto U_j V_j$ (and performing an analogous replacement for W_j) with a suitable unitary U_j. All arguments that follow do not depend on the choice of this unitary U_j. Key to the exponential decay is a lemma that is proven in the Appendix.

Lemma 4. Bound to eigenvalues of the sum of 2×2 matrices. For any positive $A, B \in \mathbb{C}^{2\times2}$ with $[A, B] \neq 0$, there exists a $\delta > 0$ such that

$$\lambda_2(A + B) \geqslant \lambda_2(A) + \lambda_2(B) + \delta. \tag{45}$$

If there exists at least one pair (i, j) for which

$$[W_i^\dagger W_i, W_j^\dagger W_j] \neq 0, \tag{46}$$

then also

$$[V_J^\dagger W_i^\dagger W_i V_J, V_J^\dagger W_j^\dagger W_j V_J] \neq 0, \tag{47}$$

and we can apply Lemma 4 directly to Eq. (44). If, in contrast,

$$[W_i^\dagger W_i, W_j^\dagger W_j] = 0 \tag{48}$$

for all pairs (i, j), all $W_i^\dagger W_i$ can be simultaneously diagonalized. This means that we can—without loss of generality—assume that

$$W_i^\dagger W_i = \text{diag}(\xi_i, \zeta_i). \tag{49}$$

Because a nonunitary W_i exists by assumption, $\min\{|\xi_i - \zeta_i| : i = 1, 2\} > 0$. In both cases we are provided with a $\nu < 1$ such that

$$p_{N+1} \leqslant \nu \sum_J \lambda_2(V_J^\dagger V_J) = \nu p_N, \tag{50}$$

where we have used the completeness relation

$$\sum_j W_j^\dagger W_j = \mathbb{1}. \tag{51}$$

This observation gives rise to the anticipated gap that proves the exponential decay of the probability of transport and, therefore, to Observation 4. The exponential decay of the subspace localizable entanglement follows directly: If there was a nondecaying localizable entanglement, this could be used to transport with a high recovery probability, in contrast to what we have shown. If this were not the case, one could use the wire to distribute entanglement, which is obviously not possible.

D. Impossibility of transport by non-Gaussian measurements in one dimension: Concluding remarks

Note, finally, that even though we have presented Observation 4 for local projective measurements—which suits the paradigm of measurement-based computing—the argument obviously holds true for positive operator-valued measurements. The proof is completely analogous, with $\sum_j |\eta_j\rangle\langle\eta_j| = \mathbb{1}$ being replaced by a more general resolution of the identity.

This argument shows that 1D GPEPSs cannot be used as quantum wires even when allowing for arbitrary non-Gaussian local measurements. Note that for this argument to hold, completeness of the measurement bases are indeed necessary: For single outcomes, the condition of the output being up to a constant unitarily equivalent to the input can well be achieved also for matrices having a different structure; but then one cannot assure that this is true for each outcome j of the measurement. This, however, is required to faithfully transport quantum information. If we allow for a finite rate of failure outcomes j in individual steps, then the overall probability of success will asymptotically again become 0 at an exponential rate.

E. Gaussian cluster states under arbitrary encodings and in higher dimensional lattices

One might wonder whether this limitation can be overcome if a large number of physical modes of a higher dimensional lattice are allowed to carry logical information. The same argument, actually, can be applied to a $k \times k \times \cdots \times k \times n$ cubic slab, as a subset of a D-dimensional cubic lattice, where one aims at transporting along the last dimension, with local measurements at each site (Fig. 7). In fact, contracting any dimension except from the last—so summing over all joint indices—one arrives at a GMPS with a bond dimension that is exponential in k. This, however, is a constant. This situation is hence again covered by a GMPS, as long as one allows for more than one physical mode and more than one virtual mode per site. Since the argument in Sec. V B does not make use of the fact that we have only a single virtual and physical mode per site: only that now $|0\rangle^{\otimes(k^{(D-1)})}$ are being fed into the sequential preparation.

Observation 5. Exponential decay of subspace localizable entanglement in a higher dimensional lattice. Let G be a one-dimensional GPEPS, and A and B two sites in a $k \times k \times \cdots \times k \times n$ slab as a subset of a D-dimensional cubic lattice, and denote by i, j the last coordinate of sites A and B. Then

$$E_S(A, B) \leqslant c_4 e^{-d(i,j)/\xi_4}, \tag{52}$$

where $c_4, \xi_4 > 0$ are constants, even if arbitrary local measurements are taken into account.

So even encodings in higher dimensional Gaussian cluster states do not alter the situation that one cannot transport along
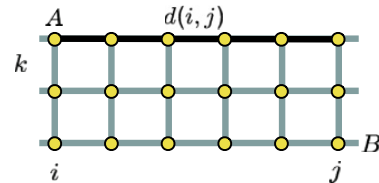

FIG. 7. (Color online) A slab of a $k \times n$ lattice, aiming at using the second dimension as a quantum wire for quantum computation. Again, the probability of transport between A and B decays exponentially with the distance along the last dimension.

a given dimension, if one wants to think of such slabs as perfect primitives being used in a universal quantum computing scheme.

F. Role of error correction and fault tolerance

Observations 2 and 5 show that, under mild conditions, Gaussian cluster states need *not be used as* or *made almost-perfect resources by* local measurements alone. This constitutes a significant challenge for MBQC with Gaussian cluster states but does not rule out this possibility. In this section, we briefly comment on ways that might allow one to overcome the limitations identified here.

Clearly, it is very much conceivable that this observation may again be overcome by concatenated encoding in fault-tolerant schemes, effectively in slabs whose width scales with the length of the computation: Rather, at the level of finite encodings, the resource cannot be uplifted to a perfect resource. The situation encountered here—having pure Gaussian states—hence has some similarity to *noisy finite-dimensional cluster states* built with *imperfect operations* [14,15]. Considering the preparation of the quantum wire and the transport by local measurements as a sequence of teleportations with not fully entangled resources, this means that every step adds a given amount of noise to the quantum information. In finite-dimensional schemes, if this noise corresponds to an error rate below the fault tolerance, a nested encoding with an error correction code allows one to perform computations. The size of the code grows polynomially with the size of the circuit one wishes to implement. In addition to this intrinsic error, any physical implementation will, of course, also suffer from experimental errors which must also be compensated by error correction schemes. Thus, the combined error rate must be below the fault-tolerance threshold. It is therefore possible that recognizing all finite squeezings as full quantum errors—which has to be done in the light of the results of the present work—and using suitable concatenated encodings over polynomially many slabs, there exists a finite squeezing allowing for full universal quantum computation with eventual polynomial overhead. The question whether schemes such as these—or ones where suitable polynomially sized complex structures are "pinched" out of a large lattice—that are universal can be constructed remains a challenging and interesting open question.

G. Ideas on percolation

One possible way forward toward the goal of achieving a fully universal resource under local non-Gaussian measurements is to think of first performing local measurements at each site, aiming at filtering an imperfect qubit, $\mathbb{C}^2$ cluster from a Gaussian cluster state. Ideally, one would arrive at the situation on, say, a cubic lattice of some dimension, where one could extract a *graph state* [46] corresponding to having an edge between nearest neighbors with some finite probability. If this probability p_s is sufficiently high—higher than the appropriate threshold for *edge percolation*—and if one can ensure suitable independence, an asymptotically perfect cluster on a renormalized lattice can be obtained [47–49]. When trying to identify such percolation schemes, one does

not have to rely solely on classical percolation schemes, but can also make use of more general repeater-type schemes as in Ref. [50], referred to as *quantum percolation* (see also Ref. [48]). To identify such maps, either classical or quantum, however, appears to be a very challenging task.

One might also ask whether TMSS bonds as such can be transformed into suitable maximally entangled pairs of $\mathbb{C}^2 \otimes \mathbb{C}^2$ systems. This, however, clearly is the case. Again applying a result for finite-dimensional systems to infinite-dimensional ones by making use of appropriate nets of Hilbert spaces, one finds that given a state vector $|\psi_\lambda\rangle$ of a TMSS of some squeezing parameter $\lambda > 0$, the transformation $|\psi_\lambda\rangle$ to $(|0,0\rangle + |1,1\rangle)/\sqrt{2}$ is possible with a generalized local filtering on A only, together with a suitable unitary in B, with a probability of success of [51,52]

$$p = \min[1, 2(1 - \lambda^2)]. \tag{53}$$

Hence, whenever $\lambda \geqslant 1/\sqrt{2}$, this transformation can be done deterministically. This has interesting consequences for quantum repeaters. The protocol performing the transformation

$$|\psi_\lambda\rangle_{A,B} \mapsto \frac{1}{\sqrt{2}} (|0,0\rangle + |1,1\rangle) \tag{54}$$

can be implemented by combining A with an ancillary system C, performing a joined unitary transform on A,C, measuring C, and applying another unitary gate on B classically conditioned on the measurement result.

But even if $\lambda < 1/\sqrt{2}$, one can still distill a resource from a collection of TMSSs distributed in a graph, performing an argument involving percolation here. This, however, merely shows that Gaussian states as such can be resources for information processing. Most importantly, this is not the resource anticipated, so not the actual GPEPS, but a collection of suitable TMSS. Thus, non-GPEPS projections cannot be implemented with linear optics without a massive overhead. Finally, the eventually created qubit cluster state would be obtained in a *single-rail* representation where measurements in the superposition bases, which are needed for the actual computation, are experimentally very difficult and require additional photons. So the question of actual universality of the Gaussian cluster state, under all fair meaningful ways of defining a set of rules, remains an interesting and challenging question.

H. Remarks on 1D Gaussian quantum repeaters

We finally briefly reconsider the question of a quantum repeater setting based on general non-Gaussian operations. We have shown that it is not possible to obtain a finitely entangled state for an arbitrary long 1D GPEPS. However, what is also true at the same time is that a sequential repeater scheme based on sufficiently entangled TMSSs *before* the PEPS projection does yield a nondecaying entangled bond between the end points. That is, using only projective local measurements of each of the sites, one can transform a collection of distributed TMSSs in a 1D setting into a maximally entangled qubit pair shared between the end sites. To show this, it suffices to revisit the situation for three sites, as the general statement on N sites follows immediately by iteration.

Now consider the quantum repeater setting and assume for simplicity that we already have a qubit Bell pair, $|\phi\rangle_{A,B_1} = (|0,0\rangle + |1,1\rangle)/\sqrt{2}$, which we want to swap through a TMSS $|\psi_\lambda\rangle_{B_2,C}$ with $\lambda \geqslant 1/\sqrt{2}$. We can use the higher, unoccupied Fock levels of the state vector $|\phi\rangle_{A,B_1}$ as an ancilla to transform $|\psi_\lambda\rangle_{B_2,C}$ according to Eq. (54). As the final unitary on C after local operations and classical communication with one-way classical communication does not change the entanglement, we can also omit it. As the unitary, the ancilla measurement, and the final Bell measurement on B_1, B_2 are equivalent to a single projective measurement on B_1, B_2, it is possible to swap entanglement through a physical TMSS perfectly. Needless to say, this will be a highly non-Gaussian, complicated operation and will not overcome the limitation of Gaussian cluster states discussed.

VI. DISCUSSION AND SUMMARY

In this article, we have assessed the requirements for possible architectures when using Gaussian states as resources for MBQC and for entanglement distribution by means of quantum repeater networks. Using a framework of GPEPSs, we have shown that under Gaussian measurements only, the localizable entanglement decays exponentially with the distance in arbitrary graphs. This rules out the possibility of processing or even transporting quantum information with Gaussian measurements only.

The preceding results also show that Gaussian cluster states—under mild conditions of the encoding of logical information in slabs, rather than general encodings in the entire lattice—cannot be used as or made perfect universal resources for MBQC. No information can be transmitted beyond a certain influence region, and hence, no arbitrarily long computation can be sustained. Now if one allows for a higher energy, and hence larger two-mode squeezing, in the resource states, this influence region will become larger. In other words, small-scale implementations as proof-of-principle experimental realizations of such an idea will be entirely unaffected by this: Any state with finite energy will constitute some approximation of the idealized improper state having infinite energy, and its outcomes in measurements will approximate the idealized ones. However, with this state, one could not go ahead with an arbitrarily long computation. This observation shows that Gaussian cluster states are fine examples of states that eventually allow for the demonstration of the functioning of a CV quantum computer, possibly realized using the many modes available in a *frequency comb* [5–7].

Also, we have discussed the requirements for fault tolerance and quantum error correction for such schemes, yet to be established, in that any finite squeezings essentially have to be considered full errors in a concatenated encoding scheme. This work motivates further studies of the fault tolerance of systems with a finite-dimensional logical encoding in infinite-dimensional systems. But it also strongly suggests that it could be a fruitful enterprise to further at alternative CV schemes not directly involving Gaussian states, but other relatively feasible classes of states, such as coherent superpositions of a few Gaussian states like the so-called cat states, which have turned out to be very useful within another computation paradigm [53]. We hope that this article will contribute to clarifying the requirements that any architecture eventually must meet based on the interesting idea of doing quantum computing by performing local measurements on Gaussian or non-Gaussian states of light.

APPENDIX: PROOF OF LEMMA 4

Let $A, B \in \mathbb{C}^{2\times2}$, with $A, B \geqslant 0$. We set

$$c = \frac{\|A^{1/2}B^{1/2}\|^2}{\|A\|\|B\|}. \tag{A1}$$

The inequality $c \leqslant 1$ follows directly from the submultiplicativity of the operator norm, while equality holds if and only if A and B commute. Rewriting

$$\lambda_n(A + B) = \text{tr}(A + B) - \lambda_1(A + B)$$
$$= \text{tr}(A + B) - \|A + B\|, \tag{A2}$$

we can now use a sharpened form of the triangle inequality for the operator norm of 2×2 matrices in Ref. [54] to obtain

$$\lambda_2(A + B) = \text{tr}(A + B) - \|A + B\|$$
$$\geqslant \text{tr}(A + B) - \tfrac{1}{2}(\|A\| + \|B\|)$$
$$+ \tfrac{1}{2}(\|A\| - \|B\|)^2 + 4\|A^{1/2}B^{1/2}\|^2)^{1/2}. \tag{A3}$$

If now $c < 1$, then there exists a $\delta > 0$ such that

$$\lambda_2(A + B) \geqslant \text{tr}(A + B) - ((\|A\| - \|B\|)^2$$
$$+ 4\|A\|\|B\|)^{1/2} + \delta$$
$$= \text{tr}(A + B) - (\|A\| + \|B\|) + \delta \tag{A4}$$
$$= \lambda_2(A) + \lambda_2(B) + \delta, \tag{A5}$$

which proves Lemma 4.

[1] E. Knill, R. Laflamme, and G. J. Milburn, Nature (London) **409**, 46 (2001).

[2] R. Raussendorf and H. J. Briegel, Phys. Rev. Lett. **86**, 5188 (2001).

[3] D. E. Browne and T. Rudolph, Phys. Rev. Lett. **95**, 010501 (2005).

[4] D. Gross, K. Kieling, and J. Eisert, Phys. Rev. A **74**, 042343 (2006).

[5] N. C. Menicucci, P. van Loock, M. Gu, C. Weedbrook, T. C. Ralph, and M. A. Nielsen, Phys. Rev. Lett. **97**, 110501 (2006).

[6] N. C. Menicucci, S. T. Flammia, and O. Pfister, Phys. Rev. Lett. **101**, 130501 (2008).

[7] S. T. Flammia, N. C. Menicucci, and O. Pfister, J. Phys. B **42**, 114009 (2009).

[8] R. Ukai, J.-I. Yoshikawa, N. Iwata, P. van Loock, and A. Furusawa, Phys. Rev. A **81**, 032315 (2010).

[9] G. Giedke and J. I. Cirac, Phys. Rev. A **66**, 032316 (2002).

[10] G. Giedke, J. Eisert, J. I. Cirac, and M. B. Plenio, Quantum Inf. Comput. **3**, 211 (2003).

[11] J. Eisert and M. B. Plenio, Int. J. Quantum Inf. **1**, 479 (2003).

[12] S. L. Braunstein and P. van Loock, Rev. Mod. Phys. **77**, 513 (2005).

[13] G. Adesso and F. Illuminati, J. Phys. A **40**, 7821 (2007).

[14] M. A. Nielsen and C. M. Dawson, Phys. Rev. A **71**, 042323 (2005).

[15] R. Raussendorf and J. Harrington, Phys. Rev. Lett. **98**, 190504 (2007).

[16] D. Gottesman, A. Kitaev, and J. Preskill, Phys. Rev. A **64**, 012310 (2001).

[17] D. Gross and J. Eisert, Phys. Rev. Lett. **98**, 220503 (2007).

[18] D. Gross, J. Eisert, N. Schuch, and D. Perez-Garcia, Phys. Rev. A **76**, 052315 (2007).

[19] D. Gross and J. Eisert, Phys. Rev. A **82**, 040303(R) (2010).

[20] G. K. Brennen and A. Miyake, Phys. Rev. Lett. **101**, 010502 (2008).

[21] X. Chen, B. Zeng, Z. C. Gu, B. Yoshida, and I. L. Chuang, Phys. Rev. Lett. **102**, 220501 (2009).

[22] C. E. Mora, M. Piani, A. Miyake, M. Van den Nest, W. Dür, and H. J. Briegel, Phys. Rev. A **81**, 042315 (2010).

[23] M. Keyl, D. Schlingemann, and R. F. Werner, e-print arXiv:quant-ph/0212014.

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

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

[26] M. A. Martin-Delgado, M. Roncaglia, and G. Sierra, Phys. Rev. B **64**, 075117 (2001).

[27] F. Verstraete and J. I. Cirac, e-print arXiv:cond-mat/0407066.

[28] F. Verstraete and J. I. Cirac, Phys. Rev. A **70**, 060302(R) (2004).

[29] N. Schuch, J. I. Cirac, and M. M. Wolf, Commun. Math. Phys. **267**, 65 (2006).

[30] A. Botero and B. Reznik, Phys. Rev. A **67**, 052311 (2003).

[31] J. Eisert, S. Scheel, and M. B. Plenio, Phys. Rev. Lett. **89**, 137903 (2002).

[32] J. Fiurášek, Phys. Rev. Lett. **89**, 137904 (2002).

[33] P. van Loock, C. Weedbrook, and M. Gu, Phys. Rev. A **76**, 032321 (2007).

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

[35] M. Popp, F. Verstraete, M. A. Martín-Delgado, and J. I. Cirac, Phys. Rev. A **71**, 042306 (2005).

[36] J. Fiurášek and J. L. Mišta, Phys. Rev. A **75**, 060302 (2007).

[37] J. L. Mišta and J. Fiurášek, Phys. Rev. A **78**, 012359 (2008).

[38] J. Eisert, Ph.D. thesis, University of Potsdam, 2001.

[39] G. Vidal and R. F. Werner, Phys. Rev. A **65**, 032314 (2002).

[40] M. B. Plenio, Phys. Rev. Lett. **95**, 090503 (2005).

[41] P. van Loock, Fortsch. Phys. **50**, 1177 (2002).

[42] J. Niset, J. Fiurášek, and N. J. Cerf, Phys. Rev. Lett. **102**, 120501 (2009).

[43] S. Perseguers, J. I. Cirac, A. Acin, M. Lewenstein, and J. Wehr, Phys. Rev. A **77**, 022308 (2008).

[44] S. L. Braunstein and H. J. Kimble, Phys. Rev. Lett. **80**, 869 (1998).

[45] J. Eisert, D. E. Browne, S. Scheel, and M. B. Plenio, Ann. Phys. (NY) **311**, 431 (2004).

[46] M. Hein, J. Eisert, and H. J. Briegel, Phys. Rev. A **69**, 062311 (2004).

[47] K. Kieling, T. Rudolph, and J. Eisert, Phys. Rev. Lett. **99**, 130501 (2007).

[48] K. Kieling and J. Eisert, *Quantum and Semi-classical Percolation and Breakdown in Disordered Solids* (Springer, Berlin, 2009), pp. 287–319.

[49] D. E. Browne, M. B. Elliott, S. T. Flammia, S. T. Merkel, A. Miyake, and A. J. Short, New J. Phys. **10**, 023010 (2008).

[50] A. Acin, J. I. Cirac, and M. Lewenstein, Nature Phys. **3**, 256 (2007).

[51] G. Vidal, Phys. Rev. Lett. **83**, 1046 (1999).

[52] M. A. Nielsen, Phys. Rev. Lett. **83**, 436 (1999).

[53] T. C. Ralph, A. Gilchrist, G. J. Milburn, W. J. Munro, and S. Glancy, Phys. Rev. A **68**, 042319 (2003).

[54] F. Kittaneh, J. Operator Theory **48**, 95 (2002).

Chapter 4

Efficient measurement-based quantum computing with continuous-variable systems

Efficient measurement-based quantum computing with continuous-variable systems

M. Ohliger[1,2] and J. Eisert[1]

[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*
(Received 16 December 2011; published 21 June 2012)

We present strictly efficient schemes for scalable measurement-based quantum computing using continuous-variable systems: These schemes are based on suitable non-Gaussian resource states, ones that can be prepared using interactions of light with matter systems or even purely optically. Merely Gaussian measurements such as optical homodyning as well as photon counting measurements are required, on individual sites. These schemes overcome limitations posed by Gaussian cluster states, which are known not to be universal for quantum computations of unbounded length, unless one is willing to scale the degree of squeezing with the total system size. We establish a framework derived from tensor networks and matrix product states with infinite physical dimension and finite auxiliary dimension general enough to provide a framework for such schemes. Since in the discussed schemes the logical encoding is finite dimensional, tools of error correction are applicable. We also identify some further limitations for any continuous-variable computing scheme from which one can argue that no substantially easier ways of continuous-variable measurement-based computing than the presented one can exist.

I. INTRODUCTION

To realize a quantum computer in the circuit model, it is crucial to have precise control over each of the carriers of quantum information. In addition to keeping the quantum state of the computer protected from unavoidable noise induced by the environment, it is necessary to implement suitable quantum gates, usually in the form of one- and two-qubit gates. The latter is of particular difficulty, especially when optical quantum systems are used because their interaction is either weak or merely induced by measurements.

The paradigm of measurement-based quantum computing (MBQC) as pioneered by Raussendorf and Briegel [1,2] and substantially generalized by Gross and Eisert [3,4] allows us to get rid of the necessity of performing unitary operations to implement a quantum circuit. Instead, the actual computation is performed by preparing a multipartite entangled state, the resource, in a first step followed by adaptively chosen local measurements on this resource. The important improvement stems from the fact that the resource is universal, i.e., it can be prepared independently of the algorithm one wants to perform. This means that the presumably difficult step, the one which involves entangling operations, can be performed off-line. This resource-preparation may also be probabilistic as it is possible to wait with the implementation of the algorithm until the resource is available. What is more, individual addressing in the final read-out step is also required in the circuit model, so this step does not add any further difficulty to the scheme.

Quantum computing based on continuous-variable (CV) quantum optical systems differs from the more conventional notion based on their discrete analogs in not making use of single photons as the carriers of quantum information [5,6]. Instead, it relies to a large extent on Gaussian states and their manipulation. The notable advantage is that Gaussian states are easier to prepare in the laboratory, the corresponding interactions are often stronger and easier to accomplish, and some measurements, e.g., homodyne detection, can be

performed with an efficiency substantially surpassing the one of single-photon measurements. However, quantum information protocols using Gaussian states only, Gaussian operations, and Gaussian measurements suffer from serious limitations as in this setting neither entanglement distillation [7–9] nor error correction against Gaussian errors [10] is possible. Also, since one can easily efficiently keep track of first and second moments, any Gaussian evolution of Gaussian states can be efficiently simulated on a classical computer [11], clearly ruling out as the possibility of universal quantum computing.

MBQC based on Gaussian resource states has been extensively discussed in the literature [12–16] due to several features: They are a direct generalization of the well-known qubit cluster state to the continuous-variable regime, they allow for universal quantum computing with Gaussian measurements and a single non-Gaussian one, and they can be prepared with present-day experimental techniques, as already demonstrated. However, if they are not formed from idealized infinitely squeezed states (ones that are not contained in Hilbert space and would require infinite energy in preparation) but rather from physical states possessing only finite squeezing, they suffer from exponentially decaying localizable entanglement. This limitation, which applies to any Gaussian resource state irrespectively of the permitted class of measurements [17], implies that full error correction and the machinery of fault-tolerant quantum computing [18,19], in parts yet to be developed for such continuous-variable systems, appears to be necessary even when both state preparation and measurements are perfect. Notably, those restrictions only apply when trying to perform a quantum computation of an *unbounded* length. For any *finite* length, there exists, for any required accuracy, a physical cluster state such that any quantum operation up to this length is possible with this accuracy. Because the widely assumed superior power of quantum computers compared to classical ones manifests itself most prominently in the scaling of the runtime with the problem size, the situation of a quantum

computation with unbounded length is in the focus of attention in this work.

In addition to schemes based on Gaussian cluster states, two more classes of schemes relying on CV-quantum optics have been proposed. The first one is based on superpositions of coherent states which are called Schrödinger cat states or, when they have very low amplitude, "kitten" states [20]. They can be, in approximation, created by subtracting photons from squeezed states and allow for universal quantum computing by passive operations such as beam splitters and photon-counting detection only. However, they also suffer from quite severe problems: The probabilistic nature of the quantum operations stemming from the use of nonoverlapping basis elements seems to be the most fundamental one, giving rise to significant overheads. The second approach combines the advantageous features of both discrete and continuous-variable optics [21,22]. In these schemes, quantum information is carried by qubits and single-qubit operations are performed directly on them. Two-qubit operations, in contrast, are performed by letting both qubits interact with a strong CV-mode, the qubus.

With this article, we pursue two different, but complementary, goals: On the one hand, we aim at clarifying the boundary between settings where CV-MBQC is possible and such situations where this is not the case. For this reason, we will develop a general framework capable of describing quantum computation in the measurement-based paradigm, regardless of the dimension of the carriers of quantum information. Within this picture, we can identify some serious limitations giving rise to challenges that have to be overcome. On the other hand, we introduce a strictly efficient scheme relying only on a simple controlled rotation, which can be realized by an atom-light interaction of the Jaynes-Cummings type or purely optically by the Kerr effect, for the creation of the resource. On the measurement side, we require beam splitters, phase-space displacers, and photon counting measurements. While this scheme is not fleshed out in all detail of its concrete physical implementation, it should be clear that quantum optical implementations of such ideas are conceivable.

The present article is organized as follows: First, we discuss what operations are possible with continuous-variable quantum systems and introduce a framework based on matrix product states (MPS) to describe MBQC in a general setting. After a discussion of the properties which a CV-MBQC scheme needs to posses in order to be called theoretically efficient, we show the problems of achieving those requirements with Gaussian measurements on non-Gaussian resource states. Last, we provide an example for a feasible scheme and discuss in detail how efficient MBQC can be performed in this situation.

II. FEASIBLE PRIMITIVES FOR CV QUANTUM COMPUTING

When performing tasks of quantum information with continuous-variable quantum optics, different classes of operations are considered to be of different difficulty. This is also true for measurements where the achievable efficiency greatly differs among the various methods.

A. Gaussian operations

Before continuing with the discussion, we remind the reader of some basic properties of Gaussian states and operations while also taking the opportunity to set the notation. We consider a single light mode. The energy eigenvectors of the unit oscillator are denoted by $|n\rangle$ with $n = 0, 1, \ldots$. The annihilation operator $\hat{a}$ acts on them according to $\hat{a}|n\rangle = \sqrt{n}|n-1\rangle$. The commutator relation with its adjoint, the creation operator, is $[\hat{a}, \hat{a}^\dagger] = 1$, setting $\hbar = 1$. The eigenvectors of the photon-number operator $\hat{n} = \hat{a}^\dagger \hat{a}$ with $\hat{n}|k\rangle = k|k\rangle$ are called the number states or Fock states. One can now define the canonical operators or quadratures as

$$\hat{q} = \frac{1}{\sqrt{2}}(\hat{a} + \hat{a}^\dagger), \tag{1}$$

$$\hat{p} = -\frac{i}{\sqrt{2}}(\hat{a} - \hat{a}^\dagger), \tag{2}$$

and for $\theta \in [0, 2\pi]$ the family of rotated quadrature operators

$$\hat{q}_\theta = \hat{q}\cos(\theta) + \hat{p}\sin(\theta). \tag{3}$$

The latter family of observables is the one that captures homodyne detection [23]. Gaussian unitaries are the ones which can be written as $\hat{U} = e^{i\hat{H}(\hat{q},\hat{p})}$, where $\hat{H}$ contains no terms in higher than quadratic order in $\hat{q}$ and $\hat{p}$ (or, equivalently, in $\hat{a}$ and $\hat{a}^\dagger$). There are three classes of Gaussian single-mode unitary operators into which all Gaussian unitary gates can be decomposed. The first ones are corresponding to the application of the displacement operator $\hat{D}(\alpha) = \exp(\alpha\hat{a}^\dagger - \alpha^*\hat{a})$ for $\alpha \in \mathbb{C}$. Such a transformation is reflected in the Heisenberg picture by a map of the form

$$\hat{q} \mapsto \hat{q} + \sqrt{2}\,\mathrm{Re}\,\alpha, \tag{4}$$

$$\hat{p} \mapsto \hat{p} + \sqrt{2}\,\mathrm{Im}\,\alpha. \tag{5}$$

Optically, such a transformation can be realized by mixing the mode with a second mode, which is in a strong coherent state, on a beam splitter in the limit of vanishing reflectivity. In the second one are the transformations generated by the clockwise rotation operator $R(\theta) = \exp(-i\theta\hat{n})$, which can be realized by a phase shifter, acting on the canonical operators as $\hat{q} \mapsto \hat{q}_\theta$, $\hat{p} \mapsto \hat{q}_{\theta+\pi/2}$, while in the last one are those generated by the squeezing operator

$$\hat{S}(\xi) = \exp\left[\frac{r}{2}(\hat{a}^2 - \hat{a}^{\dagger 2})\right], \tag{6}$$

acting as $\hat{q} \mapsto e^r\hat{q}$, $\hat{p} \mapsto e^{-r}\hat{p}$. The single-mode Gaussian unitary operations form a (noncompact) group which we denote by $\mathbb{U}_G$. To complete the set of operators, which are necessary to implement arbitrary *multimode* Gaussian operations, we introduce the (absorption-free) beam splitter acting on two modes 1 and 2 by

$$\hat{B} = \exp\left[\frac{\theta}{2}(\hat{a}_1^\dagger\hat{a}_2 - \hat{a}_1\hat{a}_2^\dagger)\right], \tag{7}$$

where $t = \cos(\theta/2)$ and $r = \sin(\theta/2)$ are the transmission and reflection coefficient, respectively.

The Gaussian operations can be divided into two classes: Phase shifters and beam splitters do not change the total number of photons and are called passive operations. To

implement a squeezer or a displacer, on the other hand, one does need additional photon sources which makes them more difficult to realize.

B. Other Hamiltonian building blocks

There are also non-Gaussian operations which are within the realm of present experimental techniques: The cross Kerr effect gives rise to a nonlinear coupling between two modes; its Hamiltonian reads

$$\hat{H} = \chi \hat{n}_1 \otimes \hat{n}_2. \tag{8}$$

It can be realized by transmitting two modes together through an optically nonlinear medium. As nonlinear optical effects are weak and absorption needs to be kept low, the achievable effective coupling strength is quite small, i.e., $t\chi \ll 1$, where t is the interaction time. For the most relevant situation, where one of the modes carries a state with many photons while the other one is in the single-photon regime, the Hamiltonian (8) has already been experimentally implemented to perform quantum nondemolition measurement (QND) [24]. The related Hamiltonian

$$\hat{H} = \chi |1\rangle\langle 1| \otimes \hat{n} \tag{9}$$

for a positive χ describes the coupling of a two-level atom to a light mode in the dispersive limit of the Jaynes-Cummings model. In this situation, the effective coupling strength can be increased by placing the mode and the atom inside an optical cavity. A particularly strong interaction can be achieved with a technique known as electromagnetically induced transparency (EIT), which is routinely used to exchange quantum information between between light modes and atomic vapors and for which also experiments with single atoms exist [25,26].

C. Measurements

We now turn to the kind of measurements that we will consider feasible for the purposes of this work. The most important Gaussian measurement scheme is homodyne detection, which correponds to the observable $\hat{q}_\theta$ in Eq. (3). It is realized by combining the mode with a strong laser, called the local oscillator, in an interferometer, measuring the intensities on both output ports and subtracting the results. Another important type of Gaussian measurement is eight-port homodyning corresponding to a direct measurement of the Q function, i.e., $Q_\rho(\alpha) = \langle\alpha|\rho|\alpha\rangle$ for $\alpha \in \mathbb{C}$, where

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_n \frac{\alpha^n}{\sqrt{n!}} |n\rangle \tag{10}$$

are the state vectors of the non-orthogonal and overcomplete coherent states [23].

Non-Gaussian measurements are in many instances more difficult to realize than Gaussian ones, and usually with significantly lower detection efficiencies. The most feasible instance of a non-Gaussian measurement reflects a single-photon detector which can distinguish only between the absence of photons and the presence of one or more photons. The corresponding POVM elements are $|0\rangle\langle 0|$ and $\mathbb{1} - |0\rangle\langle 0|$. A photon-number resolving detector is more challenging to implement but, with time-multiplexing [27] or superconducting

nanowires [28], it is possible to perform photon counting for the first few number states with reasonable efficiency.

III. FRAMEWORK FOR CV-MBQC

In this section, we introduce a general framework to describe measurement-based quantum computing. A MBQC scheme consists of two ingredients: First, a resource state and, second, a set of possible, or allowed, local measurements. We start with a description of quantum wires which are used for single-qudit processing [29] and discuss their coupling to fully universal resources afterward.

A. Matrix product states

The formalism of matrix product states, which was originally introduced to describe a certain class of entangled one-dimensional many-body states [30,31], is extremely useful to capture the essentials of quantum computing in the measurement-based model [4]. We start by describing one-dimensional continuous-variable quantum wires which can be used to carry a single qubit of quantum information. The techniques we develop are independent of the actual physical implementation but we will often describe them with terms of quantum optics and have such a system in mind.

We consider a chain of L quantum systems with dimension d_p which are called lattice sites. We also allow for the situation $d_p = \infty$ which describes CV light modes. We say that a state is physical if it has finite mean energy, i.e., for a single mode: $\text{Tr}(\hat{n}\rho) \leqslant \infty$. We choose a basis (or a countably infinite Hilbert-space basis) $\{|i\rangle : i = 1, \ldots, d_p\}$ of $\mathbb{C}^{d_p}$, which we call the computational basis, associate to any basis element a D-dimensional matrix $A[i] \in \mathcal{M}_D(\mathbb{C})$, and write a translationally invariant matrix product state (MPS) as

$$|\Psi\rangle = \sum_{i_L,\ldots,i_2,i_1=1}^{d_p} \langle L|A[i_L]\ldots A[i_2]A[i_1]|R\rangle|i_L,\ldots,i_2,i_1\rangle, \tag{11}$$

where $|L\rangle, |R\rangle \in \mathbb{C}^D$. This vector space is the one where the quantum computation will take place and is called correlation space. The matrices must fulfill the completeness relation

$$\sum_{i=1}^{d_p} A[i]^\dagger A[i] = \mathbb{1}, \tag{12}$$

for rendering deterministic computation feasible.

An important fact is that all physical states can be approximated arbitrary well by a finite-dimensional MPS. As any finite-dimensional state can be written as a MPS [31], it only remains to show that one can truncate states with finite energy. We write the mean total photon number of a state ρ with k modes as

$$N_{\text{mean}} = \sum_{n_1,\ldots,n_k=0}^{\infty} (n_1 + \ldots + n_k)\rho_{n_1,\ldots,n_k;n_1,\ldots,n_k}$$

$$\geqslant k(n_{\max} + 1) \sum_{n_1,\ldots,n_k=n_{\max}+1}^{\infty} \rho_{n_1,\ldots,n_k;n_1,\ldots,n_k}, \tag{13}$$

where the matrix elements of ρ are the ones of the Fock basis.

Let ρ_{trunk} be the non-normalized state obtained from ρ by setting $\rho_{n_1,\ldots,n_k;n'_1,\ldots,n'_k} = 0$ if one of the indices is larger than $n_{\max}$. This state fulfills

$$\text{Tr}(\rho - \rho_{\text{trunc}}) \leqslant \frac{N_{\text{mean}}}{k(n_{\max}+1)}. \quad (14)$$

Denoting the trace norm, which is the relevant norm for the distinguishability of quantum states, by $\|\cdot\|_1$, we get, from Ref. [32],

$$\|\rho_{\text{trunc}} - \rho\|_1 \leqslant 3\left(\frac{N_{\text{mean}}}{k(n_{\max}+1)}\right)^{1/2}. \quad (15)$$

If we assume the mean photon number (the energy) *per mode* to be constant, the truncation error is independent of the system size. This means, for a given required accuracy, we can choose a cut-off number $n_{\max}$ independently of the length of the intended computation.

B. MPS as quantum wires

We now show how single-qudit MBQC can be performed with such a matrix product state, following and extending the line of reasoning used in Ref. [4]. Let us assume that the first mode is measured in the computational basis with result k. The remaining system then is described by the state vector

$$|\Psi\rangle \propto \sum_{i_L,\ldots,i_2=1}^{d_p} \langle L|A[i_L]\ldots A[i_2]A[k]|R\rangle |i_L,\ldots,i_2\rangle. \quad (16)$$

This can be viewed as the action of the matrix $A[k]$ on the right boundary vector $|R\rangle$. To interpret this as the action of a quantum gate it is necessary that $A[k]$ is proportional to a unitary matrix, i.e., $A[k]^\dagger A[k] = p(k)\mathbb{1}$, where $p(k)$ is the probability with which the measurement result k is obtained. When the measurement basis is continuous, we denote by p the probability density while P denote the corresponding probability measure. Because we will also discuss measurements where the corresponding matrices are not proportional to unitary ones, we note that the probability in this more general situation is $p(k) = \|A[k]|\psi\rangle\|^2$, where $|\psi\rangle$ is the state vector of the correlation system. If the measurement is not performed in the computational basis but in another one and the result corresponds to a projection to the state vector $|x\rangle$, the matrix applied on $|R\rangle$ reads

$$A_{\mathcal{B}}[x] = \sum_{i=1}^{d_p} \langle x|i\rangle A[i]. \quad (17)$$

Note that the basis $\mathcal{B} = \{|x\rangle\}$ may be continuous and/or overcomplete and that measurements in such bases naturally occur in CV-quantum optics. We also allow this basis to consist of improper eigenstates reflecting an idealized homodyne detection. We can now already note an almost trivial but very important necessary requirement for a single-qudit MBQC scheme: If there exists no allowed measurement basis such that $A[x]$ is proportional to a unitary matrix for almost all x, it is not even possible to transport a single D-dimensional qudit. An equivalent formulation is that $p(x)$ must not depend on the state vector $|\psi\rangle$ of the correlation space. If this was the case, a
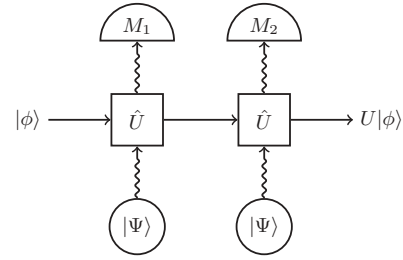

FIG. 1. Interpretation of single-qudit MBQC as sequential interaction with an ancilla. The physical sites are initially in the state vector $|\Psi\rangle$ and interact with the correlation system through $\hat{U}$. Depending on the measurement results in M_1 and M_2, the unitary gate U is applied on $|\phi\rangle$.

measurement would yield information about $|\psi\rangle$, which would clearly destroy coherence.

C. Sequential preparation

A possible way of preparing a matrix product state with bond dimension D is a sequential interaction with a D-dimensional auxiliary system [31,33]. This picture proves to be extraordinarily useful when discussing quantum wires. Let the interaction between this auxiliary system and a local physical system be described by

$$\hat{U}: \mathbb{C}^D \otimes \mathbb{C}^{d_p} \to \mathbb{C}^D \otimes \mathbb{C}^{d_p} \quad (18)$$

and assume the physical system to be initialized in the state vector $|\Psi\rangle$. The matrix elements of the MPS matrices then read

$$A[i]_{j,k} = (\langle k| \otimes \langle i|)\hat{U}(|j\rangle \otimes |\Psi\rangle). \quad (19)$$

In this picture, which is sketched in Fig. 1, the correlation system is identified with the auxiliary system, and a single-qudit gate is performed by first letting the auxiliary system interact with a mode in a known state and then measuring the mode. It is important to note that this picture works regardless of the actual way the resource state is prepared. For this reason, we will use the picture of sequential preparation to be more intuitive while maintaining full generality. For example, resource states can also arise as ground states of local Hamiltonians or be prepared by the action of translationally invariant quantum cellular automata [34,35].

D. Encoding of quantum information in correlation spaces

If the goal is not to process a D-dimensional qudit but only a d-dimensional one with $d < D$, the above requirement is too strong. The d-dimensional space is called the computational subspace. We introduce the encoding

$$V: \mathbb{C}^d \to \mathbb{C}^D \quad (20)$$

with $V^\dagger V = \mathbb{1}_d$ and the encoded matrices

$$B[i] = V^\dagger A[i] V, \quad B \in \mathcal{M}_d(\mathbb{C}). \quad (21)$$

If they are proportional to unitary matrices for all i, it is possible to process a single d dimensional qudit. One can

observe that

$$\langle\phi|V^\dagger A_{b_L}[i_L]\dots A_{b_1}[i_1]V|\psi\rangle = \langle\phi|B_{b_L}[i_L]\dots B_{b_1}[i_1]|\psi\rangle, \tag{22}$$

where $|\phi\rangle, |\psi\rangle \in \mathbb{C}^d$. The subscripts b_i denote the chosen measurement bases. This means, for all measurements, the state with the matrices B behaves exactly like the one with the matrices A. Thus, we can just proceed as if we had a state with a d-dimensional correlation space from the very beginning.

This situation is not the most general one. The isometry (20) can depend on the measurement basis and the outcome of the previous steps. In this case, the computational subspace is not fixed but changes during the computation. It is reasonable to demand that at least the initial and the final encoding coincide so it is possible to work with a fixed in- and out-coupling, provided by V. In this case Eq. (21) becomes

$$\langle\phi|V^\dagger A_{b_L}[i_L]V_{b_{L-1}}[i_{L-1}]\dots V_{b_1}[i_1]V_{b_1}^\dagger[i_1]A_{b_1}[i_1]V|\psi\rangle$$
$$= \langle\phi|B_{b_L}[i_L]\dots B_{b_1}[i_1]|\psi\rangle, \tag{23}$$

where $B[i_k] = V_{b_k}^\dagger[i_k]A[i_k]V_{b_{k-1}}[i_{k-1}]$. When several quantum wires are used to perform a quantum computation, it is also necessary to demand the changing encoding to return to the initial encoding given by V, whenever a coupling occurs. If this was not the case, the couplings would have to depend on the history of measurements and their results in both of the wires. This would clearly be against the spirit of a measurement-based protocol, where all the adaptation lies in the choice of measurement bases only.

The infinite-dimensional matrix product states used in this work bear some resemblance to the continuous matrix product states which were introduced in Refs. [36,37] to describe one-dimensional bosonic quantum fields. In contrast to the discrete structure of bosonic modes described here, they are formulated without an underlying lattice. However, such cMPS can be obtained as a suitable continuum limit from infinite-dimensional matrix product states. In this limit, only two independent MPS matrices remain while the rest can be calculated from them, which is an important difference to the present situation.

E. Coupling of wires

To go beyond single-qudit processing and achieve quantum computational universality, wires have to be coupled. Physically, this can be performed in different ways. One possibility is to let the auxiliary systems of the sequential preparation scheme interact. However, this might be very challenging to do in reality, both when working with atoms, due to the difficulty of controlling them, and with light, due to weak nonlinear interaction. For this reason, joint measurements are preferable to couple two wires which is also shown in Fig. 2. In an optical situation, this could mean to combine two modes belonging to two different wires on a beam splitter and measuring both output ports. This is a broadening of the usual definition of MBQC where only local measurements are performed while we now also allow for two local ones. On the other hand, there also exists a closely related third way which is also based on the beam-splitter interaction but strictly stays within the measurement-based paradigm. The idea is to perform the
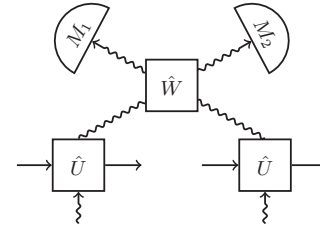

FIG. 2. Coupling of two quantum wires. The interaction of two physical systems through $\hat{W}$ and the subsequent measurements (M_1 and M_2) induces a entangling gate on the two correlation systems.

coupling independently of the executed algorithm according to some fixed scheme. This step then belongs to the preparation of the resource while the subsequently performed measurements are purely local.

Let $\hat{W}$ be the coupling between the two modes, $\{|x_1\rangle\}$ and $\{|x_2\rangle\}$ the measurement bases after interaction. The applied operation then reads

$$A[x_1, x_2] = \sum_{i,j}\langle x_1, x_2|\hat{W}|i, j\rangle A[i] \otimes A[j]. \tag{24}$$

IV. REQUIREMENTS FOR EFFICIENT MBQC SCHEMES

In this section, we give three condition a MBQC scheme has to fulfill to be called efficient and comment on their importance. To facilitate the derivation, we restrict ourselves to the arguably most important case of $D = 2$ while stressing that the requirements for universal computing on qudits with larger dimension are completely analogous. When discussing requirements, one has to distinguish between two different notions: One the one hand, there are *theoretical* requirements, of the kind that, if they are met, efficient scalable quantum computing is possible in the absence of noise. Going away from this idealized situation, on the other hand, two new features arise. First, in practice, any quantum operation is affected by noise, which error correction may counter to some extent. Second, any actual quantum computation is, needless to say, finite. Thus, the practical requirements that might be sufficient to sustain a computation may differ from the theoretical ones. However, as we are primarily interested in scalable quantum computing, for reasons of clarity we will focus on the theoretical requirements and comment on other practical implications whenever it is in order.

A. Conditions

We start with the theoretical requirements for quantum wires (which process single qubits) and state the first necessary condition.

Requirement 1 (Transport). Let $|\psi_n\rangle$ be the state vector of the correlation system after n steps in the quantum wire when the initial state vector was $|\psi\rangle$. We say that transport of a single qubit is possible if there exists some $\varepsilon > 0$ such that for all n

$$F_{\min} := \max_{U \in \mathbb{SU}_2} \min_{|\psi\rangle} |\langle\psi|U|\psi_n\rangle|^2 \geqslant \frac{1}{2} + \varepsilon. \tag{25}$$

Here, U is the best attempt at undoing the action of the quantum wire on $|\psi\rangle$. This requirement ensures that computations are not limited in length. If encodings of single logical qubits in multiple wires are used, this requirement can be weakened to $1/\varepsilon = \text{poly}(n)$.

For universal single-qubit quantum computing to be possible, it is not enough to merely transport quantum information. It is also necessary that any one-qubit unitary $U \in \mathbb{SU}_2$ can be efficiently approximated. Because it depends on the measurement outcome which operation is applied, we cannot hope for this to be possible deterministically. Instead, the measurements induce a random walk on $\mathbb{SU}_2$ which can be controlled by choosing the measurement basis in the next step depending on the gate implemented so far. For efficient quantum computing, the average number of steps must not increase too fast with the desired accuracy. We set

$$\bar{A}_\mathcal{B}[k] := p(k)^{-1/2} A_\mathcal{B}[k], \tag{26}$$

where $\mathcal{B}$ denotes the measurement basis and define

$$\bar{A}_\mathcal{B}[\mathbf{k}] = \prod_{i=1}^{n} \bar{A}_{\mathcal{B}_i}[k_i]. \tag{27}$$

Requirement 2 (Single-qubit universality). Let $U \in \mathbb{SU}_2$ and $\varepsilon > 0$. There exits a sequence (possibly adaptive) of bases $\mathcal{B}_i$ such that the expectation value for an approximation $\tilde{U} = \bar{A}_\mathcal{B}[\mathbf{k}]$ fulfilling

$$\|\tilde{U} - U\| \leqslant \varepsilon \tag{28}$$

after at most n measurements in the wire satisfies

$$\mathbb{E}(n) = O\left[\text{polylog}\left(\frac{1}{\varepsilon}\right)\right]. \tag{29}$$

Here, $\|\cdot\|$ denotes the operator norm, which is the meaningful figure of merit in this case. Of course, single-qubit operations are not sufficient for universality but an entangling operation is also needed:

Requirement 3 (Coupling). Denote with

$$C_Z = \text{diag}(1,1,1,-1) \tag{30}$$

the controlled-Z gate. The expectation value for an approximation $\tilde{C}_Z$ fulfilling

$$\|\tilde{C}_Z - C_Z\| \leqslant \varepsilon \tag{31}$$

after at most n measurements on both of the quantum wires involved, where both single-site and two-site measurements are allowed, must satisfy

$$\mathbb{E}(n) = O\left[\text{polylog}\left(\frac{1}{\varepsilon}\right)\right]. \tag{32}$$

Requirements 2 and 3 guarantee that every quantum circuit can be simulated in correlation space with a polylogarithmic overhead. This is the commonly used requirement in quantum computing. When one is content with a polynomial overhead, (29) and (32) can both be relaxed to

$$\mathbb{E}(n) = O\left[\text{poly}\left(\frac{1}{\varepsilon}\right)\right] \tag{33}$$

for some constant $c > 0$. A computational model fulfilling only this weaker condition is still of interest when the ultimate

aim is to perform a quantum algorithm which provides exponential speed-up over the corresponding classical one (like Shor's algorithm). However, whenever the speed-up is only polynomial (like in Grover's search algorithm), the stronger polylog scaling is clearly necessary for the quantum algorithm to have any advantage. When one aims at the implementation of constant size quantum circuits only, Requirements 2 and 3 are not necessary as one does not need to implement the quantum gates to arbitrary accuracy but only to some $\varepsilon > 0$ which is determined by the total length of the computation.

The last feature to demand from a MBQC model is the ability to initialize the correlation system into some known state and to perform a measurement of the correlation space in the computational basis.

Requirement 4 (Initialization and read-out). For efficient initialization of the correlation system to be possible, there must exist some measurement sequence such that, independently from its initial state vector $|\psi\rangle$,

$$\|\,|\psi_n\rangle - |0\rangle\| \leqslant \varepsilon. \tag{34}$$

Here $|\psi_n\rangle$ is the state vector after n steps in the wire and it must be true that $\mathbb{E}(n) = O(\text{polylog}(1/\varepsilon))$.

For read-out, it is necessary that after n steps with

$$\mathbb{E}(n) = O\left[\text{polylog}\left(\frac{1}{\varepsilon}\right)\right] \tag{35}$$

the combined normalized action

$$\bar{A}_\mathcal{B}[\mathbf{k}] = \prod_{j=n}^{1} \bar{A}_{\mathcal{B}_j}[\mathbf{k}_j] \tag{36}$$

fulfills either $\|\bar{A}_\mathcal{B}[\mathbf{k}] - |0\rangle\langle 0|\| \leqslant \varepsilon$ or $\|\bar{A}_\mathcal{B}[\mathbf{k}] - |1\rangle\langle 1|\| \leqslant \varepsilon$.

The first requirement just means that one can approach one of the computational basis states fast enough, while the second one ensures an efficient implementation of an approximative measurement in the computational basis. While measurement-based computational schemes that do not respect these requirements may be conceivable in principle, within the framework presented here they are very natural indeed and necessary for universal quantum computing.

B. Consequences

In the previous section we have discussed the requirements a MBQC scheme has to fulfill in order to be called efficient. We now show some consequences of these requirements, which will help us find classes of schemes that cannot be efficient and some that are. This seems an important enterprise in order to arrive at "no-go results" to identify serious limitations that have to be circumvented in one way or the other. For many important situations, $F_{\min}$ in Requirement 1 vanishes exponentially in n as the following observation shows, which by virtue of the above statement implies that efficient quantum computing is not possible.

Observation 5 (Impossibility of transport). Given a basis $\mathcal{B}$, let $\mathcal{C} \subset \mathcal{B}$, with $P(\mathcal{C}) > 0$. If for all $x \in \mathcal{C}$, $A[x] \not\propto U$ with $U \in \mathbb{SU}_2$, transport is impossible when measuring in this basis.

Proof. We calculate the fidelity

$$f_{\min}(x) := \max_{U[x] \in \mathbb{SU}_2} \min_{|\psi\rangle} \frac{|\langle\psi|U[x]A[x]|\psi\rangle|^2}{p(x)}. \quad (37)$$

Because $P(\mathcal{C}) > 0$ and $f_{\min}(x) < 1$ for all $x \in \mathcal{C}$, there exists a nonempty compact interval $\mathcal{I} \subset \mathcal{C}$ such that there is some $\delta < 1$ with $f_{\min}(x) \leqslant \delta$ for all $x \in \mathcal{I}$. Using an argument from Ref. [17], one can show that this implies the exponential decay of $F_{\min}$, showing that Requirement 1 is not satisfied. Thus, for transport to be possible, there must exist a basis such that almost all measurement outcomes correspond to some $A[x]$ with $A[x]^\dagger A[x] \propto \mathbb{1}$. For transport over a fixed length, Observation 5 does not apply because even if the fidelity decays exponentially fast, it might be still large enough for the task under question.

We show a simple sufficient condition, which will be used later when discussing examples for which Requirement 2 is true.

Observation 6 (Condition for an efficient random walk). Assume that for any target gate $U \in \mathbb{SU}_2$ there exists some basis $\mathcal{B}$ and some $\mathcal{C} \subset \mathcal{B}$ for any $P(\mathcal{C}) \geqslant p(\varepsilon)$ with $1/p(\varepsilon) = O\left(\text{polylog}(1/\varepsilon)\right)$ where ε is the desired accuracy. When

$$\|\bar{A}_\mathcal{B}[k] - U\| \leqslant \varepsilon \quad (38)$$

for all $k \in \mathcal{C}$, Requirement 2 is fulfilled.

Proof. The expected number of tries can be bounded by

$$\mathbb{E}(n_\varepsilon) \leqslant \sum_{k=1}^\infty k p(\varepsilon)[1 - p(\varepsilon)]^{k-1} = \frac{1}{p(\varepsilon)}$$
$$= O\left[\text{polylog}\left(\frac{1}{\varepsilon}\right)\right]. \quad (39)$$

The most important important situation covered by Observation 6 is the one where the measurements have discrete results and for every $U \in \mathbb{SU}_2$ a basis exists which contains at least a single element which fulfills the requirements of Observation 6. Sufficient conditions for initialization and read-out of the correlation system to be possible are provided by the following observation:

Observation 7 (Sufficient conditions for Requirement 4). Let there exist some $\delta > 1$ and a basis $\mathcal{B}$ containing some subset $\mathcal{C} \subset \mathcal{B}$ of "nonunitary measurements." Assume that

$$A[x] \propto U[x] \begin{bmatrix} s_1(x) & 0 \\ 0 & s_2(x) \end{bmatrix} V^\dagger \quad (40)$$

for all $x \in \mathcal{C}$, where $U[x], V \in \mathbb{SU}_2$, and $s_1(x), s_2(x)$ are the (not necessarily ordered) singular values of $A[x]$ which fulfill either $s_1/s_2 > \delta$ or $s_2/s_1 > \delta$. In this case, efficient read-out and initialization of the correlation system is possible.

Proof. The situation is most transparent when $s_2(x) = 0$. In this case, a measurement result x corresponds to a projection of the correlation space state to the state vector $V|0\rangle$. In the more general case the measurement is not projective but merely weak. By performing the unitary operation $VU^\dagger[x]$, which is efficiently possible due to Requirement 2, and repeating the measurement, a projective measurement in the basis $\{V|0\rangle, V|1\rangle\}$ can be approximated. This approximation is efficient, due to the existence of a finite gap between s_1 and s_2 occurring with finite probability. The independence of V of x is

crucial: If this is not the case, the basis in which the correlation space measurement will occur cannot be fixed. This will lead to a destruction of quantum information resulting in a failure of the MBQC scheme. Initialization of the correlation system in the state vector $|0\rangle$ can be performed in the very same manner. One just performs the measurement procedure outlined above after which the correlation system is in $U[x]|0\rangle$ or $U[x]|1\rangle$, depending on the outcome. Applying now the gates $U[x]^\dagger$ or $XU[x]^\dagger$, respectively, achieves the required initialization.

V. LIMITATIONS

When devising a MBQC scheme, there are several fundamental limitations concerning the use of Gaussian states and Gaussian measurements. The first important fact is that all protocols involving only Gaussian measurements on Gaussian states can be simulated efficiently on a classical computer, ruling out the possibility for any quantum speed-up [11]. However, for an *ideal* Gaussian cluster state, universality can be achieved by adding a single non-Gaussian measurement to the toolbox [12]. On the other hand, any *physical* Gaussian quantum wire, including an one-dimensional Gaussian cluster state cannot fulfill Requirement 1 even when allowing for non-Gaussian measurements [17].

A. Controlled Gaussian operations

We now discuss further limitations and consider the situation where the interaction (18) can be written as

$$\hat{U} = |0\rangle\langle0| \otimes \hat{G}_0 + |1\rangle\langle1| \otimes \hat{G}_1, \quad (41)$$

up to local unitaries on the correlation system, where $\hat{G}_0, \hat{G}_1 \in \mathbb{U}_G$. This class contains states which allow for transport of a qubit in a wire by Gaussian measurements only. An example is given by

$$\hat{U} = |0\rangle\langle0| \otimes \exp(-i\theta\hat{n}) + |1\rangle\langle1| \otimes \exp(i\theta\hat{n}), \quad (42)$$

where $\theta > 0$ is a parameter characterizing the interaction strength. Taking a coherent state vector $|\alpha\rangle$ with $\alpha \in \mathbb{R}, \alpha > 0$ as input, $q_0 = \sqrt{2}\alpha\cos(\theta)$, $p_0 = \sqrt{2}\alpha\sin(\theta)$, a q-quadrature measurement with result x applies in correlation space

$$\bar{A}[x] = \text{diag}\left(e^{ip_0(x-\frac{q_0}{2})}, e^{-ip_0(x-\frac{q_0}{2})}\right) \quad (43)$$

with probability $p(x) = \exp(-(x-q_0)^2)/\sqrt{\pi}$. Thus, Requirement 1 is fulfilled. However, efficient single-qubit gates are impossible due to the lack of control. The only possible measurement (up to squeezing) for which $\bar{A}[x]$ is unitary are phase-space displacements followed by q-quadrature homodyne detection as for all others, the projectors do not fulfill

$$|\langle\psi(x)|e^{i\theta}\alpha\rangle| = |\langle\psi(x)|e^{-i\theta}\alpha\rangle|, \quad (44)$$

which implies that the probabilities of the measurements depend on the state of the correlation system and, therefore, the applied matrix is not unitary. Writing the displacement as $\Delta = (\Delta_q + i\Delta_p)/\sqrt{2}$, we get

$$\bar{A}_\Delta[x] = \text{diag}\left(e^{i(p_0+\Delta_p)(x-\frac{q_0+\Delta_q}{2})}, e^{-i(p_0-\Delta_p)(x-\frac{q_0+\Delta_q}{2})}\right) \quad (45)$$

with probability $p_\Delta(x) = \exp(-(x-q_0-\Delta_q)^2)/\sqrt{\pi}$. Up to a redefinition $q_0 \mapsto q_0 + \Delta_q$ and some global phase, (45) is

constant in Δ, i.e., there is no way of controlling the random walk, making it impossible to meet Requirement 2. The same argument holds for any controlled Gaussian operation and for any Gaussian input state.

When allowing for non-Gaussian input states, some amount of control is indeed possible, as the following example shows: Consider the interaction of Eq. (42) together with a superposition of two different photon numbers as an input, $|\psi\rangle = (|0\rangle + |2\rangle)/\sqrt{2}$. Choosing $\theta = \pi/4$, the postinteraction state reads

$$a|0\rangle \frac{1}{\sqrt{2}}(|0\rangle - i|2\rangle) + b|1\rangle \frac{1}{\sqrt{2}}(|0\rangle + i|2\rangle) \qquad (46)$$

and the normalized matrices $\bar{A}_q$ and $\bar{A}_p$ for the q- and p-quadrature measurements are

$$\bar{A}_q[x] = \mathrm{diag}\,(e^{-i\phi(x)}, e^{i\phi(x)}), \qquad (47)$$

$$\bar{A}_p[x] = \mathrm{diag}\,(e^{i\phi(x)}, e^{-i\phi(x)}), \qquad (48)$$

with

$$\phi(x) = \arctan \frac{2x^2 - 1}{\sqrt{2}} \qquad (49)$$

and

$$p(x) = (4x^4 - 4x^2 + 3)\exp(-x^2)/(4\sqrt{\pi}). \qquad (50)$$

Thus, there exist two Gaussian measurements which allow to control the random walk but even this control is not enough to meet Requirement 2. Because this is a general feature of any MBQC-scheme relying solely on Gaussian measurements we will discuss it in detail.

B. General resource states with Gaussian measurements

We consider an MPS with $D = 2$ with the continuous family of matrices $A_q[x]$ with $x \in \mathbb{R}$. As the projectors describing eight-port homodyning are not orthogonal, they will turn the state of the remaining system into a mixed one when measuring a mode, and we do not need to consider this measurement. Because all relevant Gaussian measurements can be expressed as the application of squeezing, rotating, and displacing, followed by a q-quadrature measurement, we can write

$$\begin{aligned} A_{\theta, \Delta_q, \Delta_p, \lambda}[x] \\ = \cos(\theta) A_q[\lambda x + \Delta_q] + \sin(\theta) A_p[x/\lambda + \Delta_p], \quad (51) \end{aligned}$$

where

$$A_p[q] = \frac{1}{\sqrt{2\pi}} \int dx\, A_q[x] e^{iqx} \qquad (52)$$

and where we have omitted global phase factors. If single-qubit transport is possible in this wire, we can, without loss of generality, assume that $A_q[x] := A_{0,0,0,0}[x]$ is proportional to a unitary matrix for all relevant x. As both squeezing by a finite λ and shifting in phase space merely results in a redefinition

of x in Eq. (51), we can restrict ourselves to the discussion of

$$\begin{aligned} A_\theta[x] &= \cos(\theta) A_q[x] + \sin(\theta) A_p[x] \\ &= \sum_{n=0}^{d_p - 1} e^{i\theta n} \psi_n(x) A[n] \end{aligned} \qquad (53)$$

where $\psi_n(x)$ are the energy eigenfunctions of the harmonic oscillator. We assume that the physical dimension is finite and discuss the required changes due to an infinite physical dimension afterward.

We now argue that Requirement 2 cannot be fulfilled in the present situation:

Observation 8 (Impossibility of control). Let $A_\theta[x]$ be as in (53) with finite d_p. Then, there exist some constants $C > 0$, $\lambda > 0$, and $\varepsilon_0 > 0$ such that for all $U \in \mathbb{SU}_2$, $\theta \in [0, \pi]$, and all $0 < \varepsilon \leqslant \varepsilon_0$

$$\mathbb{P}_\theta(\|\bar{A}_\theta[x] - U\| \leqslant \varepsilon) < C\varepsilon^\lambda, \qquad (54)$$

where $\mathbb{P}_\theta$ denotes the probability for the random variable x which corresponds to a quadrature measurement with angle θ.

If this is true, the expected number of tries to implement any unitary is bounded as $\mathbb{E}(n_\varepsilon) > C^{-1}(1/\varepsilon)^\lambda$ for $0 < \varepsilon \leqslant \varepsilon_0$, and Requirement 2 is not fulfilled. The proof is given in the appendix.

This result, which can be summarized as "no quantum computation with Gaussian measurements only" is complementary to the one reported in Ref. [17] which can, in turn, be summarized as "no quantum computing with Gaussian states only." These results apply only to the model where all operations are noiseless and error correction is not performed. Together, this means that both non-Gaussian states and non-Gaussian measurements are simultaneously necessary for continuous-variable MBQC. However, the two limitations differ markedly. For the former, not enough localizable entanglement exists, which means that the quantum information gets destroyed along the wire and transport or teleportation is not possible. The root of the latter limitation is the insufficient possibility to control what quantum gate is performed.

C. Infinite-dimensional physical systems

Up to now, we have considered only resource states where the physical dimension is finite. We demonstrate the difference between states with and without support on a finite number of Fock states with the help of an example where the resource state is described by the matrices

$$A[x] = \frac{1}{\sqrt{\mathcal{N}}}(f(x)\mathbb{1} + ig(x)Z) \qquad (55)$$

with $\mathcal{N} = \int dx\,(|f(x)|^2 + |g(x)|^2)$ to ensure the normalization condition

$$\int dx\, A[x]^\dagger A[x] = \mathbb{1}. \qquad (56)$$

Both f and g are chosen to be real, even functions. Thus, also their Fourier transforms $\tilde{f}$, $\tilde{g}$ are real and the matrices corresponding to $\hat{q}_\theta$ are

$$\begin{aligned} A_\theta[x] = \frac{1}{\sqrt{\mathcal{N}}}\{&[(f(x)\cos(\theta) + \tilde{f}(x)\sin(\theta)]\mathbb{1} \\ &+ i[g(x)\cos(\theta) + \tilde{g}(x)\sin(\theta)]Z\}, \quad (57) \end{aligned}$$

which are all unitary. If the resource state is finite dimensional, f, g, $\tilde{f}$, and $\tilde{g}$ are of the form

$$h(x) = \text{poly}(x) \exp(-x^2/2) \tag{58}$$

and the probability for $\bar{A}_\theta[x]$ to differ too much from the desired unitary is too large. If we allow the resource state to be infinite dimensional, the situation changes. In particular, it is possible that $f(x) = \text{const}$ for $|x| \leqslant c$ for some constant $c > 0$. This allows us to find a resource state where a q-quadrature measurement implements some gate with a finite probability. This is a striking example for a qualitative difference between a proper continuous-variable state and all of its finite dimensional truncations.

One might now be tempted to use this new insight to devise a MBQC scheme which allows for theoretically efficient single-qubit processing with Gaussian measurements only, but this is impossible. This can be immediately seen by truncating the state according to (15). The above results then imply that single-qubit MBQC is impossible when the energy per mode stays constant when scaling the system. If, on the other hand, the energy per mode is increased with the length of the computation, this no-go result no longer holds.

The reasons for the existence of such a severe limitation lies in the strong notion of efficiency. If Eq. (32) in Requirement 2 is weakened to (33), this restriction is no longer present, and it is possible to devise a MBQC-scheme based on Gaussian measurements on a non-Gaussian resource. To do so, one can take the resource defined by the matrices (43). After changing the variable and neglecting global phases, one gets

$$\bar{A}[x] = S(-p_0 q_0) S(-2 p_0 x) \tag{59}$$

with $p(x) = \exp(-x^2)/\sqrt{\pi}$, where

$$S(\phi) = \text{diag}[\exp(-i\phi/2), \exp(i\phi/2)]$$
$$= \exp(-i\phi/2)\text{diag}[1, \exp(i\phi)] \tag{60}$$

is the phase gate. Assume that the target gate is $S(\phi)$. Chose q_0 and p_0 such that $\theta = -q_0 p_0$ and, therefore, $\bar{A}[0] = S(\theta)$. Using $\|S(\alpha) - S(\beta)\| \leqslant |\alpha - \beta|$, one can bound the probability that the distance of the actually implemented gate to the desired one is larger than ε as

$$\mathbb{P}(\|\bar{A}[x] - S(\theta)\| > \varepsilon) \leqslant \mathbb{P}(|2 p_0 x| > \varepsilon) \leqslant C p_0 \varepsilon, \tag{61}$$

where C is some constant. This means it is possible to approximate an arbitrary phase gate in $O(1/\varepsilon)$ steps. Changing the interaction unitary to contain an additional Hadamard gate on the correlation space and using the decomposition detailed below together with the composition law for errors, it is easy to show that any $U \in \mathbb{SU}_2$ can be approximated in $O((1/\varepsilon)^4)$ steps.

VI. FEASIBLE, EFFICIENT CV-MBQC SCHEMES

In this section, we present a scheme which allows for efficient CV-MBQC. The limitations discussed above force us to use both non-Gaussian resource states and non-Gaussian measurements. Even though the experimental requirements for its realization are indeed challenging, the scheme uses only primitives which all have already been demonstrated to be feasible in proof-of-principle experiments.

A. Single-qubit operations

We choose the interaction unitary in Eq. (18) to be

$$U = (H \otimes \mathbb{1})(|0\rangle\langle 0| \otimes \exp(-i\theta\hat{n}) + |1\rangle\langle 1| \otimes \exp(i\theta\hat{n})), \tag{62}$$

where $\theta > 0$ is a parameter. We initialize the modes in a coherent state vector $|\alpha\rangle$ with $\alpha > 0$. Equation (62) describes a controlled rotation in phase space. The class of measurements corresponding to unitary evolution is given by the displaced photon-counting measurements, i.e., projections onto the state vector $|x,n\rangle := \hat{D}(x)|n\rangle$, where $x \in \mathbb{R}$. With the help of (10), the applied gates read

$$\bar{A}_x[n] = H \begin{bmatrix} e^{-in\phi(x)} & 0 \\ 0 & e^{in\phi(x)} \end{bmatrix}, \tag{63}$$

where

$$\phi(x) = \arctan\left(\frac{\alpha \sin(\theta)}{\alpha \cos(\theta) + x}\right). \tag{64}$$

The corresponding probabilities are

$$p_x(n) = e^{-(\alpha^2 + x^2 + 2x \cos(\theta))} \frac{(\alpha^2 + x^2 + 2x \cos(\theta))^n}{n!} \tag{65}$$

and fulfill, for every x, the normalization condition $\sum_n p_x(n) = 1$.

Any single-qubit unitary operation can be efficiently approximated with those operations: Every $U \in \mathbb{SU}_2$ can be written as

$$U = S(\phi_1) H S(\phi_2) H S(\phi_3), \tag{66}$$

where H is the Hadamard gate and $\phi_i \in [0, 2\pi]$. Rewriting this as

$$U = H S(0) H S(\phi_1) H S(\phi_2) H S(\phi_3), \tag{67}$$

we have decomposed every gate into four applications of (63). Inverting (64) yields

$$x(\phi) = \frac{\alpha \sin(\theta)}{\tan(\phi(x))} - \alpha \cos(\theta). \tag{68}$$

To implement $S(\phi)$, one only has to choose x such that $-2\phi(x)n = \phi$ for some small n where $n = 0$ and $n = 1$ suffice and the former is only needed for $\phi = 0$. For any fixed $\alpha \neq 0$ and $\theta \neq 0$, the maximally necessary shift $|x|$ is bounded and with (65) there exists a lower bound $0 < p_0 \leqslant p_x(n)$ for all relevant x and n. Using (67), one can see that the probability to obtain any chosen $U \in \mathbb{SU}_2$ in four steps is lower bounded by p_0^4. Thus, the conditions for Observation 6 are fulfilled and efficient single-qubit processing is possible.

B. Coupling of wires

We now turn to the coupling of two wires. Let the two correlation space qubits be prepared in $|\psi\rangle_j = a_j|0\rangle + b_j|1\rangle$ for $j = 1,2$. Then, the state vectors after the interaction with the light mode reads

$$a_j|0\rangle|e^{-i\theta}\alpha\rangle + b_j|1\rangle|e^{i\theta}\alpha\rangle. \tag{69}$$

Displacing now both modes by $\Delta_x = \alpha \cos(\theta)$, rotating the second mode by an angle of $\pi/2$, and setting $\gamma = \alpha \sin(\theta)$, the

joint state vector becomes

$$a_1a_2|0,0\rangle|\gamma\rangle|i\gamma\rangle + a_1b_2|0,1\rangle|\gamma\rangle|-i\gamma\rangle$$
$$+ b_1a_2|1,0\rangle|-\gamma\rangle|i\gamma\rangle + b_1b_2|1,1\rangle|-\gamma\rangle|-i\gamma\rangle. \quad (70)$$

A balanced beam splitter transforms two coherent states as

$$|\alpha\rangle|\beta\rangle \mapsto \left|\frac{\alpha+\beta}{\sqrt{2}}\right\rangle\left|\frac{\alpha-\beta}{\sqrt{2}}\right\rangle. \quad (71)$$

Applying this to (70) yields

$$a_1a_2|0,0\rangle\left|\frac{\gamma}{\sqrt{2}}(1+i)\right\rangle\left|\frac{\gamma}{\sqrt{2}}(1-i)\right\rangle$$
$$+ a_1b_2|0,1\rangle\left|\frac{\gamma}{\sqrt{2}}(1-i)\right\rangle\left|\frac{\gamma}{\sqrt{2}}(1+i)\right\rangle$$
$$+ b_1a_2|1,0\rangle\left|\frac{\gamma}{\sqrt{2}}(-1+i)\right\rangle\left|\frac{\gamma}{\sqrt{2}}(-1-i)\right\rangle$$
$$+ b_1b_2|1,1\rangle\left|\frac{\gamma}{\sqrt{2}}(-1-i)\right\rangle\left|\frac{\gamma}{\sqrt{2}}(-1+i)\right\rangle. \quad (72)$$

The matrices corresponding to a photon-counting measurement in both modes with results n_1 and n_2 respectively read

$$A[n_1,n_1] = \text{diag}\left(e^{-\frac{i\pi n_1}{4}}, e^{-\frac{i7\pi n_1}{4}}, e^{-\frac{i3\pi n_1}{4}}, e^{-\frac{i5\pi n_1}{4}}\right)$$
$$\times \text{diag}\left(e^{-\frac{i7\pi n_2}{4}}, e^{-\frac{i\pi n_2}{4}}, e^{-\frac{i5\pi n_2}{4}}, e^{-\frac{i3\pi n_2}{4}}\right), \quad (73)$$

which is, up to a local unitary, a controlled phase gate, i.e.,

$$A[n_1,n_2] = \text{diag}\left(e^{-\frac{i\pi(n_1+7n_2)}{4}}, e^{-\frac{i\pi(3n_1+5n_2)}{4}}\right)$$
$$\otimes \text{diag}\left(1, e^{-\frac{i\pi(6n_1-6n_2)}{4}}\right)$$
$$\times \text{diag}\left(1,1,1,e^{-i\pi(-n_1+n_2)}\right). \quad (74)$$

If $n_1 + n_2$ is even, $A[n_1,n_2]$ is not entangling, and the local operations can be efficiently undone as described above. If $n_1 + n_2$ is odd, a C_Z gate is implemented up to local corrections. As this situation occurs with finite probability for any choice of α and θ, this means that Requirement 3 is fulfilled.

It remains to show that initialization and read-out are possible. To measure in the computational basis, we shift the mode after the interaction by $\Delta = \alpha(\cos(\theta) - i\sin(\theta))$ which turns the joint state of qubit and mode to

$$a|0\rangle|0\rangle + b|1\rangle|2i\alpha\sin(\theta)\rangle. \quad (75)$$

Counting the photons leads to a matrix fulfilling the conditions of Requirement 4. Note that there is some asymmetry: If the counted number of photons is larger than one, the qubit is in $|1\rangle$. If the result is zero, one has to undo the gate $H \text{diag}(1,-i)$ and repeat the procedure. Due to the gap between 1 and $|\langle 0|2i\alpha\sin(\theta)\rangle|$, this is efficiently possible. This concludes the proof that the proposed scheme is efficient.

As already mentioned, the coupling scheme discussed so far is not completely in the MBQC paradigm because it involves the adaptive coupling of sites. However, this is not necessary: Consider two wires which are coupled every $k + 1$ sites by the beam splitter described by Eq. (71). Because $C_Z^2 = \mathbb{1}$, two consecutive couplings can be undone by appropriately choosing local operations between them. From Eq. (65) it follows that for every block of four sites, there exists
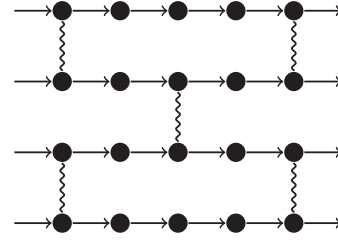

FIG. 3. Nonadaptive coupling of four quantum wires where the wiggly lines denote the probabilistic measurement-based C_Z gate. By scaling the distance between the single couplings polylogarithmically in the length of the computation, the total probability of failure can be made arbitrarily small.

a probability $p_0 > 0$ that the desired gate is realized and probability of failure in at least one of the wires is upper bounded by $2(1 - p_0)^{k/4}$. Let n be total number of gates, the total probability of failure is bounded by

$$p_{\text{failure}} \leqslant 2n(1-p_0)^{k/4} \leqslant c_1 n e^{-c_2 k} \quad (76)$$

for suitable constants c_1 and c_2. Thus, for a fixed probability of success, k has only to grow as $k = O(\text{polylog}(n))$, which means that the need for two-site measurements can be removed for the price of a polylog overhead. For a quantum circuit consisting of more than two wires, as depicted in Fig. 3, every wire should be alternately coupled to its left and right neighbor. When the circuits consists of m quantum wire with all have a length n, the respective overhead behaves as $k = O(\text{polylog}(nm))$.

C. Errors

Even if a model fulfills the theoretical requirements it will still be affected by noise. Thus, error correction will be necessary to build a scalable quantum computer. In the scheme presented above, one plausible dominant source of errors is reflecting the finite efficiency of the photon-counting measurement needed to perform both single- and two-qubit gates. If photon loss happens, a "wrong" operation is applied in the two-dimensional correlation space. When the rate of these qubit errors is low enough, techniques of error correction and fault-tolerant quantum computing may be applied to make a computation with unbounded length possible at the expense of a polylogarithmic overhead [18,19]. This contrasts with the errors stemming from the use of finitely squeezed states in the scheme based on Gaussian cluster states for which no method of error correction exists yet to achieve fault tolerance.

In addition to the obvious errors stemming from decoherence and imperfections in the measurement procedure, the presented continuous variable schemes have an important additional source of errors. For example, when performing a shifted photon counting, uncertainty in the parameter of the shift operator $D(x)$ results in uncertainty of the applied operator. We now investigate how those errors scale with the length of the computation. We restrict ourselves to the single-qubit case while noting that the error analysis for multiple qubits is completely analogous. Consider the situation

where the target gate reads $U = \prod_{i=1}^{n} U_i$ while the gate actually applied is $\bar{A} = \prod_{i=1}^{n} \bar{A}_i$. We assume that the error induced by a single step is bounded, i.e., for all i,

$$\|\bar{A}_i - U_i\| \leqslant \varepsilon. \tag{77}$$

If all $\bar{A}_i$ are also unitary, the total error grows linearly in n and can be bounded as $\|\bar{A} - U\| \leqslant n\varepsilon$. If this is not the case, an tight upper bound to the error is

$$\|\bar{A} - U\| \leqslant n\varepsilon(1 + \varepsilon)^n, \tag{78}$$

i.e., it grows, in the worst case, exponentially in n. In the MBQC-scheme based on Gaussian cluster states as discussed, e.g., in Refs. [12–15], the implemented gates are not unitary for physical resource states, which is the reason for the exponential decay of the transport fidelity [17].

VII. CONCLUSION

In this work, we have investigated the possibilities provided by measurement-based quantum computing with continuous-variable resources. After introducing a framework allowing for the description of a huge class of possible resource states, we have clarified which conditions are necessary for a CV-MBQC scheme to be viable. Those conditions lead to new limitations. Especially, we have shown that, without the use of not-yet-developed continuous-variable error correction, scalable quantum computing with Gaussian measurements alone is impossible, even if the resource state is non-Gaussian, complementing the prior result which prohibits MBQC with Gaussian resources, even if the measurements are non-Gaussian. As the second main result, we gave an explicit example of an efficient MBQC scheme where the resource could be created by a simple interaction between light modes in coherent states and some qubit degrees of freedom. Processing is then performed by shifted photon counting while entangling two-qubit gates rely on simple beam splitters. An analysis of errors highlighted the qualitative differences of the two major kinds of errors.

Two major conclusions can be drawn from our findings: First, continuous-variable measurement based quantum computing is an extraordinarily difficult enterprise. Second, when non-Gaussian resource states are combined with non-Gaussian measurements, efficient schemes do exist, and in this work, we introduce such strictly efficient schemes. While this may be disappointing when aiming at a scalable optical quantum computer, one does need to keep in mind that our statements about efficiency address only the asymptotic behavior. Thus, the limitation do not at all rule out the usefulness of much simpler protocols in situations where only a few modes are used. Such situations include quantum repeaters and hybrid entanglement-distribution protocols for applications particularly in quantum cryptography. Hence, for such purposes, both the purely Gaussian setting as well as those non-Gaussian schemes that asymptotically do not give rise to universal quantum computing can well be feasible as elementary tools for entanglement distillation schemes. It is the hope that the framework established here fosters further such investigations.

APPENDIX: PROOF OF OBSERVATION 8

To find an upper bound to $\mathbb{P}(\|\bar{A}_\theta[x] - U\| \leqslant \varepsilon)$, we first lower bound the probability that $|x| > x_0$ for some suitably chosen x_0. Then, for $|x| \leqslant x_0$, we upper bound $\|\bar{A}_\theta[x] - U\|$ by a polynomial in x and use that the probability for a polynomial to be larger than ε cannot grow too fast for small ε.

We calculate

$$\|\bar{A}_\theta[x] - U\| \geqslant \|\frac{A_\theta[x]}{\sqrt{p_\theta(x)}} - U\|_2/\sqrt{2}, \tag{A1}$$

where $\|O\|_2 = \sqrt{\sum_{i,j} |O_{ij}|^2}$ denotes the Frobenius norm. We now use (53) and the fact that

$$\psi_n(x) = \frac{1}{\sqrt{2^n n! \sqrt{\pi}}} H_n(x) e^{-x^2/2}, \tag{A2}$$

where $H_n(x)$ denotes the nth Hermite polynomial. We also use $0 \leqslant \sqrt{x} \leqslant x + 1$ and that $p_\theta(x)$ is upper bounded by a constant independent of θ. This allows us to find non-negative polynomials $f_{U,\theta}$ and $g_{U,\theta}$, where the coefficients are continuous functions of U and θ with $\|\bar{A}_\theta[x] - U\| \geqslant f_{U,\theta}(x)/g_{U,\theta}(x)$. As the probability density $p_\theta(x)$ is of the form $p_\theta(x) = \text{poly}(x) \exp(-x^2)$ we get

$$\mathbb{P}_\theta\left[|x| > C_1 \ln\left(\frac{C_2}{\varepsilon}\right)\right] \leqslant \varepsilon \tag{A3}$$

for some suitable $C_1 > 0$, $C_2 > 0$. Now we assume that $|x| \leqslant C_1 \log(C_2/\varepsilon)$. In this case, $g_{U,\theta}$ can be upper bounded by $C_3 (C_1 \log(C_2/\varepsilon))^r$ for some $C_3 > 0$ and some even r. Note that the maximal degree of all polynomials only depends on d_p. Assume that $f_{U,\theta}$ has M zeros $x_1, \ldots, x_M$. Around any of these zeros, $f_{U,\theta}$ can be lower bounded by $C_4(x - x_i)^s$ for some $C_4 > 0$ and some even s. Using this and (A3), we chose some ε_0 such that

$$\mathbb{P}_\theta(\|\bar{A}_\theta[x] - U\| \leqslant \varepsilon) \leqslant \varepsilon \tag{A4}$$

$$+ \sum_{i=1}^{M} \mathbb{P}_\theta\left\{C_4(x - x_i)^s \leqslant C_3\left[C_1 \ln\left(\frac{C_2}{\varepsilon}\right)\right]^r \varepsilon\right\}. \tag{A5}$$

for all $\varepsilon \leqslant \varepsilon_0$. Because $p_\theta(x)$ is upper bounded, there exist a constant C_5 such that

$$\mathbb{P}((x - x_i)^s \leqslant \varepsilon) \leqslant C_5 \varepsilon^{1/s}. \tag{A6}$$

for all $\varepsilon \leqslant \varepsilon_0$. Inserting (A6) into (A5), one obtains (54). As all constants depend on U and θ in a continuous way, and U and θ are both taken from compact domains, there exist a set of constants such that (54) is simultaneously true for all U and θ. Thus, Observation 8 holds.

[1] R. Raussendorf and H. J. Briegel, Phys. Rev. Lett. **86**, 5188 (2001).

[2] R. Raussendorf and H. J. Briegel, Quant. Inf. Comp. **2**, 443 (2002).

[3] D. Gross and J. Eisert, Phys. Rev. Lett. **98**, 220503 (2007).

[4] D. Gross, J. Eisert, N. Schuch, and D. Perez-Garcia, Phys. Rev. A **76**, 052315 (2007).

[5] J. Eisert and M. B. Plenio, Int. J. Quantum. Inform. **1**, 479 (2003).

[6] S. L. Braunstein and P. van Loock, Rev. Mod. Phys. **77**, 513 (2005).

[7] J. Eisert, S. Scheel, and M. B. Plenio, Phys. Rev. Lett. **89**, 137903 (2002).

[8] J. Fiurášek, Phys. Rev. Lett. **89**, 137904 (2002).

[9] G. Giedke and J. I. Cirac, Phys. Rev. A **66**, 032316 (2002).

[10] J. Niset, J. Fiurášek, and N. J. Cerf, Phys. Rev. Lett. **102**, 120501 (2009).

[11] S. D. Bartlett, B. C. Sanders, S. L. Braunstein, and K. Nemoto, Phys. Rev. Lett. **88**, 097904 (2002).

[12] N. C. Menicucci, P. van Loock, M. Gu, C. Weedbrook, T. C. Ralph, and M. A. Nielsen, Phys. Rev. Lett. **97**, 110501 (2006).

[13] N. C. Menicucci, S. T. Flammia, and O. Pfister, Phys. Rev. Lett. **101**, 130501 (2008).

[14] S. T. Flammia, N. C. Menicucci, and O. Pfister, J. Phys. B **42**, 114009 (2009).

[15] R. Ukai, J.-I. Yoshikawa, N. Iwata, P. van Loock, and A. Furusawa, Phys. Rev. A **81**, 032315 (2010).

[16] M. Gu, C. Weedbrook, N. C. Menicucci, T. C. Ralph, and P. van Loock, Phys. Rev. A **79**, 062318 (2009).

[17] M. Ohliger, K. Kieling, and J. Eisert, Phys. Rev. A **82**, 042336 (2010).

[18] M. A. Nielsen and C. M. Dawson, Phys. Rev. A **71**, 042323 (2005).

[19] R. Raussendorf and J. Harrington, Phys. Rev. Lett. **98**, 190504 (2007).

[20] T. C. Ralph, A. Gilchrist, G. J. Milburn, W. J. Munro, and S. Glancy, Phys. Rev. A **68**, 042319 (2003).

[21] T. P. Spiller, K. Nemoto, S. L. Braunstein, W. J. Munro, P. van Loock, and G. J. Milburn, New J. Phys. **8**, 30 (2006).

[22] P. van Loock, W. J. Munro, K. Nemoto, T. P. Spiller, T. D. Ladd, S. L. Braunstein, and G. J. Milburn, Phys. Rev. A **78**, 022303 (2008).

[23] W. P. Schleich, *Quantum Optics in Phase Space* (Wiley-VCH, New York, 2001).

[24] P. Grangier, J. A. Levenson, and J.-P. Poizat, Nature **396**, 537 (1998).

[25] S. Rebic, S. M. Tan, A. S. Parkins, and D. F. Walls, J. Opt. B. **1**, 490 (1999).

[26] S. Parkins, Nature **465**, 699 (2010).

[27] D. Achilles, C. Silberhorn, C. Sliwa, K. Banaszek, I. A. Walmsley, M. J. Fitch, B. C. Jacobs, T. B. Pittman, and J. D. Franson, J. Mod. Opt. **51**, 1499 (2004).

[28] A. Divochiy, F. Marsili, D. Bitauld, A. Gaggero, R. Leoni, F. Mattioli, A. Korneev, V. Seleznev, N. Kaurova, O. Minaeva, G. Goltsman, K. G. Lagoudakis, M. Benkhaoul, F. Levy, and A. Fiore, Nature Photonics **2**, 302 (2008).

[29] D. Gross and J. Eisert, Phys. Rev. A **82**, 040303(R) (2010).

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

[31] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Quant. Inf. Comp. **7**, 401 (2007).

[32] M. Ohliger, V. Nesme, D. Gross, Y.-K. Liu, and J. Eisert, arXiv:1111.0853.

[33] C. Schön, K. Hammerer, M. M. Wolf, J. I. Cirac, and E. Solano, Phys. Rev. A **75**, 032311 (2007).

[34] G. K. Brennen and A. Miyake, Phys. Rev. Lett. **101**, 010502 (2008).

[35] X. Chen, B. Zeng, Z. C. Gu, B. Yoshida, and I. L. Chuang, Phys. Rev. Lett. **102**, 220501 (2009).

[36] F. Verstraete and J. I. Cirac, Phys. Rev. Lett. **104**, 190405 (2010).

[37] T. J. Osborne, J. Eisert, and F. Verstraete, Phys. Rev. Lett. **105**, 260401 (2010).

Chapter 5

Continuous-variable quantum compressed sensing

Continuous-variable quantum compressed sensing

M. Ohliger[1,2], V. Nesme[1,3], D. Gross[4], Y.-K. Liu[5], and J. Eisert[1]

[1]*Dahlem Center for Complex Quantum Systems, Physics Department, Freie Universität Berlin, 14195 Berlin, Germany*
[2]*Institute for Physics and Astronomy, University of Potsdam, 14476 Potsdam, Germany*
[3]*Laboratoire d'Informatique de Grenoble, 38400 Saint-Martin-d'Heres, France*
[4]*Institute of Physics, University of Freiburg, 79104 Freiburg, Germany*
[5]*Applied and Computational Mathematics Division, National Institute of Standards and Technology, Gaithersburg, MD, USA*

Abstract

We significantly extend recently developed methods to faithfully reconstruct unknown quantum states that are approximately low-rank, using only a few measurement settings. Our new method is general enough to allow for measurements from a continuous family, and is also applicable to continuous-variable states. As a technical result, this work generalizes quantum compressed sensing to the situation where the measured observables are taken from a so-called tight frame (rather than an orthonormal basis) — hence covering most realistic measurement scenarios. As an application, we discuss the reconstruction of quantum states of light from homodyne detection and other types of measurements, and we present simulations that show the advantage of the proposed compressed sensing technique over present methods. Finally, we introduce a method to construct a certificate which guarantees the success of the reconstruction with no assumption on the state, and we show how slightly more measurements give rise to "universal" state reconstruction that is highly robust to noise.

1 Introduction

One of the most fundamental tasks in quantum mechanics is that of quantum state tomography, i.e., reliably reconstructing an unknown quantum state from measurements. Specifically in the context of quantum information processing in most experiments one has to eventually show what state had actually been prepared. Yet, surprisingly little attention has so far been devoted to the observation that standard methods of quantum state tomography scale very badly with the system size. Only quite recently, novel more efficient methods have been introduced which solve this problem in a more favorable way in the number of measurement settings that need to be performed [1, 2, 3, 4, 5, 6, 12]. This development is more timely than ever, given that the experimental progress with controlled quantum systems such as trapped ions is so rapid that traditional methods of state reconstruction will fail: E.g., 14 ions can already be controlled in their quantum state [7]. Hence, further experimental progress appears severely challenged as long as ideas of reconstruction cannot keep up. Such new methods are based on ideas of quantum compressed sensing [1, 2, 6] — inspired by recent work on low-rank matrix completion [8, 9] — or on ideas of approximating unknown

quantum states with matrix-product states [4]. Indeed, using methods of quantum compressed sensing, one can reduce the number of measurement settings from $n^2 - 1$ in standard methods to $O(rn \log^2 n)$ for a quantum system with Hilbert space dimension n, if the state is of rank r. This is efficient in the sense that the number of measurements required is only slightly greater (by an $O(\log^2 n)$ factor) than the number of degrees of freedom in the unknown state.

These ideas have so far been tailored to the situation where observables are taken from an orthonormal operator basis, which is not always the natural situation at hand. In this paper, we introduce a theory of state reconstruction based on quantum compressed sensing that allows for continuous families of measurements, referred to as *tight frames*, which can be thought of over-complete, non-orthogonal generalization of operator bases. These settings are particularly important in the context of continuous-variables, which are notably used to describe quantum optical systems beyond the single-photon regime. These have drawn a considerable amount of research, both experimentally and theoretically, due to very desired features such as easy preparation and highly efficient detection. Note that when talking about a measurement, we always mean the estimation of an expectation value of an observable for which, of course, several repetition of some experimental procedure are necessary. In this paper we are mainly concerned with the number of distinct observables or *measurement settings* that are needed for tomography.[1]

In this work, we make significant progress towards a full theory of efficient state reconstruction via compressed sensing:

1. We introduce *new incoherence properties* for tight frames, that are sufficient to ensure efficient compressed sensing for low-rank states. This uses an extension of the "golfing" proof technique of [1, 2]. We give examples of tight frames that satisfy these properties. In addition, we show that, if one only wishes to reconstruct "typical" or "generic" low-rank states, there is a much larger class of tight frames that also lead to efficient compressed sensing.

2. We also describe a way to *certify* a successful reconstruction of the state, making our protocol unconditional and heralded. In this way, one does not need to make any a priori assumptions on the unknown state. Our method uses convex duality, and is different from other approaches to certification that focus mainly on pure states [1, 4, 11, 12]. Also, we discuss the *robustness* of the procedure under decoherence, imperfect measurements, and statistical noise. We show that, as long as all those effects are small, it is possible to certify that the reconstructed state is close to the true state.

3. We show that, using an incoherent tight frame, and a slightly larger number of measurements, one can achieve *universal* state reconstruction: a single fixed set of measurements can simultaneously distinguish among all possible low-rank states. This is a qualitatively stronger claim than those shown above, and it is obtained using a different technique, based on the "restricted isometry property" (RIP) [6, 13]. This implies strong error bounds, showing that our procedure for state reconstruction is robust to statistical noise, and that it works even when the true state is full-rank with rapidly decaying eigenvalues (in which case our procedure returns a low-rank approximation to the true state).[2]

---

[1]Other work addresses the number of copies of the unknown state that must be provided [10] — that is, the *sample complexity* of tomography.

[2]As a side note, the RIP-based analysis also shows that the compressed sensing state estimator is nearly

4. We show how our theory can be applied in realistic experimental scenarios, involving pointwise measurements of the Wigner function, and homodyne detection.

5. We demonstrate *numerically* that compressed sensing outperforms the naive approach to tomography not only in the asymptotic limit of large systems but also for system sizes commonly accessible in present day experiments.

This article is organized as follows: We start by introducing quantum compressed sensing in the general setting described by tight frames in Section 2. After discussing a suitable notion of efficiency, we show in Section 3 that efficient compressed sensing is possible if the tight frame fulfills certain incoherence properties. Section 4 is devoted to certified compressed sensing. We discuss how to certify the success of the reconstruction without prior assumptions on the tight frame, both in the ideal case and under the effects of errors. In Section 5 we show universal state reconstruction and error bounds. In Section 6, we investigate applications of the formalism to two common classes of quantum optical experiments; and in Section 7, numerical data, showing the efficiency for small systems, is presented.

2   Quantum compressed sensing

Consider a quantum system with Hilbert space dimension n. In most cases of interest, n is very large, but the states one wants to reconstruct are approximately low-rank, that is, they are well-approximated by density matrices having rank $r \ll n$. (Pure states correspond to the case where $r = 1$.) When dealing with continuous-variable systems, we will truncate the infinite-dimensional Hilbert space and choose n to be some large but finite cutoff. This is unavoidable, if one wants to do tomography as one cannot reconstruct a state that contains an infinite number of completely independent parameters. However, in most experimentally relevant situations, e.g., continuous-variable light modes with finite mean energy, all states can be arbitrarily well approximated by finite-dimensional ones. We will elaborate on this claim when discussing other sources of errors such as decoherence or imperfect measurements.

Compressed sensing contains two key ideas. First, rather than measuring all n^2 degrees of freedom, it is sufficient to measure a randomly chosen subset of about rn degrees of freedom, provided these degrees of freedom satisfy certain *incoherence properties*. Secondly, one can reconstruct the state using an efficient algorithm. The obvious approach of searching for the lowest-rank state compatible with the measurement results leads to a computationally intractable problem (generally NP-hard). Instead, one can perform a *convex relaxation*, and minimize $\|.\|_1$ instead of the rank. Here $\|.\|_p$ stands for the Schatten p-norm: $\|.\|_1$, $\|.\|_2$, and $\|.\| = \|.\|_\infty$ are respectively the trace norm, Frobenius norm, and spectral (or operator) norm.

Let us denote the m measured observables, i.e. Hermitian matrices, by $w_1, \ldots, w_m$, and suppose that we estimate their expectation values (by measuring many copies of the unknown state). Knowing these expectation values (for an unknown state ρ) is equivalent to knowing the value of the *sampling operator* $\mathcal{R}(\rho)$, where we define

$$\mathcal{R} : \sigma \mapsto \frac{n^2}{m} \sum_{i=1}^{m} (w_i, \sigma) w_i \tag{1}$$

---

minimax-optimal [13], and it implies nearly-optimal bounds on the sample complexity of low-rank quantum state tomography [10].

where $(A, B) = \text{Tr}(A^\dagger B)$ is the Hilbert-Schmidt scalar product. In all of our compressed sensing schemes, $w_1, \ldots, w_m$ will be chosen independently at random from some distribution μ. The sampling operator $\mathcal{R}$ is a linear super-operator on the d^2-dimensional real vector space of Hermitian matrices, or operators, $\mathcal{B}(\mathbb{C}^n)$. Such super-operators will always be denoted by capital script letters. Sometimes we will use the notation $\mathcal{R}\sigma$, multiplying the "matrix" $\mathcal{R}$ by the "vector" σ; this means the same thing as $\mathcal{R}(\sigma)$.

Let ρ be the unknown state. In the ideal case, with perfect measurements and no statistical noise, we measure $\mathcal{R}(\rho)$ exactly. Then the procedure to reconstruct ρ can be written as

$$\min_{\sigma \in \mathcal{B}(\mathbb{C}^n)} \|\sigma\|_1, \quad \text{subject to } \mathcal{R}(\sigma) = \mathcal{R}(\rho). \tag{2}$$

Note that a quantum state ρ is a Hermitian matrix with the additional properties $\rho \geq 0$ and $\text{Tr}\,\rho = 1$. However, the reconstruction procedure (2) does not make use of this property and is, therefore, also applicable in more general settings, e.g. matrix completion. This problem can be stated as a semi-definite program (SDP) and, therefore, solved efficiently with many well-developed tools.

In the case of noisy data, we measure $\mathcal{R}(\rho)$ approximately, that is, we measure some b such that $\|b - \mathcal{R}(\rho)\| \leq \delta$, for some norm $\|\cdot\|$ and tolerance δ that are chosen depending on the kind of noise that is expected. The constraint $\mathcal{R}\sigma = \mathcal{R}\rho$ in (2) can then be replaced by $\|\mathcal{R}(\sigma) - b\| \leq \delta$, which implies $\|\mathcal{R}(\sigma - \rho)\| \leq 2\delta$.

We remark that equation (2) is the key to certifying our estimate for ρ. Notice that if the solution σ^* of (2) is unique and fulfills $\|\sigma^*\|_1 = 1$, then it must be the case that $\sigma^* = \rho$. We will show later on how one can test the uniqueness of the solution σ^*, without assuming anything about ρ. (This can be adapted to work with noisy data, without assuming anything about the noise.)

2.1 Measurements and tight frames

When we talk about a compressed sensing scheme, we mean any protocol based on the reconstruction procedure (2), with *some* choice of measurements described by the sampling operator (1). In Refs. [1, 2], the observables were required to be chosen uniformly at random from an operator basis. We substantially generalize these techniques, using the notion of a tight frame, which naturally captures many useful quantum measurements:

Definition 1 (Tight frame)**.** *Let μ be a probability measure on some set S, and for every $\alpha \in S$, let w_α be an observable, i.e., a Hermitian operator, and let $\mathcal{P}_\alpha$ be the (unnormalized) orthogonal projector which acts as $\mathcal{P}_\alpha : \sigma \mapsto (w_\alpha, \sigma)w_\alpha$. We say that $(w_\alpha)_{\alpha \in S}$ is a tight frame if*

$$\int \mathcal{P}_\alpha \mathrm{d}\mu(\alpha) = \frac{\mathbb{1}}{n^2}. \tag{3}$$

This can also be written as $\mathbb{E}_\alpha(n^2 \mathcal{P}_\alpha) = \mathbb{1}$ where α is drawn according to μ. Because we deal with randomly drawn operators very often, α will usually denote a random element of S that has distribution μ. Note that we do not require that $\|w_\alpha\|_2 = 1$ for all α as it will be convenient in many applications. However, we do require a weaker normalization condition: $\mathbb{E}_\alpha[\|w_\alpha\|_2^2] = 1$ which follows by taking the trace of (3).

We also define a generalized notion of a tight frame, where the sampling operator is not a sum of projectors; we will need this later to model homodyne detection on optical

modes, where a single measurement setting provides more information than only one expectation value.

Definition 2 (Generalized tight frame). *Let μ be a probability measure on some set S, and for every $\alpha \in S$ let $\mathcal{Q}_\alpha$ be a positive operator. We say that $(\mathcal{Q}_\alpha)_{\alpha \in S}$ forms a generalized tight frame if*

$$\int \mathcal{Q}_\alpha \mathrm{d}\mu(\alpha) = \frac{\mathbb{1}}{n^2}. \tag{4}$$

We note that the formalism can be also applied to 8-port homodyne detection which corresponds, for a single mode, to projections on coherent states $|\alpha\rangle$ with $\alpha \in \mathbb{C}$.

2.2 Uniqueness of the solution to (2)

For ρ to be the unique solution to (2), any deviation Δ must be either trace-norm increasing, i.e., $\|\rho + \Delta\|_1 > \|\rho\|_1$, or infeasible, i.e., $\mathcal{R}\Delta \neq 0$. This is done by decomposing Δ into a sum $\Delta_T + \Delta_T^\perp$, and then showing that, with high probability, in the case where Δ_T is large, Δ must be infeasible, while in the case where Δ_T is small, Δ must be trace-norm increasing. Here, we denote by T the real space of Hermitian matrices that send the kernel of ρ on its image. In other words, the elements of T are the Hermitian matrices σ whose restriction on and to the kernel of ρ, i.e. $\pi\sigma\pi$ where π is the orthogonal projection on $\mathrm{Ker}\,\rho$, is equal to 0. $\mathcal{P}_T$ denotes the projection on this space T.

Again, in the actual reconstruction, no assumptions have to be made concerning ρ or T. Theorem 1 gives a sufficient condition for uniqueness. The sign function sgn of a Hermitian matrix is defined by applying the ordinary sign function to the matrix' eigenvalues.

Theorem 1 (Uniqueness of the solution). *Let $Y \in \mathrm{range}\,\mathcal{R}$, and set (a) $c_1 := \|\mathcal{P}_T Y - \mathrm{sgn}\,\rho\|_2$, (b) $c_2 := \|\mathcal{P}_T^\perp Y\|$, and (c) $c_3 := \|\mathcal{P}_T \mathcal{R} \mathcal{P}_T - \mathcal{P}_T\|$. If*

$$\frac{1}{n^3}(1 - c_2)\sqrt{\frac{1 - c_3}{m}} - c_1 > 0, \tag{5}$$

then the solution to (2) is unique.

Proof: Δ must be infeasible if $\|\mathcal{R}\Delta\| > 0$ which is the case if

$$\|\mathcal{R}\Delta_T\|_2^2 = (\mathcal{R}\Delta_T, \mathcal{R}\Delta_T) > \|\mathcal{R}\Delta_T^\perp\|_2^2. \tag{6}$$

The right-hand side is bounded as $\|\mathcal{R}\Delta_T^\perp\|_2^2 \leq \|\mathcal{R}\|^2 \|\Delta_T^\perp\|_2^2 \leq n^8 \|\Delta_T^\perp\|_2^2$ while the left-hand side fulfills

$$\|\mathcal{R}\Delta_T\|_2^2 = (\mathcal{R}\Delta_T, \mathcal{R}\Delta_T) \geq \frac{n^2}{m}(\Delta_T, \mathcal{R}\Delta_T)$$

$$\geq \frac{n^2}{m}\left(1 - \|\mathcal{P}_T \mathcal{R} \mathcal{P}_T - \mathcal{P}_T\|\right)\|\Delta_T\|_2^2. \tag{7}$$

Thus, (6) is satisfied if

$$\frac{n^2}{m}\left(1 - \|\mathcal{P}_T \mathcal{R} \mathcal{P}_T - \mathcal{P}_T\|\right)\|\Delta_T\|_2^2 > n^8 \|\Delta_T^\perp\|_2^2, \tag{8}$$

which, using definition (c), is equivalent to

$$\|\Delta_T^\perp\|_2 < \frac{1}{n^3}\|\Delta_T\|_2\sqrt{\frac{1-c_3}{m}}. \tag{9}$$

Using the pinching [19] and Hölder's inequalities, as detailed in Ref. [2], yields

$$\|\rho + \Delta\|_1 \geq \|\rho\|_1 + (\operatorname{sgn}\rho + \operatorname{sgn}\Delta_T^\perp, \Delta). \tag{10}$$

The second term is equal to

$$(\operatorname{sgn}\rho - Y, \Delta_T) + (\operatorname{sgn}\Delta_T^\perp - Y, \Delta_T^\perp) \tag{11}$$

which is, according to (a) and (b), larger than

$$\|\Delta_T^\perp\|_2 - c_2\|\Delta_T^\perp\|_2 - c_1\|\Delta_T\|_2. \tag{12}$$

Inserting this into (9) gives rise to condition (5) and concludes the proof. If all the elements in the tight frame fulfill $\|w_\alpha\|_2 = 1$ we call it normalized and one can bound $\|\mathcal{R}\| \leq n^2$. In this case (5) in Theorem 1 can be weakened to

$$\frac{1}{n}(1-c_2)\sqrt{\frac{1-c_3}{m}} - c_1 > 0. \tag{13}$$

2.3 Efficient quantum compressed sensing

Let ρ be a state of dimension n and rank r. In the compressed sensing method of tomography, we choose m observables $w_1, \ldots, w_m$ randomly from some distribution, measure their expectation values with respect to ρ, then solve (2) to obtain σ^*, which is our estimate of ρ.

For a given state ρ, there is some probability $p_f(\rho)$ that the procedure may fail (i.e., it may return a solution σ^* that is not close to ρ). Note that this probability $p_f(\rho)$ is taken with respect to the random choice of $w_1, \ldots, w_m$, and the random outcomes of the measurements. We say that the method succeeds with high probability if, for every low-rank state ρ, the failure probability is $p_f(\rho)$ small. Equivalently, the method succeeds with high probability if,

> for every low-rank state ρ, most choices of the observables $w_1, \ldots, w_m$ can be used to successfully reconstruct ρ.

Now, the basic question is: how large does m have to be, to ensure that the method succeeds with high probability? A common situation is that the system under consideration consists of k subsystems with local Hilbert space dimension d; then $n = d^k$. Of course, no method of tomography can counter the exponential growth of the required number of measurements in k. Thus, efficiency needs to be regarded relative to the $n^2 - 1$ measurements necessary for standard tomography. As even a pure state needs $\Theta(n)$ parameters to be described, this also is a lower bound to the number of observables that need to be measured. We allow for an additional polylogarithmic overhead and define efficiency as follows:

Definition 3 (Efficient quantum compressed sensing). *Compressed sensing for a state ρ (with dimension n and rank r) is regarded as efficient if: The number of measured observables satisfies $m = O(nr\operatorname{polylog}(n))$, and the probability of failure satisfies $p_f(\rho) < 1/2$.*

If this is the case, $p_f(\rho)$ can be made arbitrarily small by repeating the protocol and using a majority vote among the reconstructed states to get the final result. Then, the probability of failure decays exponentially in the number of repetitions.

Note that this is a very stringent definition of efficiency. One can also merely ask for any scaling of m in $o(n^2)$. Of course, this weaker condition is easier to satisfy, as we shall see later on.

2.4 Sufficient conditions for efficiency

The general theory of quantum compressed sensing, which will be developed here, relies heavily on and significantly extends the analysis for the special case where the observables form an operator basis in Ref. [2]. The hypothesis for Theorem 1 is fulfilled if $c_1 \leq 1/(2n^4)$, $c_2 \leq 1/2$, $c_3 \leq 1/2$ under the additional condition $m < n^2/2$, which can be safely assumed to be true as we are only interested in the regime of $m \ll n^2$. For normalized tight frames, the first condition can be weakened to $c_1 \leq 1/(2n^2)$. We show conditions to the tight frame under which those conditions are fulfilled with high probability.

For efficient compressed sensing to be possible, the observables w_α need to fulfill certain *incoherence properties*. Roughly speaking, the observables are "incoherent" if they have small inner product with every possible state one wishes to reconstruct. For example, operator norm can be a measure of incoherence for reconstructing pure states, since $\|w_\alpha\| = \max_{\langle \psi | \psi \rangle = 1} \langle \psi | w_\alpha | \psi \rangle$. We distinguish two general cases (which we will define more precisely in the following sections):

1. "Fourier-type" compressed sensing, where almost all of the observables have small operator norm. In this case, efficient compressed sensing is possible for any low-rank state.

2. "Non-Fourier type" compressed sensing, where the observables may have large operator norm, but efficient compressed sensing is still possible for certain restricted classes of states, e.g., generic states.

2.5 Fourier-type efficient compressed sensing

The efficiency of a tomography protocol, as given in Definition 3, is a statement about a family of procedures acting on systems with growing dimension n. We now give a sufficient condition for a family of tight frames to allow for efficient compressed sensing.

Theorem 2 (Fourier type). *Let $(w_\alpha(n))_{\alpha \in S}$ be, for any $n > 0$, a tight frame. Let $\rho(n)$ be any state with dimension n and rank r. Let $\nu = O(\text{polylog}(n))$. Set $C(n) := \{\alpha \in S : \|w_\alpha(n)\|^2 > \nu/n\}$ and let $\mu(C(n))$ be the measure of this set. If*

$$\mu(C(n)) \leq \frac{1}{16\sqrt{r}n^2 m}, \tag{14}$$

efficient compressed sensing is possible for the family of states ρ.

Here, the underlying "incoherence property" is the bound on the operator norm of the observables,

$$\|w_\alpha(n)\|^2 \leq \nu/n, \tag{15}$$

which holds for "most" choices of α. If there is no risk of confusion, we will omit the explicit dependencies on n.

7

2.5.1 Perfect Fourier-type case

We have to first consider the case $\mu(C) = 0$. Even though the proof in Ref. [2] can be applied with only minor changes, we state it in a way as complete and still non-technical as possible where we focus on the asymptotic behavior and do not provide explicit constants. We need Lemma 5 from Ref. [2] which reads:

Lemma 1 (Large deviation bound for the projected sampling operator)**.** *For all $t < 2$*

$$\mathbb{P}\left[\|\mathcal{P}_T \mathcal{R} \mathcal{P}_T - \mathcal{P}_T\| > t\right] \leq 4nr \exp\left(-\frac{t^2 \kappa}{8\nu}\right), \tag{16}$$

where $\kappa = m/(nr)$ is the oversampling factor which must fulfill $\kappa = O(\text{polylog}(n))$ for efficiency.

The tool to prove Lemma 1 and other bounds of this form is provided by the operator-Bernstein inequality which was first given in Ref. [17] and which we state here without a proof.

Lemma 2 (Operator-Bernstein inequality)**.** *Let $(X_i)_{i=1,\dots,m}$ be i.i.d. Hermitian matrix-valued random variables with zero mean. Suppose there exist constants V_0 and c such that $\|\mathbb{E}(X_i^2)\| \leq V_0^2$, $\|X_i\| \leq c$ where the latter needs to be true for all realizations of the random variable. Define $A = \sum_i X_i$ and $V = mV_0^2$. Then, for all $t \leq 2V/c$*

$$\mathbb{P}\left[\|A\| > t\right] \leq 2n \exp\left(-\frac{t^2}{4V}\right). \tag{17}$$

The proof of Lemma 1 is given in Ref. [2] but we restate it here because it is quite instructive. Let α be a random variable taking values in S. We define m random variables by $Z_{\alpha_i} = (n^2/m)\mathcal{P}_T \mathcal{P}_{\alpha_i} \mathcal{P}_T$ and $X_{\alpha_i} = Z_{\alpha_i} - \mathbb{E}(Z_{\alpha_i})$. Now $S = \mathcal{P}_T \mathcal{R} \mathcal{P}_T - \mathcal{P}_T = \sum_i X_{\alpha_i}$ and we have to estimate the maximum of $\|X_{\alpha_i}\|$ and the norm of the variance of X_{α_i} in order to apply Lemma 2. From the incoherence condition (15), we get by using the matrix Hölder inequality [19]

$$\|\mathcal{P}_T w_\alpha\|_2^2 = \sup_{\sigma \in T, \|\sigma\|_2 = 1} (w_\alpha, \sigma)^2 \leq 2\nu \frac{r}{n}. \tag{18}$$

This allows us to write

$$\begin{aligned}\|\mathbb{E}(X_{\alpha_i}^2)\| =& \|\mathbb{E}(Z_{\alpha_i}^2) - \mathbb{E}(Z_{\alpha_i})^2\| \\ \leq& \frac{2n\nu r - 1}{m^2}\|\mathcal{P}_T\| \leq \frac{2\nu}{m\kappa}\end{aligned} \tag{19}$$

and

$$\begin{aligned}\|X_{\alpha_i}\| =& \frac{1}{m}\|n^2 \mathcal{P}_T \mathcal{P}_{\alpha_i} \mathcal{P}_T - \mathcal{P}_T\| \\ \leq& \frac{1}{m}\|n^2 \mathcal{P}_T \mathcal{P}_{\alpha_i} \mathcal{P}_T\| = \frac{n^2}{m}\|\mathcal{P}_T w_{\alpha_i}\|_2^2 \\ \leq& \frac{2\nu}{\kappa}.\end{aligned} \tag{20}$$

Here, and in the remainder, statements of the form (20) are meant to hold for all realization of the random variable as needed in the Operator Bernstein inequality. Inserting

now (19) and (20) into Lemma 2 yields Lemma 1 which concludes the proof. Applying Lemma 1 for $t = 1/2$ and choosing $\kappa = O(\mathrm{polylog}(n))$, the probability that $c_3 > 1/2$ can be made arbitrarily small.

Now we have to construct a certificate Y whose projection on T is close to $\mathrm{sgn}\,\rho$. This is done by an iterative process, called the golfing scheme [2]. The m samples are grouped into l groups which are indexed by i and contain m_i samples each. Let $\mathcal{R}_i$ be the sampling operator of the ith group and set $X_0 = \mathrm{sgn}\,\rho$, $X_i = (\mathbb{1} - \mathcal{P}_T \mathcal{R}_i \mathcal{P}_T) X_{i-1}$, $Y_i = \sum_{j=1}^{i} \mathcal{R}_j X_{j-1}$, and $Y = Y_l$.

Again, Lemma 1 can be used to show that with high probability (at the expense of a polylog growth of κ_i)

$$\|X_i\|_2 \le \|\mathcal{P}_T \mathcal{R}_i \mathcal{P}_T - \mathcal{P}_T\| \|X_{i-1}\|_2 \le \frac{1}{2} \|X_{i-1}\|_2, \tag{21}$$

and, therefore, $\|X_i\|_2 \le \sqrt{r} 2^{-i}$ from which we get

$$c_1 = \|X_l\|_2 \le \sqrt{r} 2^{-l} \le \frac{1}{2n^2}, \tag{22}$$

while for the final inequality to hold it is enough to set $l = \Theta(\log(\sqrt{r} n))$. For the last remaining condition we need the subsequent statement:

Lemma 3 (Bound for the orthogonal projection). *Let* $F \in T$ *and* $t \le \sqrt{2/r} \|F\|_2^2$. *Then*

$$\mathbb{P}\left[\|\mathcal{P}_T^\perp \mathcal{R} F\| > t\right] \le 2n \exp\left(-\frac{t^2 \kappa r}{4\nu \|F\|_2^2}\right). \tag{23}$$

Proof: Without loss of generality, consider $\|F\|_2 = 1$ and define the zero-mean random variables $X_{\alpha_i} = (n^2/m) \mathcal{P}_T^\perp w_{\alpha_i} (w_{\alpha_i}, F)$ which fulfill $\sum_i X_{\alpha_i} = \mathcal{P}_T^\perp \mathcal{R} F$. Their variance is bounded by

$$\begin{aligned}
\|\mathbb{E}(X_{\alpha_i}^2)\| &\le \frac{n^4}{m^2} \int \mathrm{d}(\mu)\, (w_\alpha, F)^2 \|(\mathcal{P}_T^\perp w_\alpha)^2\| \\
&\le \frac{\nu}{m\kappa r},
\end{aligned} \tag{24}$$

and their norm by

$$\|X_{\alpha_i}\| \le \frac{n^2}{m} \sqrt{\frac{\nu}{n} \frac{2\nu r}{n}} = \frac{\sqrt{2}\nu}{\sqrt{r}\kappa}. \tag{25}$$

Lemma 3 follows directly from using (24) and (25) in Lemma 2. Now we can bound

$$c_2 = \|\mathcal{P}_T^\perp Y\| \le \frac{1}{4} \sum_{i=1}^{l} 2^{-(i-1)} < \frac{1}{2}. \tag{26}$$

Again, the probability of (26) not being true can be made as small as desired by choosing $\kappa = O(rm\,\mathrm{polylog}(n))$. Of course, this is also true for the total probability of failure which concludes the proof.

2.5.2 Imperfect Fourier-type case

We now show that the incoherence condition may be violated for some of the observables and adapt a technique used in Ref. [14]. Intuitively, when $\mu(C)$ is small enough, we can just abort and restart the reconstruction procedure whenever we encounter a non-incoherent operator during our sampling process. The probability of this to happen is upper bounded by $(16\sqrt{r}n^2)^{-1}$ as obtained from (14) by a union bound over the m measurements. This is equivalent to sampling only from the set $S \setminus C$. The conditional probability distribution on the observables does fulfill the approximate tight-frame condition

$$\|\mathcal{W} - \mathbb{1}\| \leq 1/(8\sqrt{r}), \tag{27}$$

where $\mathcal{W} = n^2 \mathbb{E}(\mathcal{P}_\alpha | E)$ where E is the event that all of the m chosen operators satisfy $\|w_{\alpha_i}\|^2 \leq \nu/n$ and its complement is denoted by E^c. Let $\mathbb{1}_E$ be the indicator function of E. Then, $\mathbb{1} = n^2 \mathbb{E}(\mathcal{P}_\alpha) = n^2 \mathbb{E}(\mathcal{P}_\alpha \mathbb{1}_E) + n^2 \mathbb{E}(\mathcal{P}_\alpha \mathbb{1}_{E^c})$. This leads to

$$\|n^2 \mathbb{E}(\mathcal{P}_\alpha | E) - \mathbb{1}\|\mathbb{P}(E) = \|(1 - \mathbb{P}(E))\mathbb{1} - n^2 \mathbb{E}(\mathcal{P}_\alpha \mathbb{1}_{E^c})\|$$
$$\leq \mathbb{P}(E^c) + n^2 \|\mathbb{E}(\mathcal{P}_\alpha \mathbb{1}_{E^c})\|. \tag{28}$$

With the help of Jensen's inequality, we can simplify $\|\mathbb{E}(\mathcal{P}_\alpha \mathbb{1}_{E^c})\| \leq \mathbb{E}(\|\mathcal{P}_\alpha\|\mathbb{1}_{E^c}) = \mathbb{P}(E^c)$. Inserting this into (28) and rearranging, we get

$$\|n^2 \mathbb{E}(\mathcal{P}_\alpha | E) - \mathbb{1}\| \leq \frac{2n^2 \mathbb{P}(E^c)}{1 - \mathbb{P}(E^c)} \leq 2n^2 \mathbb{P}(E^c). \tag{29}$$

Our claim follows by taking $\mathbb{P}(E^c) = 1/(16\sqrt{r})$ which is always true by a union bound. We now have to justify why the tight frame condition (3) can be replaced by the approximate one in Eq. (27) in the proof of Lemma 1 and Lemma 3. We denote the probability measure which is conditioned on the event E by $\bar{\mu}$.

Lemma 1 provides a bound to

$$\|\mathcal{P}_T (\mathcal{R} - \mathbb{1}) \mathcal{P}_T\| \quad \leq \quad \|\mathcal{P}_T (\mathcal{R} - \mathcal{W}) \mathcal{P}_T\| \tag{30}$$
$$+ \quad \|\mathcal{P}_T (\mathcal{W} - \mathbb{1}) \mathcal{P}_T\|.$$

We define the random variables Z_{α_i} and X_{α_i} as in the proof of Lemma 1 and bound their variance as

$$\begin{aligned}
\|\mathbb{E}(X_{\alpha_i}^2)\| &= \|\mathbb{E}(Z_{\alpha_i}^2) - \mathbb{E}(Z_{\alpha_i})^2\| \\
&\leq \|\mathbb{E}(Z_{\alpha_i}^2)\| + \|\mathbb{E}(Z_{\alpha_i})^2\| \\
&\leq \frac{1}{m^2} \left(2n\nu r + \|\mathcal{W}\|^2\right) \\
&= \frac{1}{m^2} \left(2n\nu r + (\frac{1}{8\sqrt{r}} + 1)^2\right) \leq \frac{4n\nu r}{m^2}, \tag{31}
\end{aligned}$$

and their norm as $\|X_{\alpha_i}\| \leq 2\nu n r/m$. Using the operator Bernstein inequality yields

Lemma 4 (Large deviation bound for the projected sampling operator).

$$\mathbb{P}(\|\mathcal{P}_T \mathcal{R} \mathcal{P}_T - \mathcal{P}_T\| > t) \leq 4nr \exp\left(-\frac{t^2 \kappa}{64\nu}\right), \tag{32}$$

for all $1/(4\sqrt{r}) \leq t \leq 4$.

where we have also used (30) to bound the second term in (27). Thus, up to an irrelevant constant factor, Lemma 4 replaces Lemma 1 wherever it is used.

To also replace Lemma 3, let $F \in T$, $\|F\|_2 = 1$ and note that

$$\|\mathcal{P}_T^\perp \mathcal{R} F\| \leq \|\mathcal{P}_T^\perp (\mathcal{R} - \mathcal{W}) F\| + \frac{1}{8\sqrt{r}}. \tag{33}$$

The random variables are $Z_{\alpha_i} = (n^2/m)\mathcal{P}_T^\perp \mathcal{P}_{\alpha_i} F$ and $X_{\alpha_i} = Z_{\alpha_i} - \mathbb{E}(Z_{\alpha_i})$ where the variance is bounded by

$$
\begin{aligned}
\|\mathbb{E}(X_{\alpha_i}^2)\| &= \|\mathbb{E}(Z_{\alpha_i}^2) - \mathbb{E}(Z_{\alpha_i})^2\| \\
&\leq \|\mathbb{E}(Z_{\alpha_i}^2)\| + \|\mathbb{E}(Z_{\alpha_i})^2\| \\
&\leq \frac{1}{m^2}\left(n\nu + \frac{1}{64r}\right) \leq \frac{2\nu}{m\kappa r}
\end{aligned}
\tag{34}
$$

which gives, together with $\|X_{\alpha_i}\| \leq 2\sqrt{2}\nu/(\sqrt{r}\kappa)$, and an application of the operator-Bernstein inequality the subsequent statement.

Lemma 5 (Bound for the orthogonal projection). *Let $F \in T$ and $1/(2\sqrt{r}) \leq t/\|F\|_2 \leq 2\sqrt{2/r}$. Then*

$$\mathbb{P}\left[\|\mathcal{P}_T^\perp F\| > t\right] \leq 2n \exp\left(-\frac{t^2 \kappa r}{32\nu \|F\|_2^2}\right). \tag{35}$$

Lemma 5 takes the place of Lemma 3 and, again, differs only by a constant factor in the exponent which concludes the proof of Theorem 2.

An example for a Fourier-type frame for which $\mu(C) \neq 0$ is given by the following situation. Here, with some probability, every Hermitian matrix with unit Frobenius norm is drawn in the measurement.

Example 1 (Tight frame containing all Hermitian matrices). *Any $w_\alpha \in \mathcal{B}(\mathbb{C}^n)$ with $\|w_\alpha\|_2 = 1$ can be viewed as a vector on the n^2 dimensional unit sphere. Therefore, on can define a rotationally invariant Haar measure on it. The tight frame formed by the Haar measure on all Hermitian matrices with $\|w_\alpha\|_2 = 1$ fulfills Theorem 2. Therefore, it allows for efficient compressed sensing.*

In order to satisfy Theorem 2, we have to show

$$\mathbb{P}\left(\|w_\alpha\|^2 > \frac{\nu}{n}\right) \leq \frac{1}{16\sqrt{r}n^2 m}, \tag{36}$$

where $\nu = O\left(\text{polylog}\left(n\right)\right)$. To see that this is true, we note that we are dealing with a normalized version of the extensively discussed Gaussian unitary ensemble (GUE) denoted by $\{\bar{w}_\alpha\}$, $w_\alpha = \bar{w}_\alpha/\|w_\alpha\|_2$. Now for all $\delta > 0, \varepsilon > 0$ we have

$$\mathbb{P}\left(\|w_\alpha\| \geq \frac{\delta}{\sqrt{n}}\right) \leq \mathbb{P}\left(\|\bar{w}_\alpha\| > \frac{\delta\varepsilon}{\sqrt{n}}\right) + \mathbb{P}\left(\|\bar{w}_\alpha\|_2 > \varepsilon\right). \tag{37}$$

The first term can be bounded using a result from Ref. [16] yielding

$$\mathbb{P}(\|\bar{w}_\alpha\| > \delta\varepsilon/\sqrt{n}) \leq c_1 \exp(-c_2 n(\delta\varepsilon - 2)^{3/2}) \tag{38}$$

where $c_1, c_2 > 0$ are small constants while for the second term we use the properties of the χ_k^2-distribution which are given the appendix. From this, we get

$$\mathbb{P}\left(\|\bar{w}_\alpha\|_2^2 > 1 - y\right) \leq \exp(-y^2 n^3/4). \tag{39}$$

11

We set $y = 1/2$ and see that (36) is fulfilled for some constant ν when n is large enough.

Product measurements are of great experimental importance: They describe the situation of addressing individual quantum systems, say, ions in an ion trap experiment or individual modes in an optical one. They are described by tight frames which are formed as tensor products of tight frames on the local systems. Given a tight frame which fulfills $\|w_\alpha\|^2 \leq \nu/d$, one can obtain a tight frame on the $n = d^k$ dimensional Hilbert space by forming the k-fold tensor product. The strongest possible incoherence property we can obtain is $\|w_\alpha\|^2 \leq \nu^k/n$. Unless $\nu = 1$, as for the Pauli matrices, ν grows too fast to allow for efficient compressed sensing for all states. This is even true if the incoherence condition may be violated on some set C with $\mu(C) = O(1/\text{poly}(n))$.

2.6 Non-Fourier-type efficient compressed sensing

The conditions in Theorem 2 imply that efficient compressed sensing is possible for *any* low-rank state ρ. This is a quite special situation and for Theorem 2 to be fulfilled, either a very special structure, like the one of the Pauli basis [1], or a large amount of randomness, like in the above example, is needed. As an example for a very different situation, consider the state $\rho = |0\rangle\langle 0|$ together with the observables which corresponds to the sampling of single matrix-entries (or the Hermitian combinations of two of them). Here, one needs to take $\Theta(n^2)$ attempts until one "hits" the single non-zero entry in the upper-left corner. This is not surprising because the operators in this basis fulfill $\|w_\alpha\| = \Theta(1)$. However, for most of the states, efficient compressed sensing is indeed possible in this basis. In Theorem 3, we give a sufficient condition for combinations of states and tight frames to work.

Theorem 3 (Non-Fourier-type efficient compressed sensing). *For a given tight frame* $\{w_\alpha \mid \alpha \in S\}$, *and a given rank-$r$ state ρ, denote by $C \subset S$ the set of observables for which at least one of the following conditions is not fulfilled:*

$$\|\mathcal{P}_T w_\alpha\|_2^2 \leq \frac{2\nu r}{n}, \tag{40}$$

$$(w_\alpha, \text{sgn}\,\rho)^2 \leq \frac{\nu r}{n^2}. \tag{41}$$

If $\mu(C) \leq (16\sqrt{r}n^2 m)^{-1}$, efficient compressed sensing is possible for the state ρ.

The golfing scheme works exactly like in the Fourier-type case, as does the proof of Lemma 1. However, Lemma 5 must be replaced by something else. Again, we use the technique of conditioning which means that we assume the incoherence condition to hold for all operators in the tight frame and the tight frame condition to be approximately true as in (27). First, we need some preparation.

Lemma 6 (Bound to the scalar product). *Let $F \in T$ such that $\|F\|_2 \leq f$, $1/(4\sqrt{r}) \leq f/t \leq 2\sqrt{2/r}$, and*

$$(w_\alpha, F)^2 \leq \frac{\nu f^2}{n^2} \tag{42}$$

for all $\alpha \in S$. Then

$$\mathbb{P}\left(\|\mathcal{P}_T^\perp \mathcal{R}F\| > t\right) \leq 2n \exp\left(-\frac{t^2 \kappa r}{64 \nu f^2}\right). \tag{43}$$

12

Proof: We consider the same same random variables as in the proof of Lemma 4 (note that we have again set $\|F\|_2 = 1$) and bound their variance as

$$\|\mathbb{E}(X_{\alpha_i}^2)\| \leq \frac{n^4}{m^2}\Big(\max_\psi \int \mathrm{d}\mu(\alpha)\,(w_\alpha, F)^2 \langle\psi|w_\alpha^2|\psi\rangle + \frac{1}{64r}\Big)$$
$$\leq \frac{4\nu}{m\kappa r}, \tag{44}$$

where we have used the incoherence property and

$$\left\|\int \mathrm{d}\bar{\mu}(\alpha)\,w_\alpha^2\right\| \leq \frac{2}{n}. \tag{45}$$

To see that (45) holds, we start with

$$\frac{\mathbb{1}}{n} = \int \mathrm{d}\mu(\alpha)\,w_\alpha^2 = (1 - |C|)\int \mathrm{d}\bar{\mu}(\alpha)\,w_\alpha^2 + \int_C \mathrm{d}\mu(\alpha)\,w_\alpha^2 \tag{46}$$

where the first equality follows directly from the tight frame property, c.f. Ref. [2], while the second one stems from the definition of the conditional probability distribution $\bar{\mu}$. Rearranging and taking the norm yields

$$\left\|\int \mathrm{d}\bar{\mu}(\alpha)\,w_\alpha^2\right\| \leq \frac{1}{1-|C|}\left(\frac{1}{n} + |C|\right) \tag{47}$$

which implies (45). Using (44) together with $\|X_{\alpha_i}\| \leq 2\sqrt{2}\nu/(\sqrt{r}\kappa)$ in Lemma 2, we obtain Lemma 6 which concludes the proof.

The above Lemma must by applied for $F = X_0, \ldots, X_l$, i.e., the operators occurring in the golfing scheme. By the second incoherence condition, (42) is fulfilled for $F = X_0$. To ensure that incoherence is preserved during the golfing scheme, we must use a more complicated and technical argument than in Ref. [2] where a union bound over all elements of the operator basis was used which is clearly impossible in a tight frame with an infinite number of elements.

Lemma 7 (Replacing the union bound)**.**

$$\mathbb{P}_\mathcal{R}\left(\xi((\mathbb{1} - \mathcal{P}_T\mathcal{R}\mathcal{P}_T)F) > \frac{1}{2}\|F\|^2\right) \leq 16\sqrt{r}mn^2\exp\left(-\frac{\kappa}{64\xi(F)\nu}\right), \tag{48}$$

where $\xi(F)$ is the smallest number such that

$$\mathbb{P}_\alpha\left((w_\alpha, F)^2 < \xi(F)\right) \leq \frac{1}{16\sqrt{r}n^2m}. \tag{49}$$

Proof: We fix an element w_β from the tight frame and note that for $F \in T$

$$|(w_\beta, \mathcal{P}_T(\mathcal{R} - \mathbb{1})F)| \leq |(w_\beta, \mathcal{P}_T(\mathcal{R} - \mathcal{W})F)|$$
$$+ |(w_\beta, \mathcal{P}_T(\mathcal{W} - \mathbb{1})F)|. \tag{50}$$

The latter term is smaller than $\|\mathcal{W} - \mathbb{1}\|\|F\|_2$. To bound the former term, we define the random variable

$$X_{\alpha_i} = \frac{1}{m}(w_\beta, F) - (w_\beta, \frac{n^2}{m}\mathcal{P}_T w_{\alpha_i})(w_{\alpha_i}, F) \tag{51}$$

13

whose variance is bounded by

$$|\mathbb{E}[X_{\alpha_i}^2]| \leq \frac{2n\xi(F)\nu r}{m^2} + \frac{1}{m^2}\|\mathcal{W} - \mathbb{1}\|^2\|F\|_2^2 \tag{52}$$

and $\|X_{\alpha_i}\| \leq 2(1+n\nu r)\sqrt{\xi(F)}/m$. Using once again the operator Bernstein inequality yields after squaring

$$\mathbb{P}\left((w_\beta, (\mathbb{1} - \mathcal{P}_T\mathcal{R}\mathcal{P}_T)F)^2 > \frac{1}{2}\|F\|_2^2\right) \leq 2\exp\left(-\frac{m}{128nr\xi(F)\nu}\right). \tag{53}$$

Eq. (53) says that the desired property is true with high probability for any fixed w_β. To show that it is also true with high probability for most of the operators, we need a simple fact from probability theory, which is shown in the appendix.

Lemma 8 (Inverting probabilities). *Let X and Y be two measure spaces and denote by $x \sim y$ a relation between the elements $x \in X$ and $y \in Y$. If*

$$\forall x \in X : \mathbb{P}(x \nsim y | y \in Y) \leq p \tag{54}$$

then

$$\mathbb{P}\left(\mathbb{P}(x \nsim y | x \in X) > \beta | y \in Y\right) \leq \frac{p}{\beta} \tag{55}$$

Applying this to (53) and using the definition of $\xi(F)$, one directly obtains (48) which completes the proof of Lemma 7. Now, we can see that $\mu(X_i) \leq 2^{-i}\sqrt{r}\nu/n^2$ which means that Lemma 6 can be applied in the golfing scheme and we have proven Theorem 3.

2.7 Reconstructing generic quantum states

In a next step, we investigate the reconstruction of random quantum states, that are sampled from probability measures that are invariant under the action of the unitary group by conjugation. We show examples of tight frames that satisfy the incoherence properties required in Theorem 3 to allow reconstruction of *most* quantum states.

Theorem 4 (Incoherence properties of generic states). *Let $(w_\alpha)_{\alpha \in S}$ be a (family of) tight frame for which all operators fulfill $\|w_\alpha\|_1 = O(\text{polylog}(n))$, and pick a random rank r quantum state ρ, with a distribution that is invariant under the action of the unitary group. Then the probability that ρ cannot be efficiently reconstructed by compressed sensing vanishes as $O(1/\text{poly}(n))$.*

Note that Theorem 4 holds for all unitarily invariant measures on the quantum states of rank r regardless of the actual distribution of the eigenvalues.

Proof: We first show that for any fixed element of the tight frames, both incoherence properties are fulfilled with high probability. First, we turn to

$$\|P_T w\|_2^2 = \sum_{i,j | \min(i,j) \leq r} |(U^\dagger w U)_{i,j}|^2 \tag{56}$$

where U is a unitary matrix which is chosen according to the Haar measure and we have fixed an element w from the tight frame. We look at the ith row of $U^\dagger w U$ and note that $\sum_j |(U^\dagger w U)_{i,j}|^2 = \sum_j |(U^\dagger w)_{i,j}|^2$. We write w_j for the jth column vector

of w and note that $U^\dagger w_j / \|w_j\|$ is just a random vector on a sphere . Thus, the squares of its coordinates are concentrated around $1/n$, c.f. the appendix, and we get

$$\mathbb{P}_U \left(\frac{|(U^\dagger w)_{i,j}|^2}{\|w_j\|^2} > \frac{\nu}{n} \right) \leq 2 \exp \left(-\frac{\nu}{8} \right). \tag{57}$$

Using this in Eq. (56), inserting $\sum_i \|w_i\|^2 = \|w\|_2^2 = 1$, and applying a union bound yields

$$\mathbb{P}_U \left(\|\mathcal{P}_T w\|_2^2 > \frac{2\nu r}{n} \right) \leq 2nr \exp \left(-\frac{\nu}{8} \right). \tag{58}$$

Employing again Lemma 8, this implies

$$\mathbb{P}_U \left(\mathbb{P}_w \left(\|\mathcal{P}_T w\|_2^2 > \frac{2\nu r}{n} \right) > \frac{1}{16\sqrt{r}n^2 m} \right) \ \leq \ 32 r^{3/2} n^3 m \exp \left(-\frac{\nu}{8} \right). \tag{59}$$

where w is chosen according to the probability distribution of the tight frame. By allowing ν to grow polylogarithmically in n, this probability vanishes polynomially in n which means that it is violated to much only for a proportion of state vanishing polynomially as n grows. Now we turn to the second non-Fourier incoherence condition. Decomposing w as a sum of projectors on orthogonal subspaces $w = \sum_k \lambda_k |\Psi_k\rangle\langle\Psi_k|$, we can write

$$|(w, \operatorname{sgn} \rho)| \leq \sum_{i=0}^r \sum_k |\lambda_k| |\langle i|U^\dagger|\Psi_k\rangle|^2. \tag{60}$$

Using the concentration of measure on the sphere and $\sum_k |\lambda_k| = \|w\|_1$ yields

$$\mathbb{P} \left((w, \operatorname{sgn} \rho)^2 > \frac{r^2 \nu}{n^2} \|w\|_1^2 \right) \leq 2nr \, e^{-\sqrt{\nu}/8}, \tag{61}$$

which finally gives

$$\mathbb{P}_U \left(\mathbb{P}_w \left((w, \operatorname{sgn} \rho)^2 > \frac{r^2 \nu}{n^2} \|w\|_1^2 \right) > \frac{1}{16\sqrt{r}mn^2} \right)$$
$$\leq 32 r^{3/2} mn^3 \exp \left(-\frac{\sqrt{\nu}}{8} \right). \tag{62}$$

Since the additional factor of r can be absorbed in ν, Theorem 4 follows from Eq. (62).

Tight frames for which this is the case include those where the rank of the operators does not grow with n. The other extreme is given by the Pauli basis: From $\|w\|_2 = 1$ and $\|w\| = 1/\sqrt{n}$ it follows that $\|w\|_1 = \sqrt{n}$. Colloquially speaking, a small spectral norm implies a large trace norm and vice versa. Thus, we have two classes of tight frames (Fourier likes ones and the ones with small 1-norm) for which efficient compressed sensing is efficiently possible. Because they represent in some sense the two extreme cases (flat spectra vs. concentrated spectra), we have some reason to believe that this is indeed true for *any* tight frame.

3 Certification

3.1 Ideal case

Theorems 2 and 3 show that efficient compressed sensing is possible in a vast number of situations. They are stated in the asymptotic regime for clarity but could be furnished

with reasonable prefactors for finite Hilbert-space dimension n. However, when using compressed sensing in actual experiments, one encounters three main problems.

- Firstly, the necessary number of measurements as calculated from the incoherence properties of the employed tight frame might still be too large to be feasible.

- Secondly, repetition of the experiments to increase the probability of success to a satisfactory value may be expensive or difficult.

- Thirdly, it is unknown how close to low-rank the state actually is. After all, no assumptions are made about the unknown input state.

The solutions to those problems is provided by certification. Instead of theoretically constructing some certificate based on ρ with the help of the golfing scheme, we use the solution of the minimization problem σ^* to explicitly check whether the conditions for Theorem 1 are satisfied for σ^*. The candidate for the certificate can be calculated as

$$Y = \mathcal{R}\mathcal{P}_{T'}(\mathcal{P}_{T'}\mathcal{R}\mathcal{P}_{T'})^{-1}\operatorname{sgn}\sigma^* \tag{63}$$

where $\mathcal{P}_{T'}$ is obtained like $\mathcal{P}_T$ but with ρ replaced by σ^* and M^{-1} denotes the Moore-Penrose pseudo inverse of M. One can now check whether (5) is fulfilled. If the conditions for Theorem 1 are fulfilled and $\|\sigma^*\|_1 = 1$, the solution must be unique and equal to the state ρ, i.e. tomography was successful.

3.2 Errors and noise

For compressed sensing to work in a realistic setting, the reconstruction procedure must be robust, i.e., small errors introduced by decoherence, errors stemming from imperfect measurements, and statistical noise due to the fact that every observable is only measured a finite number of times, should only lead to small errors in the reconstructed state. In addition, the Hilbert space might be infinite-dimensional. When the mean energy, and therefore, the mean photon number N_{mean}, is finite, the error made by truncating the Hilbert space at photon number N vanishes as

$$\|\rho_{\mathrm{trunc}} - \rho\|_1 \leq 3\sqrt{\frac{N_{\mathrm{mean}}}{N+1}} = \varepsilon \tag{64}$$

which is shown in the appendix. This means that the expectation values with respect to the truncated state are close to the actually measured ones, i.e.,

$$|\operatorname{Tr}(w\rho_{\mathrm{trunc}}) - \operatorname{Tr}(w\rho)| \leq \varepsilon, \tag{65}$$

for all w such that $\|w\| \leq 1$.

We assume that the observed data correspond to a matrix $\tilde{\rho}$ (not necessarily a state) with $\|\mathcal{P}_{\mathcal{R}}(\tilde{\rho} - \rho)\|_2 \leq \delta$ where ρ is the low-rank, infinite-dimensional state, i.e., the errors made by truncating to a finite-dimensional Hilbert space are already included in δ, and where we denote by $\mathcal{P}_{\mathcal{R}}$ the projection to the image of the sampling operator. Such a tube condition is satisfied with very high probability for realistic error models like Gaussian noise [1, 18]. We relax the conditions in (2) to

$$\|\mathcal{P}_{\mathcal{R}}(\sigma - \tilde{\rho})\|_2 \leq \delta. \tag{66}$$

The solution of the SDP might not be of low rank. Because a low-rank state is needed for the construction of the certificate Y, we truncate σ^* to the q largest eigenvalues

(call this σ_q^*) and obtain $\mathcal{P}_{T'}$ as above. As $r = \text{rank}\,\rho$ is in general not known, one has to perform the truncation of σ^* and the subsequent construction of the certificate Y for $q = 1, 2, \ldots$ until a valid Y, as to be specified below, has been found. If this is not the case, the number of measurements was not enough and needs to be increased.

To provide an error bound, we denote the 2-norm error made by the truncation of σ^* to rank q by ε and obtain from the triangle inequality

$$\|\mathcal{P}_{\mathcal{R}}(\sigma_q^* - \rho)\|_2 = \|\mathcal{P}_{\mathcal{R}}(\sigma_q^* - \sigma^*)\|_2 + \|\mathcal{P}_{\mathcal{R}}(\sigma^* - \tilde\rho)\|_2 + \|\mathcal{P}_{\mathcal{R}}(\tilde\rho - \rho)\|_2 \leq \varepsilon + 2\delta. \quad (67)$$

We calculate a candidate for a certificate as $Y = \mathcal{R}\mathcal{P}_{T'}(\mathcal{P}_{T'}\mathcal{R}\mathcal{P}_{T'})^{-1}\,\text{sgn}\,\sigma_q^*$ where T' is obtained from σ_q^*. If Y is valid, i.e., $\|\mathcal{P}_{T'^\perp}Y\| \leq 1/2$, and $\mathcal{P}_{T'}\mathcal{P}_{\mathcal{R}}\mathcal{P}_{T'} \geq (p/2)\mathcal{P}_{T'}$ with $p = m/n^2$, then the proof of Theorem 7 in Ref. [18] yields the robustness result

$$\|\sigma_q^* - \rho\|_2 \leq \left(4\sqrt{\frac{(2+p)n}{p}} + 2\right)(2\delta + \varepsilon). \quad (68)$$

By the equivalence of the norms, this also provides a 1-norm bound at the expense of an additional factor $\sqrt{n}$.

Thus, with no further assumption than 2-norm closeness of the observations to the state of interest it is possible to obtain a certified reconstruction which is also close to the state of interest. In this sense, quantum compressed sensing can achieve assumption-free certified quantum state reconstruction in the presence of errors. This discussion applies to box errors, where each of the expectation values is assumed to be contained in a certain interval. The discussion of other error models will be the subject of forthcoming work.

4 Universal quantum compressed sensing

4.1 Universal quantum state reconstruction

The preceding discussion has focused on claims of the following form:

> For every low-rank state ρ, most choices of the observables $w_1, \ldots, w_m$ can be used to successfully reconstruct ρ.

In some situations, however, one can actually prove a much stronger statement, in which the order of the quantifiers is reversed:

> Most choices of the observables $w_1, \ldots, w_m$ will have the property that, for every low-rank state ρ, the observables $w_1, \ldots, w_m$ can be used to successfully reconstruct ρ.

This is known as *universal* reconstruction; more simply, it says that a fixed set of observables $w_1, \ldots, w_m$ can distinguish among all low-rank states *simultaneously*. Besides being of conceptual interest, universal reconstruction also implies stronger error bounds for reconstruction from noisy data.

Formally, we say that our method performs universal compressed sensing if $p_{fu} < 1/2$, where p_{fu} is the "universal" failure probability. That is, we define p_{fu} to be the probability (with respect to the choice of observables $w_1, \ldots, w_m$) that there exists a state ρ (with dimension n and rank r) such that the method fails with probability $> 1/2$ (where this last probability is taken with respect to the random measurement outcomes).

17

Definition 4 (Efficient universal quantum compressed sensing). *Universal compressed sensing (with dimension n and rank r) is regarded as efficient if: the number of measured observables satisfies $m = O(nr \operatorname{polylog}(n))$, and the probability of failure satisfies $p_{fu} < 1/2$.*

4.2 Universal reconstruction using any Fourier-type tight frame

In this section, we show that measurements using a Fourier-type tight frame lead to efficient universal quantum compressed sensing. This result can be viewed as a companion to Theorem 2. Essentially, it says that, by using a slightly larger number of measurements (by a polylog(n) factor), one can construct (with high probability) a single, *fixed* set of measurements that can reconstruct *all* possible states of rank r and dimension n. In addition, universal reconstruction implies very strong error bounds, in the case of noisy data; we will discuss this in the following section.

Theorem 5 (Universal reconstruction). *Let $(w_\alpha)_{\alpha \in S}$ be a tight frame. Let $\nu = O(\operatorname{polylog}(n))$, and suppose that, for all $\alpha \in S$, $\|w_\alpha\|^2 \leq \nu/n$. Then efficient universal compressed sensing (for states of rank r and dimension n) is possible.*

This proof of this theorem is a straightforward generalization of [6]. First, we define the sampling operator to be $\mathcal{A} : \mathbb{C}^{n \times n} \to \mathbb{R}^m$,

$$(\mathcal{A}(\sigma))_i = \frac{n}{\sqrt{m}}(w'_i, \sigma), \quad i = 1, \dots, m. \tag{69}$$

This is related to the notation used in previous sections by $\mathcal{A}^\dagger \mathcal{A} = \mathcal{R}$. (As before, the observables $w'_1, \dots, w'_m$ are sampled independently from the distribution μ on the tight frame, and $(A, B) = \operatorname{Tr}(A^\dagger B)$ is the Hilbert-Schmidt inner product.)

A key tool in the proof is the *restricted isometry property* (RIP) [23]. We say that $\mathcal{A}$ satisfies the RIP if there exists some constant $\delta \in [0, 1)$ such that, for all rank-r n-dimensional states σ,

$$(1 - \delta)\|\sigma\|_2 \leq \|\mathcal{A}(\sigma)\|_2 \leq (1 + \delta)\|\sigma\|_2. \tag{70}$$

In geometric terms, the set of all low-rank states forms an $O(rn)$-dimensional manifold in $\mathbb{C}^{n \times n}$, and $\mathcal{A}$ satisfies the RIP if it embeds this manifold into $\mathbb{C}^m$, with constant-factor distortion.

The importance of the RIP stems from the following fact: when $\mathcal{A}$ satisfies RIP, one can reconstruct any low-rank state ρ from noiseless data $\mathcal{A}(\rho)$, by solving a trace-minimization convex program:

$$\min \|\sigma\|_1, \quad \text{subject to } \mathcal{A}(\sigma) = \mathcal{A}(\rho). \tag{71}$$

This follows from a standard argument of [23]. This result can be generalized to the case of noisy data; we will discuss this in the following section.

It now remains to prove that, when the observables w_α are chosen from a Fourier-type tight frame (i.e., they satisfy $\|w_\alpha\|^2 \leq \nu/n$), the sampling operator $\mathcal{A}$ satisfies RIP with high probability. Intuitively, one first shows that, for any particular low-rank state σ, and a random choice of measurements $w'_1, \dots, w'_m$, the sampling operator $\mathcal{A}$ satisfies equation (70) with high probability. After this comes the main part of the argument. Let $p_f(\sigma)$ denote the probability of failure on a given state σ. One now needs to upper-bound the probability of a failure on any one of the states σ. The simplest

approach is to assume that the failure events are disjoint, and so the probabilities sum up — this is the union bound, and it does not give a useful bound in this case. Instead, one uses an "entropy argument" that exploits the fact that failure events are not disjoint: failures on nearby states are correlated.

Formally, the entropy argument is carried out using Gaussian processes and Dudley's inequality (following [20, 21]), and using bounds on covering numbers of the trace-norm ball due to [22]. The proof is essentially the same as in [6]; the original proof in [6] handles the case where the w_α form an incoherent orthonormal basis, but the same proof goes through unchanged for a Fourier-type tight frame. This shows that, if the number of measurements satisfies $m \geq C\nu r n \log^6 n$ (for some constant C), then with high probability the sampling operator $\mathcal{A}$ satisfies the RIP (for rank r and dimension n).

4.3 Robust reconstruction from noisy data

More interestingly, RIP implies strong error bounds in the case of noisy data [13]. We sketch the basic idea here. Suppose one observes $y = \mathcal{A}(\rho) + z$, where z denotes a noise component. Then one can replace (71) with other estimators, such as the matrix Dantzig selector:

$$\min \|\sigma\|_1 \quad \text{such that} \quad \|\mathcal{A}^\dagger(y - \mathcal{A}(\sigma))\| \leq \lambda, \tag{72}$$

or the matrix Lasso:

$$\min \tfrac{1}{2}\|\mathcal{A}(\sigma) - y\|_2^2 + \mu\|\sigma\|_1. \tag{73}$$

(See Ref. [13] for details about setting the parameters λ and μ.)

When the noise vector z is normally distributed, one can show particularly nice error bounds. These hold even for states ρ that are *full-rank* (though ρ must at least have decaying eigenvalues, for the bounds to be useful) [13] (see also [6]). Suppose that ρ is arbitrary, and one simply assigns some value for r, and measures $m = O(\nu r n \log^6 n)$ observables. Then let σ^* denote the solution returned by either of the above estimators. Intuitively, one expects that σ^* should reconstruct the first r eigenvectors of ρ. One can prove a bound that is consistent with this intuition: the squared 2-norm error $\|\sigma^* - \rho\|_2^2$ will be nearly proportional (up to log factors) to the total variance of the noise acting on the first r eigenvectors of ρ, plus the squared 2-norm of the "tail" of ρ (consisting of its last $n - r$ eigenvectors).

5 Applications

We now demonstrate how our theory can be applied to some common experimental setups in quantum optics. We show how pointwise measurements of the Wigner function, and histograms obtained using homodyne detection, can be expressed as measurements using tight frames, and generalized tight frames. Furthermore, we propose efficient compressed sensing schemes (with Fourier-type tight frames) using these measurements.

5.1 Homodyne detection

The most common way to do quantum state tomography on continuous-variable light modes is based on homodyne detection, which is done by combining the light field with

a mode in a strong coherent state, called the local oscillator, in an interferometer and measuring the difference of the intensities on the two output ports [29, 30, 31]. This amounts to sampling $x \in \mathbb{R}$ according to the one-dimensional probability distribution given by the Radon transform (at angle θ) of the Wigner function, i.e.,

$$\mathbb{P}_\theta(x) = \int W(x \cos \theta - p \sin \theta, x \sin \theta + p \cos \theta) \mathrm{d}p. \tag{74}$$

The angle θ is chosen by phase-shifting the mode with respect to the local oscillator.

For a general quantum state with maximal photon number N, $N + 1$ equidistant choices of $\theta \in [0, \pi)$ are sufficient and necessary to reconstruct the state by an inverse Radon transform of Eq. (74) or by using pattern functions [29, 30]. Here we show how these measurements can be described by a generalized tight frame. A tight frame by itself does not suffice because here every measurement setting, i.e., every choice of θ, does not only give a single number as a result but an entire distribution $\mathbb{P}_\theta$.

A key observation is that the Fourier transform of the probability distribution (74) is identical to the characteristic function, i.e., the Fourier transform of the Wigner function, written in radial coordinates. We define

$$\tilde{W}(u, v) = \int \mathrm{d}x \, \mathrm{d}p \, W(x, p) \exp[-i(ux + vp)] \tag{75}$$

which fulfills

$$\tilde{\mathbb{P}}_\theta(\zeta) = \tilde{W}(\zeta \cos \theta, \zeta \sin \theta) \tag{76}$$

where $\tilde{\mathbb{P}}_\theta(\zeta) = \int dx \mathbb{P}_\theta(\zeta) \exp(-i\zeta x)$.

This allows us to write the projector (corresponding to measurement setting θ and outcome ζ) as

$$(\mathcal{P}_\theta(\zeta))_{(i,j),(k,l)} = \tilde{W}_{|j\rangle\langle i|}(\zeta \cos \theta, \zeta \sin \theta) \tilde{W}^*_{|l\rangle\langle k|}(\zeta \cos \theta, \zeta \sin \theta). \tag{77}$$

Because choosing a measurement setting does not mean choosing values for θ and ζ, but rather only choosing a phase θ and obtaining a whole "slice" of the characteristic function, the operator corresponding to a measurement setting is

$$\mathcal{P}_\theta = \int \mathrm{d}\zeta \, \mathcal{P}_\theta(\zeta). \tag{78}$$

It is easy to check that $\mathcal{P}_\theta$ fulfills

$$\frac{1}{\pi} \int_0^\pi \mathrm{d}\theta \, \mathcal{P}_\theta = \frac{\mathbb{1}}{n^2}, \tag{79}$$

which implies that it satisfies Definition 2 and forms a generalized tight frame.

5.2 Efficient compressed sensing using homodyne measurements

In the previous subsection, we have introduced the generalized tight frame corresponding to homodyne detection. This can be combined with the convex program in equation (2) to perform state reconstruction. In Section 6, we show by means of a numerical simulation that this procedure performs well in practice. However, our theoretical analysis does not apply to this procedure, due to the generalized tight frame; it would be interesting to try to extend our theoretical results to this case.

20

In this section, we present a different way of using homodyne detection to reconstruct low-rank states, which is a little less direct, but does have a rigorous guarantee of success. We will do three things. First, we will show how homodyne measurements can be used to estimate expectation values of displacement operators. Then, we will use (scaled) displacement operators to construct a tight frame. Finally, we will show that this tight frame has Fourier-type incoherence. By combining these pieces, we will then get an efficient compressed sensing scheme.

Before continuing, we note that $D(\alpha)$ cannot be directly measured as it is not Hermitian. However, one can also use 8-port homodyning to directly measure the observables $|\alpha\rangle\langle\alpha| = D(\alpha)|0\rangle\langle0|D^\dagger(\alpha)$ [28]. Because the experimental effort is higher, compared to standard homodyning, we will not discuss this scheme here.

Define the displacement operators

$$D(\alpha) = e^{-|\alpha|^2/2}e^{\alpha a^\dagger}e^{-\alpha^* a}, \qquad \alpha \in \mathbb{C}. \tag{80}$$

Note that we have the identities $D(\alpha) = e^{\alpha a^\dagger - \alpha^* a} = e^{|\alpha|^2/2}e^{-\alpha^* a}e^{\alpha a^\dagger}$.

Now recall the definition of the characteristic function [32]:

$$C^{(s)}(\beta) = \text{Tr}(e^{i\beta a^\dagger + i\beta^* a}\rho), \qquad \beta \in \mathbb{C}. \tag{81}$$

Setting $\alpha = i\beta$, we see that $C^{(s)}(\beta)$ is precisely the expectation value of the displacement operator $D(\alpha)$. On the other hand, $C^{(s)}(\beta)$ is also equal to $\tilde{W}(\beta)$, the (two-dimensional) Fourier transform of the Wigner function $W(\xi)$. This in turn is related, via equation (76), to the probability distribution $\mathbb{P}_\theta(x)$, which we can sample using homodyne detection.

Thus, we can estimate the expectation value of a displacement operator $D(\alpha)$ as follows: set $\beta = -i\alpha$, and make homodyne measurements with phase angle $\theta = \arg(\beta)$. This produces several points $x_1, \ldots, x_\ell \in \mathbb{R}$ sampled from the distribution $\mathbb{P}_\theta(x)$. Then set $\zeta = |\beta|$, and compute $\frac{1}{\ell}\sum_{i=1}^{\ell}\exp(-i\zeta x_i)$. This gives an estimate for $\tilde{\mathbb{P}}_\theta(\zeta) = \tilde{W}(\beta) = C^{(s)}(\beta)$, which is the desired expectation value.

Note that a lossy detector (i.e., one with efficiency less than 1) has the effect of convolving the true Wigner function $W(\xi)$ with a Gaussian, to produce the empirically observed Wigner function [33]. This is equivalent to pointwise multiplying the characteristic function $C^{(s)}(\beta)$ with a Gaussian envelope. We can compensate for this by re-scaling $C^{(s)}(\beta)$ at each point β, provided that our raw estimates of $C^{(s)}(\beta)$ are sufficiently precise, and the detector efficiency is not too poor.

Next, we will construct a tight frame using the displacement operators $D(\alpha)$. Note that the $D(\alpha)$ form an orthonormal basis for the state space [32]:

$$\rho = \frac{1}{\pi}\int_{\mathbb{C}} D(\alpha)\text{Tr}(D(\alpha)^\dagger\rho)\mathrm{d}\alpha, \qquad \text{for all states } \rho, \tag{82}$$

where we are taking a 2-dimensional integral over the complex plane. Now suppose we sample α from a 2-dimensional Gaussian distribution on the complex plane with width σ (which we will choose later). This distribution has probability density

$$P_G(\alpha) = \frac{1}{2\pi\sigma^2}e^{-|\alpha|^2/2\sigma^2}. \tag{83}$$

Define scaled displacement operators

$$\tilde{D}(\alpha) = \sqrt{2}\sigma e^{|\alpha|^2/4\sigma^2}D(\alpha). \tag{84}$$

21

Then we can rewrite (82) as

$$\rho = \int_{\mathbb{C}} \tilde{D}(\alpha) \text{Tr}(\tilde{D}(\alpha)^\dagger \rho) P_G(\alpha) \mathrm{d}\alpha, \qquad \text{for all states } \rho. \tag{85}$$

This is (up to normalization) a tight frame for the full, infinite-dimensional state space.

In fact, we are only interested in the finite-dimensional subspace consisting of states with at most N photons; this subspace is isomorphic to $\mathbb{C}^{(N+1)\times(N+1)}$. So we will truncate the above operators. Let Π_N be the projector onto span$\{|0\rangle, |1\rangle, \ldots, |N\rangle\}$ (where the $|j\rangle$ are Fock basis states). Then define truncated displacement operators

$$D_N(\alpha) = \Pi_N D(\alpha) \Pi_N, \qquad \text{and} \qquad \tilde{D}_N(\alpha) = \Pi_N \tilde{D}(\alpha) \Pi_N. \tag{86}$$

Then the operators $w_\alpha = \frac{1}{N+1} \tilde{D}_N(\alpha)$ form a tight frame for $\mathbb{C}^{(N+1)\times(N+1)}$, as desired:

$$\frac{1}{(N+1)^2} \rho = \int_{\mathbb{C}} w_\alpha \text{Tr}(w_\alpha^\dagger \rho) P_G(\alpha) d\alpha, \qquad \text{for all } \rho \in \mathbb{C}^{(N+1)\times(N+1)}. \tag{87}$$

Finally, we set $\sigma = \sqrt{2N \log(1 + 4N)}$, and we claim that the above tight frame $\{w_\alpha\}$ is Fourier-type incoherent, in the sense of Theorems 2 and 5. More precisely, we claim that

$$\|\tilde{D}_N(\alpha)\| \leq \sqrt{2} e \sigma = 2e\sqrt{N \log(1 + 4N)}, \qquad \text{for all } \alpha \in \mathbb{C}; \tag{88}$$

we will prove this below. This directly implies

$$\|w_\alpha\| \leq \frac{2e\sqrt{\log(1 + 4N)}}{\sqrt{N}}, \qquad \text{for all } \alpha \in \mathbb{C}. \tag{89}$$

Then, by Theorems 2 and 5, we have an efficient compressed sensing scheme.

We now show why (88) holds. First, note that while the displacement operators $D(\alpha)$ are unitary, the scaled operators $\tilde{D}(\alpha)$ are unbounded. However, when α is small, this is not a problem. In particular, when $|\alpha| \leq 2\sigma$, we can just use the trivial bound

$$\|\tilde{D}_N(\alpha)\| \leq \|\tilde{D}(\alpha)\| \leq \sqrt{2}\sigma e^{|\alpha|^2/4\sigma^2}, \tag{90}$$

which implies (88).

It remains to consider the case where $|\alpha| > 2\sigma$. In this case, $\tilde{D}(\alpha)$ is large, but it acts mostly on states with more than n photons, so the truncated operator $\tilde{D}_N(\alpha)$ is small. Using a straightforward calculation, we can bound $\tilde{D}_N(\alpha)$ in the 2-norm, which implies (88). See the appendix for details.

5.3 Pointwise measurements of the Wigner function

A quantum state ρ of a single optical mode can be represented in phase space by a real Wigner function $W_\rho : \mathbb{R}^2 \to \mathbb{R}$ [28]. For a single mode it is given by, c.f. Ref. [26, 27],

$$W_\rho(\xi) = \frac{2}{\pi} \text{Tr}\left((-1)^{\hat{n}} D(\xi)^\dagger \rho D(\xi)\right) \tag{91}$$

where $(-1)^{\hat{n}}$ is the parity operator where $\xi = (x, p) \in \mathbb{R}^2$, $D(\xi)$ is the displacement operator which becomes the one defined in (80) by setting $\alpha = (1/\sqrt{2})(\xi_1 + i\xi_2)$. With

the same convention, we will, whenever it is convenient, regard the Wigner function as a function of a complex variable.

Eq. (91) allows pointwise measurement of the Wigner function by a displacement in phase space followed by a measurement of the parity of the photon number. This has already been experimentally performed for optical fields in a cavity [24] and for pulsed single photons (for the special case of a rotationally invariant state) [25]. We consider a single mode containing up to N photons and, therefore, Hilbert space dimension $n = N + 1$. Measuring the Wigner function at a point α amounts to a measurement of the observable

$$w_\alpha = \sqrt{2\pi} \frac{2}{\pi} D(\alpha)(-1)^{\hat{n}} D^\dagger(\alpha). \tag{92}$$

We make again use of the probability density P_G of Eq. (83) and define scaled, truncated operators $\tilde{w}_\alpha = n^{-1} P_G(\alpha)^{-1/2} \Pi_N w_\alpha \Pi_N$. They form a tight-frame on the truncated Hilbert space when the sampling is performed according to P_G.

We now proceed exactly as in the previous section to show that the operator norm of the $\tilde{w}$ is small enough for the Fourier type incoherence property of Theorem 2. We do not give explicit constants but focus on the asymptotic scaling in n. If $|\alpha| \leq 2\sigma$, we get the bound $\|\tilde{w}_\alpha\| \leq 4e\sigma/n$. We will show that if we set $\sigma = \sqrt{n} \log n$ one has $\|\exp(|\alpha|^2/(2\sigma^2))w_\alpha\| \leq 1$ for all α with $|\alpha| > 2\sigma$ whenever n is large enough which implies the requirements of Theorem 2. We need the matrix elements $\langle l|w_\alpha|k\rangle = W_{|k\rangle\langle l|}(\alpha)$. To calculate them, first let $\xi = (x, p)$ and recall the definition of the Wigner function [28]:

$$W_{|l\rangle\langle k|}(x, p) = \frac{1}{\pi} \int dy\, \psi_l^*(x + y)\psi_k(x - y)e^{2ipy}. \tag{93}$$

where we remember the identification $\alpha = (1/\sqrt{2})(x + ip)$. Inserting the eigenfunctions of the harmonic oscillator ψ_i, using the properties of the occurring Hermite polynomials, and performing the integral allows to write

$$W_{|l\rangle\langle k|}(x, p) = \frac{(-1)^{l+k} e^{x^2}}{\pi \sqrt{2^{l+k} l! k!}} \frac{\partial^{l+k}}{\partial x^k \partial x'^l} G(x, x, p')\bigg|_{x'=x} \tag{94}$$

with the generating function

$$G(x, x', p) = e^{-p^2 + 2ip(x-x') - 2xx'}. \tag{95}$$

From this, one gets the bound, which is by no means tight but strong enough, $|W_{|k\rangle\langle l|}(\alpha)| \leq n^n(2|\alpha|)^n \exp(-2|\alpha|^2)$ which allows us to write

$$\|\exp(|\alpha|^2/(2\sigma^2))w_\alpha\| \leq \|\exp(|\alpha|^2/(2\sigma^2))w_\alpha\|_2$$
$$\leq \exp\left(-2|\alpha|^2 + (n+2)\log n + 2n\log(2|\alpha|) + \frac{|\alpha|^2}{2\sigma^2}\right). \tag{96}$$

We now set $\sigma = \sqrt{n} \log n$ and get, for large enough n, a bound valid for all α with $|\alpha| > 2\sigma$ which reads

$$\|\exp(|\alpha|^2/(2\sigma^2))w_\alpha\| \leq \exp\left(-2n\log^2 n + (n+2)\log n + 2n\log(4\sqrt{n}\log n)\right). \tag{97}$$

As the first term in the exponent grows fastest, one has $\|\exp(|\alpha|^2/(2\sigma^2))w_\alpha\| \leq 1$ for sufficiently large n. Thus, there is some $C > 0$ such that $\|\tilde{w}_\alpha\| \leq C \log n/\sqrt{n}$ which means that the pointwise Wigner function measurement is of Fourier type and, therefore, can be used for efficient compressed sensing.
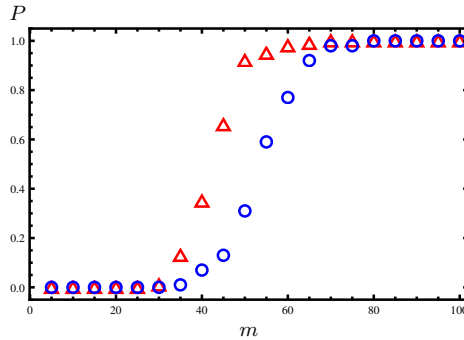
23

Figure 1: (color online) Reconstruction of a random pure state on 4 qubits by Pauli-measurements. Red triangles: Probability of successful state recovery. Blue circles: Probability of successful certification.

6   Numerical examples

We now present some examples which show the performance of certified compressed sensing for randomly chosen states. We demonstrate the method for small-dimensional noiseless states and defer a detailed analysis of the method, especially in the presence of noise and decoherence, to a subsequent publication. For small systems, the condition $c_3 < 1$ in Theorem 1 is hard to satisfy. However, the conditions for uniqueness can be replaced by (a') $c_1 := \|\mathcal{P}_T Y - \operatorname{sgn} \rho\|_2 = 0$ and (b') $c_2 := \|\mathcal{P}_T^\perp Y\| < 1$, discarding the condition on c_3, because these conditions imply that the expression in (11) is positive, which guarantees that any feasible change in the solution will be 1-norm increasing.

Figure 1 demonstrates certified compressed sensing for the very important case of the Pauli basis. It is clearly visible that the certificate is only a sufficient condition and not a necessary one as it is possible that the reconstruction is successful but no valid certificate is produced. It is also apparent that the overhead in the number of queries needed for certification is actually quite reasonable.

For the tight frame consisting of all Hermitian matrices, as shown in Figure 2, it is interesting to note that taking global random observables performs superior to taking tensor products of local random observables. The intuitive reason for this is provided by concentration of measure. By considering a distribution of observables which is invariant under the action of the unitary group on the full system, the proportion of observables that are not Fourier-like, i.e., whose operator norms are too large, is much smaller. Thus, more information is obtained per observable which leads to a faster reconstruction.

Figure 3 illustrates that compressed sensing also works using optical homodyne detection with a generalized tight frame, c.f. Subsection 5.1. In Figure 4, we show the reconstruction of a single mode optical state based on the measurement of expectation values of displacement operators as discussed in Subsection 5.2.
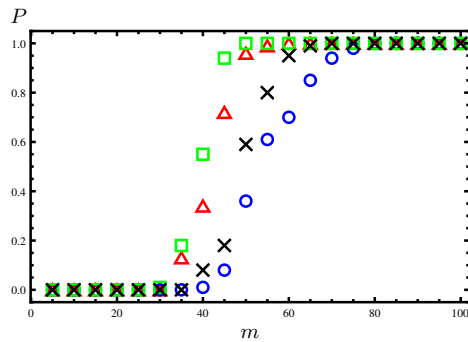
24

Figure 2: (color online) Reconstruction of a pure state on 4 qubits. Red triangles (blue circles): Probability of successful state recovery (certification) for *local* random measurements. Green squares (black crosses): Successful state recovery (certification) for *global* random measurements.
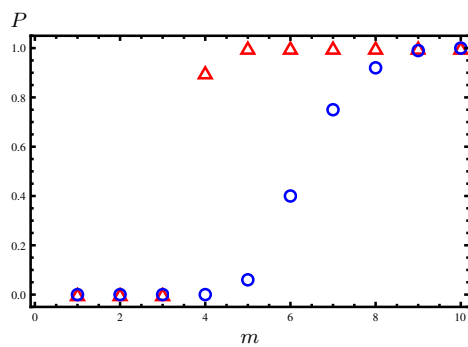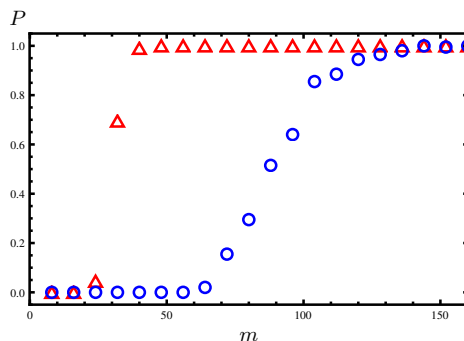

Figure 3: (color online) Reconstruction of a random state with rank 5 on 3 modes with up to 2 photons each by optical Homodyne detection. Red triangles: Probability of successful state recovery. Blue circles: Probability of successful certification.

Figure 4: (color online) Reconstruction of a random state of a single optical mode, truncated at the 15-th number state, by measuring expectation values of 2-norm normalized displacement operators $D(\alpha)$ where $|\alpha\|$ is chosen uniformly at random between 0 and 5 while $\arg(\alpha)$ is chosen uniformly between 0 and 2π. Red triangles: Probability of successful state recovery. Blue circles: Probability of successful certification.

7 Summary

In this article, we have presented a general theory of quantum state tomography for continuous-variable systems using compressed sensing. We have used tight frames to describe continuous measurement families, which are very natural in a plethora of physical situations. We have shown how our theory applies to prominent and frequently used techniques in quantum optics, in particular, pointwise measurements of the Wigner function, and homodyne detection.

- We have explored different incoherence properties sufficient for efficient compressed sensing. Improved results using Fourier-type tight frames were presented in Theorem 2. Also, it was shown in Theorem 3 that for every tight frame whose operators fulfill

$$\|w_\alpha\|_1 = O(\text{polylog}(n)), \tag{98}$$

most states (i.e., all but a proportion $1/\text{poly}(n)$ thereof) can be reconstructed from merely $O(n\,\text{polylog}(n))$ expectation values. It would be interesting to extend these results to generalized tight frames.

- We have introduced the idea of certified compressed sensing which allows to get rid of all assumptions and guarantee successful state reconstruction a posteriori. This assumption-free certified quantum state reconstruction is possible even in the presence of errors.

- Furthermore, we have shown universal compressed sensing results for any Fourier-type tight frame in Theorem 5. This implies strong error bounds in the case of noisy data.

- We have presented numerical results showing the practical (non-asymptotic) performance of these methods. It would be interesting to investigate this further, in

particular to other types of feasible measurements, and to apply these ideas on other physical systems as well.

8 Acknowledgements

References

[1] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Phys. Rev. Lett. **105**, 150401 (2010).

[2] D. Gross, IEEE Trans. on Inf. Th. **57**, 1548 (2011).

[3] A. Shabani, R. L. Kosut, M. Mohseni, H. Rabitz, M. A. Broome, M. P. Almeida, A. Fedrizzi, A. G. White, Phys. Rev. Lett. **106**, 100401 (2011).

[4] M. Cramer, M. B. Plenio, S. T. Flammia, D. Gross, S. D. Bartlett, R. Somma, O. Landon-Cardinal, Y.-K. Liu, and D. Poulin, Nat. Commun. **1**, 149 (2010).

[5] A. Shabani, M. Mohseni, S. Lloyd, R. L. Kosut, and H. Rabitz, Phys. Rev. A **84**, 012107 (2011).

[6] Y.-K. Liu, Adv. in Neural Information Processing Systems **24**, 1638–1646 (2011).

[7] T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hänsel, M. Hennrich, and R. Blatt, Phys. Rev. Lett. **106**, 130506 (2011).

[8] E. J. Candès and B. Recht, Found. of Comput. Math. **9**, 717 (2009).

[9] E. J. Candès and T. Tao, IEEE Trans. on Inf. Th. **56**, 2053 (2010).

[10] S. T. Flammia, D. Gross, Y.-K. Liu, and J. Eisert, arXiv:1205.2300.

[11] S. T. Flammia and Y.-K. Liu, Phys. Rev. Lett. **106**, 230501 (2011).

[12] M. P. da Silva, O. Landon-Cardinal, and D. Poulin, Phys. Rev. Lett. **107**, 210404 (2011).

[13] E. J. Candès and Y. Plan, IEEE Trans. on Inf. Th. **57**, 2342 (2011).

[14] E. J. Candès and Y. Plan, IEEE Trans. on Inf. Th. **57**(11), 7235 (2011).

[15] B. Laurent and P. Messart, Ann. Statist. **28**, 1302 (2000).

[16] M. Émery, M. Ledoux, and M. Yor, Séminaire de Probabilités XXXVIII (2004).

[17] R. Ahlswede and A. Winter, IEEE Trans. on Inf. Th. **48**, 569 (2002).

[18] E. J. Candès and Y. Plan, Proc. IEEE **98**(6), 925 (2010).

[19] R. Bhatia, *Matrix Analysis*, Springer (1997).

[20] M. Rudelson and R. Vershynin, Commun. Pure and Applied Math. **61**, 1025 (2008).

[21] E. J. Candès and T. Tao, IEEE Trans. In. Th. **52**, 5406 (2004).

[22] O. Guédon, S. Mendelson, A. Pajor and N. Tomczak-Jaegermann, Rev. Mat. Iberoamericana **24** (3), 1075 (2008).

[23] B. Recht, M. Fazel and P. A. Parrilo, SIAM Rev. **52**(3), 471 (2010).

[24] P. Bertet, A. Auffeves, P. Maioli, S. Osnaghi, T. Meunier, M. Brune, J. M. Raimond, and S. Haroche, Phys. Rev. Lett. **89**, 200402 (2002).

[25] K. Laiho, L. N. Cassemiro, D. Gross, and C. Silberhorn, Phys. Rev. Lett. **105**, 253603 (2010).

[26] K. Banaszek and K. Wodkiewicz, Phys. Rev. Lett. **76**, 4344 (1996).

[27] S. Wallentowitz and W. Vogel, Phys. Rev. A **53**, 4528 (1996).

[28] W. P. Schleich, *Quantum Optics in Phase Space*, Wiley-VCH (2001).

[29] U. Leonhardt, *Measuring the Quantum State of Light*, Cambridge University Press (1997).

[30] U. Leonhardt and M. G. Raymer, Phys. Rev. Lett. **76**, 1985 (1996).

[31] A. Mari, K. Kieling, B. Melholt Nielsen, E. S. Polzik, and J. Eisert, Phys. Rev. Lett. **106**, 010403 (2011).

[32] M. O. Scully and M. S. Zubairy, *Quantum Optics*, Cambridge Univ. Press, 1997.

[33] M. G. Raymer and M. Beck, Springer Lect. Notes Phys. **649**, 235 (2004).

Appendix

Properties of the χ_k^2-distribution

In order to be self-contained, we repeat two simple bounds to the tails of a χ_k^2 distributed random variable X which can be found in Ref. [15]. A right-sided bound is

$$\mathbb{P}\left(X - k > 2\sqrt{kx} + 2x\right) \leq e^{-x}, \tag{99}$$

while a left-sided one is

$$\mathbb{P}\left(k - X > 2\sqrt{kx}\right) \leq e^{-x}. \tag{100}$$

Random vectors on a sphere

A random vector $v \in \mathbb{C}^n$ on a sphere can be obtained by choosing an vector $\bar{v} \in \mathbb{R}^{2n}$ with Gaussian entries and normalizing. Doing so yields

$$\mathbb{P}\left(|v_i| \geq \frac{\delta}{\sqrt{n}}\right) \leq \mathbb{P}\left(|\bar{v}_i| > \frac{\delta\varepsilon}{\sqrt{n}}\right) + \mathbb{P}\left(\|\bar{v}\| < \varepsilon\right). \tag{101}$$

To bound the first term, one can use (99), obtaining

$$\mathbb{P}\left(|\bar{v}_i| > \frac{1}{\varepsilon\sqrt{n}}\right) \leq \exp\left(-\frac{\delta^2\varepsilon}{2}\right) \tag{102}$$

while for the second terms the inequality (100) leads to

$$\mathbb{P}\left(\|\bar{v}\|^2 < 1 - y\right) < \exp\left(-\frac{y^2 n}{2}\right). \tag{103}$$

Setting $\varepsilon = 1/2$ finally gives

$$\mathbb{P}\left(|v_i| > \delta/\sqrt{n}\right) \leq 2\exp\left(-\frac{\delta^2}{8}\right). \tag{104}$$

Proof of Lemma 8

Proof: From

$$\mathbb{P}\left(\mathbb{P}(x \not\sim y | x \in X) > \beta | y \in Y\right) \leq \frac{p}{\beta} \tag{105}$$

it follows that

$$\mathbb{P}(x \not\sim y | x \in X, y \in Y) \leq p. \tag{106}$$

We assume now the contrary of (105), i.e.,

$$\mathbb{P}\left(\mathbb{P}(x \not\sim y | x \in X) > \beta | y \in Y\right) > \frac{p}{\beta} \tag{107}$$

from which follows

$$\mathbb{P}(x \not\sim y | x \in X, y \in Y) > p, \tag{108}$$

which is a contradiction to (106) and, therefore, concludes the proof.

Truncating the Hilbert space of a continuous-variable-light mode

We show how large the Hilbert space must be to describe a continuous-variable-light mode with bounded energy, i.e., bounded photon number. Let ρ be the state of interest, N_{mean} its mean photon number, and ρ_{trunc} the truncation of ρ to the first N Fock layers which is not normalized

$$\begin{aligned} N_{\text{mean}} &= \sum_{n=0}^{\infty} n\,\rho_{n,n} \geq (N+1) \sum_{n=N+1}^{\infty} \rho_{n,n} \\ &\geq (N+1)\text{Tr}(\rho_{\text{trunc}} - \rho). \end{aligned} \tag{109}$$

From this we obtain

$$\text{Tr}(\rho - \rho_{\text{trunc}}) \leq \frac{N_{\text{mean}}}{N+1}. \tag{110}$$

To get from (110) an error to the 1-norm we need

Lemma 9 (Truncation of matrices)**.** *Let M be a positive semidefinite matrix, or a trace-class operator, written as*

$$M = \begin{pmatrix} A & B \\ B^\dagger & C \end{pmatrix}, \tag{111}$$

where A and C are square matrices. It is true that

$$\|B\|_1^2 \leq \|A\|_1 \|C\|_1. \tag{112}$$

Inserting (112) with $M = \rho$ into (110) and employing the triangle inequality yields with $\|A\|_1 \leq 1$.

$$\|\rho_{\text{trunc}} - \rho\|_1 \leq \frac{N_{\text{mean}}}{N+1} + 2\sqrt{\frac{N_{\text{mean}}}{N+1}} \leq 3\sqrt{\frac{N_{\text{mean}}}{N+1}}, \tag{113}$$

as long as $N + 1 \geq N_{\text{mean}}$.

Proof of Lemma 9

We decompose the Hilbert space according to the block structure of (111) as $E \oplus F$ and write M as $M = \sum_k \lambda M_k$ where the M_k are rank one projectors with A_k, B_k, and C_k as in (111) and $\lambda \geq 0$. Now, we write $\lambda M_k = |\Psi_k\rangle\langle\Psi_k|$ with $|\Psi_k\rangle = a_k|\phi_k\rangle + b_k|\psi_k\rangle$ where $|\phi_k\rangle \in E$ and $|\psi_k\rangle \in F$. From this, one obtains immediately

$$\|B_k\|_1^2 = |a_k|^2 |b_k|^2 = \|A_k\|_1 \|C_k\|_1. \tag{114}$$

To conclude the proof, we write

$$\|B\|_1 \leq \sum_k \|B_k\|_1 \leq \sum_k \sqrt{\|A_k\|_1}\sqrt{\|C_k\|_1}$$
$$\leq \sqrt{\sum_k \|A_k\|_1}\sqrt{\sum_k \|C_k\|_1} = \sqrt{\|A\|_1}\sqrt{\|C\|_1}, \tag{115}$$

where we have used the Cauchy-Schwarz inequality.

Proof of equation (88)

It remains to consider the case where $|\alpha| \geq 2\sigma$. We start by bounding the matrix elements of the displacement operator $D(\alpha)$:

$$\langle k|D(\alpha)|\ell\rangle = e^{-|\alpha|^2/2}\langle k|e^{\alpha a^\dagger}e^{-\alpha^* a}|\ell\rangle, \tag{116}$$

$$e^{-\alpha^* a}|\ell\rangle = \sum_{i=0}^{\ell} \frac{(-\alpha^*)^i}{i!}\sqrt{\ell\cdots(\ell-i+1)}|\ell-i\rangle = \sum_{i=0}^{\ell} \frac{(-\alpha^*)^{\ell-i}}{(\ell-i)!}\sqrt{\ell\cdots(i+1)}|i\rangle, \tag{117}$$

$$\langle k|e^{\alpha a^\dagger} = \sum_{j=0}^{k} \frac{\alpha^j}{j!}\sqrt{k\cdots(k-j+1)}\langle k-j| = \sum_{j=0}^{k} \frac{\alpha^{k-j}}{(k-j)!}\sqrt{k\cdots(j+1)}\langle j|, \tag{118}$$

$$\langle k|D(\alpha)|\ell\rangle = e^{-|\alpha|^2/2}\sum_{j=0}^{\min(k,\ell)} \frac{\alpha^{k-j}}{(k-j)!}\frac{(-\alpha^*)^{\ell-j}}{(\ell-j)!}\sqrt{k\cdots(j+1)}\sqrt{\ell\cdots(j+1)}. \tag{119}$$

Using the Cauchy-Schwarz inequality, and the binomial theorem,

$$|\langle k|D(\alpha)|\ell\rangle| \leq e^{-|\alpha|^2/2}\Big[\sum_{j=0}^{\min(k,\ell)}\Big(\frac{|\alpha|^{k-j}}{(k-j)!}\Big)^2 \cdot k\cdots(j+1)\Big]^{1/2}\Big[\sum_{j=0}^{\min(k,\ell)}\Big(\frac{(|\alpha|)^{\ell-j}}{(\ell-j)!}\Big)^2\cdot\ell\cdots(j+1)\Big]^{1/2}$$

$$= e^{-|\alpha|^2/2}\Big[\sum_{j=0}^{\min(k,\ell)}\binom{k}{j}\frac{|\alpha|^{2(k-j)}}{(k-j)!}\Big]^{1/2}\Big[\sum_{j=0}^{\min(k,\ell)}\binom{\ell}{j}\frac{|\alpha|^{2(\ell-j)}}{(\ell-j)!}\Big]^{1/2}$$

$$\leq e^{-|\alpha|^2/2}\Big[\sum_{j=0}^{k}\binom{k}{j}|\alpha|^{2(k-j)}\Big]^{1/2}\Big[\sum_{j=0}^{\ell}\binom{\ell}{j}|\alpha|^{2(\ell-j)}\Big]^{1/2}$$

$$= e^{-|\alpha|^2/2}(1+|\alpha|^2)^{k/2}(1+|\alpha|^2)^{\ell/2}. \tag{120}$$

Note that, for any fixed k and ℓ, this quantity decays exponentially as $|\alpha|$ becomes large.

We now consider the N-photon truncated operator $D_N(\alpha)$. We can bound it in 2-norm as follows:

$$\|D_N(\alpha)\|_2 \leq e^{-|\alpha|^2/2}\Big[\sum_{k,\ell=0}^{N}(1+|\alpha|^2)^k(1+|\alpha|^2)^\ell\Big]^{1/2}$$

$$= e^{-|\alpha|^2/2}\sum_{k=0}^{N}(1+|\alpha|^2)^k = e^{-|\alpha|^2/2}\frac{(1+|\alpha|^2)^{N+1}-1}{(1+|\alpha|^2)-1} \quad (\text{since } |\alpha| > 0)$$

$$\leq e^{-|\alpha|^2/2}(1+|\alpha|^2)^{N+1}|\alpha|^{-2} = e^{-|\alpha|^2/2}(1+|\alpha|^2)^N(1+|\alpha|^{-2}). \tag{121}$$

Then we can bound the scaled truncated operator $\tilde{D}_N(\alpha)$ as follows:

$$\|\tilde{D}_N(\alpha)\|_2 \leq \sqrt{2}\sigma\exp(\frac{|\alpha|^2}{4\sigma^2}-\frac{|\alpha|^2}{2})(1+|\alpha|^2)^N(1+|\alpha|^{-2})$$

$$= \sqrt{2}\sigma\exp[\frac{|\alpha|^2}{4\sigma^2}-\frac{|\alpha|^2}{2}+N\log(1+|\alpha|^2)](1+|\alpha|^{-2}). \tag{122}$$

Let

$$E := \frac{|\alpha|^2}{4\sigma^2}-\frac{|\alpha|^2}{2}+N\log(1+|\alpha|^2) \tag{123}$$

be the quantity inside the exponential; we will upper-bound it. Note the following identity, for any $x, x_0 \in (0,\infty)$: (by approximating $\log(1+x)$ to first order at the point $x = x_0$)

$$\log(1+x) \leq \log(1+x_0) + \frac{x-x_0}{1+x_0} = \log(1+x_0) + \frac{1+x}{1+x_0} - 1. \tag{124}$$

Set $x = |\alpha|^2$ and $x_0 = 4N$, then we have

$$\log(1+|\alpha|^2) \leq \log(1+4N) + \frac{1+|\alpha|^2}{1+4N} - 1 \leq \log(1+4N) + \frac{1+|\alpha|^2}{4N} - 1 \leq \log(1+4N) + \frac{|\alpha|^2}{4N}. \tag{125}$$

Then

$$E \leq (\frac{1}{4\sigma^2}-\frac{1}{2}+\frac{1}{4})|\alpha|^2 + n\log(1+4N). \tag{126}$$

Using the fact that $\alpha \geq 2\sigma = \sqrt{8N\log(1+4N)}$, we get

$$E \leq (\frac{1}{4\sigma^2}-\frac{1}{2}+\frac{1}{4}+\frac{1}{8})|\alpha|^2 = (-\frac{1}{8}+\frac{1}{4\sigma^2})|\alpha|^2. \tag{127}$$

31

Plugging this into (122), we get

$$\|\tilde{D}_N(\alpha)\|_2 \leq \sqrt{2}\sigma \exp[(-\tfrac{1}{8} + \tfrac{1}{4\sigma^2})|\alpha|^2](1 + |\alpha|^{-2}). \qquad (128)$$

Using the fact that $\sigma \geq 2$ and $|\alpha| \geq 2\sigma \geq 4$, we have that

$$\|\tilde{D}_N(\alpha)\|_2 \leq \sqrt{2}\sigma \exp[-\tfrac{1}{16}|\alpha|^2](1 + |\alpha|^{-2}) \leq \sqrt{2}\sigma \exp(-1)\tfrac{17}{16} < \sqrt{2}\sigma. \qquad (129)$$

Since the operator norm is upper-bounded by the 2-norm, this implies (88).

Chapter 6

Efficient and feasible state tomography of quantum many-body systems

Efficient and feasible state tomography of quantum many-body systems

M. Ohliger

Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany and
Institute for Physics and Astronomy, University of Potsdam, 14476 Potsdam, Germany

V. Nesme and J. Eisert

Laboratoire d'Informatique de Grenoble, 38400 Saint-Martin-d'Heres, France
(Dated: June 21, 2012)

We present a novel method to perform quantum state tomography for many-particle systems which are particularly suitable for estimating states in lattice systems such as of ultra-cold atoms in optical lattices. We show that the need for measuring a tomographically complete set of observables can be overcome by letting the state evolve under some suitably chosen random circuits followed by the measurement of a single observable. We generalize known results about the approximation of unitary 2-designs, i.e., certain classes of random unitary matrices, by random quantum circuits and connect our findings to the theory of quantum compressed sensing. We show that for ultra-cold atoms in optical lattices established techniques like optical super-lattices, laser speckles, and time-of-flight measurements are sufficient to perform fully certified, assumption-free tomography. Combining our approach with tensor network methods – in particular the theory of matrix-product states – we identify situations where the effort of reconstruction is even constant in the number of lattice sites, allowing in principle to perform tomography on large-scale systems readily available in present experiments.

I. INTRODUCTION

Quantum state tomography is – for obvious reasons – a procedure of great importance in a large number of experiments involving quantum systems: It amounts to reconstructing an unknown quantum state entirely based on experimental data. In many situations one indeed aims at identifying what state has actually been prepared in an experiment. This seems particularly important in the context of quantum information science, where quantum state and process tomography is now routinely applied to small, precisely controlled quantum systems [1–3]. Yet, needless to say, in a number of other contexts the reliable reconstruction of quantum states is an important aim as well.

For finite-dimensional quantum systems, conventional quantum state tomography can be performed by choosing a suitable basis of $\mathcal{B}(\mathbb{C}^d)$, i.e., the operators on the d-dimensional Hilbert space of the system in question. Then, the expectation values of these d^2 observables are being measured to some required accuracy, from which one can reconstruct the unknown density matrix ρ. The same approach, however, is doomed to failure when applied to quantum many-body systems: If one has a many-body system at hand with k lattice sites of local dimension d_l, the number of necessary different measurement settings is given by $m = d_l^{2k}$, i.e., it scales exponentially with the size, rendering the treatment even of reasonably large systems impossible. Techniques of quantum compressed sensing [4–8] allow to significantly reduce the required number of measurement settings, if the state is of rank r, to $m = \Theta(rd\log^2 d)$ (where Θ denotes asymptotic equality). If $r \ll d$, this is an impressive reduction, and gives rise to feasible quantum state tomography in medium-sized quantum systems, but this number is still exponential in the number of sites. Such a scaling cannot be overcome without further restriction of the class of possible states, simply because even a pure state needs of the order of d parameters to be described. However, if the state is not only pure but also described by a
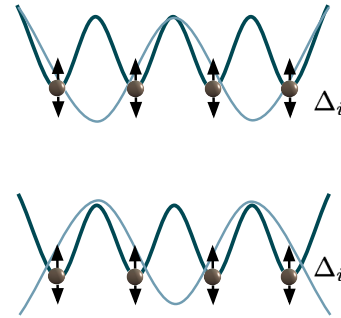

FIG. 1. Realization of a random quantum circuit by an optical superlattice. The lattice with the larger period is switched between the two depicted situations, and the lattice depth is changed locally by Δ_i which is uncorrelated between the lattice sites.

generic matrix product state (MPS), the necessary number of measurements only scales linearly with the system size and is even independent of it for the important special case of translationally invariant MPS. What is more, in several instances the classical procedure to reconstruct the MPS matrices from the measurement data is efficient [9].

This small number of parameters ought to make tomography an easier task, but in many practical settings involving quantum many-body systems, a serious challenge arises: In most interesting systems it is very difficult if not impossible to directly measure a full operator basis. Instead, merely measurements of some preferred observables might be readily available.

In the present work, we propose a solution to this problem. We do so by combining the action of a suitable random quantum circuit with a measurement of a very small number of different observables (even a single one can be enough). Such random circuits are just becoming a tool of great theoretical

importance in several subfields of quantum information theory [10–12]. Here, we show that they also offer technological advantages, in that they allow for the natural implementation of quantum state tomography in systems of ultra-cold atoms using only the techniques of super-lattices and laser speckles. Both techniques have already been experimentally proven to be feasible [13–15]. Combining these new insights with the above mentioned methods of MPS tomography brings, for the first time, full tomography of many-particle systems close to experimental reach. Needless to say, it still amounts to a very difficult prescription. But while conventional measurements in ultra-cold bosonic systems amount to measuring certain correlation functions or estimates of the temperature in thermometry, say, the path described here eventually allows for the full reconstruction of an unknown state of a quantum many-body system – quite a promising perspective.

The remainder of this article is organized as follows: First, we introduce in Section II tomographically complete sets of observables and a generalization of operator bases, called tight frames, and discuss their realization by means of random unitary matrices. In a next step, we show in Section III how efficient compressed sensing is possible with families of unitary matrices which form approximate unitary 2-designs [10–12, 16–18]. In Section IV, a way to efficiently realize such unitary 2-designs with the help of random quantum circuits is presented, before we discuss the application of this approach to ultra-cold atoms in optical lattices in Section V. Finally, Section VI shows how, under certain assumptions on the state, tomography is possible with a number of measurements which is linear or constant in the number of lattice sites before a conclusion is given in Section VII.

II.   TOMOGRAPHY BY MEANS OF UNITARY EVOLUTION

Before we turn to the question of how tomography is possible, we provide general conditions for sets of observables to be suitable for reconstructing quantum states. We discuss why tomographic completeness as such is not sufficient and introduce generalizations of operator bases which allow for robust tomography.

A.   Tomographically complete sets of observables and tight frames

We consider a quantum system consisting of k subsystems, called (lattice) sites which all have local dimension d_l. Let $\rho \in \mathcal{S}(\mathcal{H})$ be a quantum state on the Hilbert space $\mathcal{H} = \mathbb{C}^d$ with dimension $d = d_l^k$ and S the set of corresponding normalized observables, i.e., Hermitian matrices $w \in \mathcal{B}(\mathcal{H})$ with $\|w\|_2 = 1$. In this whole work, we denote by $\| \cdot \|_p$ the Schatten p-norm, where $p = 2$ and $p = \infty$ are the Frobenius norm and operator norm respectively. Furthermore, we use the Hilbert-Schmidt scalar product defined as $(A, B) = \mathrm{Tr}(A^\dagger B)$ and the projection on the subspace spanned by some $w \in S$ which is

defined as

$$\mathcal{P}_w : \rho \mapsto (w, \rho)w. \tag{1}$$

Noting that a measurement of the expectation value of w corresponds to determining (w, ρ), we define the total sampling operator

$$\mathcal{W}_d = d^2 \int \mathrm{d}\mu(w)\mathcal{P}_w \in \mathcal{B}(\mathcal{B}(\mathcal{H})), \tag{2}$$

where μ is a probability measure on S.

A finite set of observables is said to be tomographically complete if any two different quantum states have distinct expectation values for some observable: This implies that one can theoretically reconstruct the state if one knows all expectation values. If $\mathcal{W}_d$ has full rank, i.e. $\mathrm{rank}(\mathcal{W}_d) = d^2$, the state ρ can be obtained from $\mathcal{W}_d(\rho)$ by matrix inversion if issues of statistical errors and numerical imprecisions are neglected. For general probability measures we make use of the subsequent definition.

Definition 1 (Tomographic completeness). *The measure μ is said to be tomographically complete if $\mathcal{W}_d$, as defined in Eq. (2), is full rank.*

In practice, not every tomographically complete measure on the observables is necessarily useful: The observables, viewed as vectors in $\mathbb{R}^{d^2}$, should not be too unevenly distributed over the sphere. That is to say, if the ratio between the largest and smallest eigenvalues of $\mathcal{W}_d$ is large, small errors in the expectation values can lead to large errors in the reconstructed state. The ideal situation is that of a tight frame, also known as spherical 1-design:

Definition 2 (Tight frame). *A probability measure μ on the set of 2-norm normalized Hermitian matrices S is called a tight frame if $\mathcal{W}_d = \mathbb{1}$ with $\mathcal{W}_d$ given by Eq. (2).*

Examples for tight frames include any operator basis and the rotationally invariant measure on the 2-norm sphere S [8]. When observables are taken from tight frames, the reconstruction problem is well conditioned and small errors in the expectation values only lead to a small error in the reconstructed state [8, 19].

B.   Evolution of observables

A tomographically complete set of observables must contain at least d^2 observables, which might be difficult to measure directly. We introduce a way to do tomography by performing suitably random unitaries, followed by the estimation of the expectation value for a single observable. A different but related approach has been employed in Ref. [20] to perform high-fidelity quantum state reconstruction in situations where the knowledge about the state is not tomographically complete. In the present work, the time evolution is a tool to obtain knowledge about the quantum state of the system leaving the question of determining Hamiltonians aside [21].

Switching to the Heisenberg picture, the outlined procedure amounts to measuring the time-evolved observables. Simple as this idea is, it allows the economical reconstruction of unknown quantum states, as it turns out:

Definition 3 (Induced observables). *To a measure $\tilde{\mu}$ on the special unitary group $SU(d)$ and an observable w_0 we associate the following induced measure on S:*

$$\mu = \left(1 - \frac{1}{d^2}\right)\tilde{\mu} \circ f^{-1} + \frac{1}{d^2}\delta_{\mathbb{1}/\sqrt{d}} \qquad (3)$$

where $f : U(d) \to S$ is defined by $f(U) = U^\dagger w_0 U$, and δ_x denotes the Dirac in x. Notice that if $\tilde{\mu}$ is a probability measure, then so is μ.

An important insight is provided by the following observation.

Observation 1 (Tight frame induced by random unitary). *Let w be a traceless, normalized observable and μ_H be the Haar measure on $SU(d)$. The measure induced on S is a tight frame.*

Proof: To show this observation, we use (3) to calculate the sampling operator (2) for a state ρ with $\mathrm{Tr}(\rho) = 1$:

$$\mathcal{W}_{d,\mathrm{H}}(\rho) = (d^2 - 1)\int \mathrm{d}\mu_\mathrm{H}(U)(U^\dagger w_0 U, \rho)U^\dagger w_0 U + \frac{\mathbb{1}}{d}. \tag{4}$$

As w is diagonalized by a unitary matrix, we can assume it to be diagonal and obtain

$$(\mathcal{W}_{d,\mathrm{H}}(\rho))_{i,j} = (d^2 - 1)\sum_{k,l,m,n} \mathbb{E}_{U\sim\mu_\mathrm{H}}\left[U_{k,l}U_{m,j}\bar{U}_{k,n}\bar{U}_{m,i}\right]$$
$$\times w_{k,k}w_{m,m}\rho_{l,n} + \frac{\delta_{i,j}}{d} \tag{5}$$

where $U \sim \mu_\mathrm{H}$ indicates that U is distributed according to the Haar measure. The occurring expectation values can be obtained from Ref. [22]. We first consider the off-diagonal elements, i.e., the ones with $i \neq j$. Here, the expectation value vanishes unless $l = i$ and $n = j$, in which case we get

$$\mathbb{E}_{U\sim\mu_\mathrm{H}}\left[U_{k,i}U_{m,j}\bar{U}_{k,j}\bar{U}_{m,i}\right] =$$
$$\frac{1}{d(d^2-1)}\left(\delta_{k,m}(n-1) + (1-\delta_{k,m})(-1)\right). \tag{6}$$

We now turn to the diagonal elements and note that if $i = j$, one needs $l = n$ to get a non-vanishing expectation value. We consider two cases separately: If $m = k$ we get

$$\mathbb{E}_{U\sim\mu_\mathrm{H}}\left[U_{k,l}U_{k,i}\bar{U}_{k,l}\bar{U}_{k,i}\right] =$$
$$\frac{1}{d(d^2-1)}\left(\delta_{i,l}2(n-1) + (1-\delta_{i,l})(n-1)\right), \tag{7}$$

while for $m \neq k$ we obtain

$$\mathbb{E}_{U\sim\mu_\mathrm{H}}\left[U_{k,l}U_{m,i}\bar{U}_{k,l}\bar{U}_{m,i}\right] =$$
$$\frac{1}{d(d^2-1)}\left(\delta_{i,l}(n-1) + (1-\delta_{i,l})n\right). \tag{8}$$

Inserting now (6), (7), and (8) into (4) and using $\sum_i \rho_{i,i} = 1$, $\sum_i w_{i,i} = 0$, and $\sum_i |w_{i,i}|^2 = 1$, we get $\mathcal{W}_{d,\mathrm{H}}(\rho) = \rho$ which concludes the proof.

C. Tight frames under physical restrictions

In many situations of interest, the quantum state is not completely arbitrary but satisfies some additional conditions. In experiments with ultra-cold atoms, for example, the total particle number is conserved and super-selection rules forbid superpositions of states belonging to different eigenvalues of the corresponding operator. Assume that the quantum state ρ acts only on the eigenspace of some operator $\hat{N}$ corresponding to the eigenvalue N and denote the corresponding matrix valued projection operator as $\mathcal{P}_N$. In this case, Definition 2 can be relaxed to

$$\mathcal{W}_{d_N}\mathcal{P}_N = \mathcal{P}_N \tag{9}$$

where d_N denotes the dimension of the subspace corresponding to the eigenvalue N of $\hat{N}$. If one aims at realizing such a restricted tight frame by the means of Observation 1, one can replace the group $SU(d)$ by

$$SU_{\hat{N}}(d) = \left\{U \in SU(d) : [U, \hat{N}] = 0\right\}. \tag{10}$$

As this group is compact, there exists a unique Haar probability measure on it. We adapt Definition 3 to this situation.

Definition 4 (Induced observable on subspace). *Let $\hat{N} \in \mathcal{B}(\mathcal{H})$ and S_N its d_N dimensional eigenspace corresponding to the eigenvalue N. To a measure $\tilde{\mu}$ on the group $SU_{\hat{N}}(d)$ and an observable $w_0 \in S_N$ with $\|w_0\|_2 = 1$, we associate the following measure on S_N:*

$$\mu = \left(1 - \frac{1}{d_N^2}\right)\tilde{\mu} \circ f^{-1} + \frac{1}{d_N^2}\delta_{\mathbb{1}/\sqrt{d_N}} \tag{11}$$

where $f(U) = U^\dagger w_0 U$.

The matrices ρ, w_0, and all $V \in SU_{\hat{N}}(d)$ are block diagonal with respect to the eigenbasis of $\hat{N}$. We consider the block corresponding to the eigenvalue N. Invariance of the Haar measure on $SU_{\hat{N}}$ implies that if $V \sim \mu_{\mathrm{H},SU_{\hat{N}(d)}}$, then $\mathcal{P}_N(V) \sim \mu_{\mathrm{H},SU(d_N)}$. Thus, one can apply the proof of Observation 1 to this block and obtain Eq. (9).

III. UNITARY t-DESIGNS

The effort to implement random unitaries drawn from the Haar measure scales exponentially in the number of lattice sites k, making an implementation both theoretically inefficient and pratically unfeasible. However, this problem can be circumvented by replacing the Haar measure by a unitary 2-design which is much easier to sample from as we will see later on. Unitary t-designs behave like the Haar measure in specific situations [12, 17, 18]. The definition most suited to our problem is the following: Let ν be a probability measure on $SU(d)$. We define two channels on $\mathcal{B}(\mathcal{H}^{\otimes t})$:

$$\mathcal{G}_{t,\nu}(\rho) = \mathbb{E}_{U\sim\nu}\left[U^{\otimes t}\rho(U^\dagger)^{\otimes t}\right] \tag{12}$$

and $\mathcal{G}_{t,\mu_{\mathrm{H}}}$, where μ_{H} is the Haar measure. We say that ν is a unitary t-design if $\mathcal{G}_\nu = \mathcal{G}_{\mathrm{H}}$. We say that ν is an ε-approximate t-design if

$$\|\mathcal{G}_{t,\nu} - \mathcal{G}_{t,\mathrm{H}}\| \le \varepsilon \tag{13}$$

where $\|\cdot\|$ denotes the superoperator $2 \to 2$-norm which is defined as

$$\|\mathcal{O}\| = \sup_{X,\|X\|_2=1} \|\mathcal{O}(X)\|_2. \tag{14}$$

This superoperator norm is equal to the Schatten ∞-norm when the channel is seen as a mere linear operator acting on the real vector space of Hermitian matrices. In the remainder we only consider the case $t = 2$ and drop the index t for simplicity.

A. Tight frames from unitary 2-designs

For our purpose, i.e., replacing the Haar measure random unitary in Observation 1, we need an approximate unitary 2-design. This is the case because when Haar measure induces a tight frame, c.f. Eq. (4), both U and $U^\dagger$ appear twice.

Observation 2 (Tight frames induced by unitary 2-designs). *Let w_0 be a traceless, 2-norm normalized observable and ν be an ε-approximate unitary 2-design. Then the sampling operator $\mathcal{W}_{d,\nu}$ corresponding to the induced measure fulfills*

$$\|\mathcal{W}_{d,\nu} - \mathbb{1}\| \le \sqrt{d}(d^2 - 1)\varepsilon. \tag{15}$$

Proof: We have

$$\begin{aligned}\|\mathcal{W}_{d,\nu} - \mathbb{1}\| &= \|\mathcal{W}_{d,\nu} - \mathcal{W}_{d,\mathrm{H}}\| \\ &= \sup_{X,\|X\|_2=1} \|\mathcal{W}_{d,\nu}(X) - \mathcal{W}_{d,\mathrm{H}}(X)\|_2. \end{aligned} \tag{16}$$

We note that

$$\begin{aligned}\mathcal{W}_{d,\nu}(X) &= (d^2-1)\mathbb{E}_{U\sim\nu}\left[(U^\dagger w_0 U, X)U^\dagger w_0 U\right] + \frac{1}{d}\mathrm{Tr}X \\ &= (d^2-1)\sqrt{d}\,\mathrm{Tr}_1\Bigg(\mathbb{E}_{U\sim\nu}\left[(U^\dagger \otimes U^\dagger)\right. \\ &\quad \times (w_0 \otimes w_0)(U \otimes U)\Big]\left(X \otimes \frac{\mathbb{1}}{\sqrt{d}}\right)\Bigg) + \frac{1}{d}\mathrm{Tr}X, \end{aligned} \tag{17}$$

where Tr_1 denotes the partial trace with respect to the first of the two subsystems of equal dimension. Note that we have extended the definition of $\mathcal{W}_{d,\nu}$ to operators with non-unit trace. This relation yields (15) after inserting it into Eq. (16) and applying (13).

The same argument holds also for restricted tight frames as defined in (9): Let ν be a distribution on $SU_{\hat{N}}(d)$ such that if $U \sim \nu$, then $\mathcal{P}_N(U)$ is drawn from an ε-approximate 2-design on $SU(d_N)$. This implies

$$\|\mathcal{W}_{d_N,\nu}\mathcal{P}_N - \mathcal{P}_N\| \le \sqrt{d_N}(d_N^2 - 1)\varepsilon, \tag{18}$$

as follows from the application of the above proof to the block in U corresponding to the eigenvalue N of $\hat{N}$.

B. Compressed sensing

The technique of compressed sensing allows to reduce the number of measurements which are needed to reconstruct a quantum state from $\Theta(d^2)$ to $\Theta(d\,\mathrm{polylog}(d))$ if the rank of the state does not increase with d. To perform this method, one has to choose $m = \Theta(d\,\mathrm{polylog}(d))$ observables $w_1, \ldots, w_m$ randomly from the tight frame according to the corresponding probability measure and determine their expectation value (w_i, ρ) by measurement. Then, one can efficiently solve the optimization problem

$$\min_\sigma \|\sigma\|_1 \text{ s.t. } \forall i = 1, \ldots, m : (w_i, \sigma) = (w_i, \rho). \tag{19}$$

The theory has been developed for observables forming operator bases in Refs. [5, 6] and extended to tight frames in Ref. [8]. There, it was also shown that the tight frame condition may be violated and compressed sensing is still possible if

$$\|\mathcal{W}_d - \mathbb{1}\| \le \frac{1}{8\sqrt{r}}, \tag{20}$$

where r is the rank of the state we want to reconstruct. Not all tight frames are equally suited for compressed sensing as can be seen with a simple example: Let $w_1, \ldots, w_{d^2}$ be the elements of an orthonomal operator bases of $\mathcal{B}(\mathbb{C}^d)$ where w_1 is a rank-one projector and $\rho = w_1$. In this case, one has to measure of the order of d^2 observables before one "hits" w_1 and gets any information on the system. In Refs. [5, 6, 8], it has been shown that this problem cannot occur if all observables fulfill the so-called "Fourier type incoherence condition" which reads

$$\mathbb{P}_{w\sim\mu}\left(\|w\|_\infty^2 > \frac{\lambda}{d}\right) = 0, \tag{21}$$

where λ must fulfill $\lambda = O(\mathrm{polylog}(d))$. Note that statements like (21) make only sense when considering families of tight frames with growing dimension d. As we are mainly interested in the asymptotic efficiency of our scheme, we restrict ourselves to the scaling behavior and omit explicit prefactors. We now give a condition under which (21) is fulfilled in the situation of interest.

Observation 3 (Compressed sensing with induced observables). *Let w_0 be a traceless, normalized observable fulfilling $\|w_0\|_\infty^2 \le \lambda/d$ with $\lambda = O(\mathrm{polylog}(d))$, and let ν be a $1/(8\sqrt{rd}(d^2-1))$-approximate 2-design. The induced tight frame fulfills (21), which implies that it allows for compressed sensing.*

Proof: Since

$$\|U^\dagger w_0 U\|_\infty^2 = \|w_0\|_\infty^2 \tag{22}$$

and $\|\mathbb{1}/\sqrt{d}\|_\infty^2 = 1/d$, condition (21) is fulfilled and from (15) it follows that (20) is satisfied which proves that compressed sensing is possible.

Observation 3 holds especially in the important case of a observable which acts non-trivially only on a few number of

lattice sites because here it is of the form

$$w_0 = v \otimes \frac{\mathbb{1}_{d_l^{k-m}}}{(d_l^{k-m})^{1/2}} \qquad (23)$$

for some normalized, traceless observable v and some small constant m.

IV. APPROXIMATION OF UNITARY 2-DESIGNS BY RANDOM QUANTUM CIRCUITS

In the previous section, we have shown that unitary 2-designs can be used to realize tight frames. We now show how they can be approximated by parallel random circuits and generalize the results of Refs. [10, 11, 16] to show the following:

Observation 4 (Random circuits). *Assume k to be even. Consider a parallel random circuit where in each step either $U_{1,2} \otimes U_{3,4} \otimes \ldots \otimes U_{k-1,k}$ or $U_{2,3} \otimes U_{4,5} \otimes \ldots \otimes U_{k-2,k-1}$ is performed with probability $1/2$ where $U_{i,i+1}$ acts on the neighboring sites i and $i + 1$. If the nearest-neighbor unitaries are drawn, in each step independently, from a probability measure ν_2 which is universal, as defined below, there exists a constant C (depending on the local dimension) such that the random circuit forms an ε-approximate unitary 2-design after $n = C \log(1/\varepsilon)k \log k$ steps.*

A finite set of nearest-neighbor unitary quantum gates is called universal if they generate a dense subgroup of $SU(d_l^2)$. For an arbitrary probability measure on $SU(d_l^2)$, the notion of universality can be generalized, according to Ref. [16]:

Definition 5 (Universality). *We say that μ is universal if for any open ball S there exists $l > 0$ such that S has a nonzero weight for the l-fold convolution product of μ.*

Proof: The proof of Observation 4 is an extension of that of similar results in Refs. [10, 11]. Readers mostly interested in the application to optical lattice systems can safely skip to the next section.

In Ref. [11], it is shown that parallel random circuits with periodic boundary conditions form ε-approximate 2-designs after $n = C \log(1/\varepsilon)k$ steps if the unitaries are drawn from the Haar measure on $SU(d_l^2)$.

We now proceed in three steps: First, we show that the nearest-neighbor unitaries can be drawn from an approximate 2-design on $SU(d_l^2)$ instead from the Haar measure. We denote the measure corresponding to a single step of the random circuit by ν_k and define the linear operator G_{ν_k} by

$$G_{\nu_k} = \int d\nu_k(U) \, U^{\otimes 2} \otimes \bar{U}^{\otimes 2}. \qquad (24)$$

This operator can be decomposed as $G_{\nu_k} = (M_e + M_o)/2$ with

$$M_e = P_{1,2} \otimes P_{3,4} \otimes \ldots \otimes P_{k-1,k}, \qquad (25)$$
$$M_o = P_{k,1} \otimes P_{2,3} \otimes \ldots \otimes P_{k-2,k-1} \qquad (26)$$

where

$$P_{i,j} = \int d\nu_2(U_{i,j}) \, U_{i,j}^{\otimes 2} \otimes \bar{U}_{i,j}^{\otimes 2}. \qquad (27)$$

We have to bound

$$\|G_{\nu_k}^n - G_{\mathrm{H}}\|_\infty = \lambda_2(G_{\nu_k})^n \qquad (28)$$

where n is the depth of the circuit and λ_2 denotes the second largest eigenvalue. In Ref. [11], it is shown that if the nearest-neighbor unitaries are drawn from the Haar measure, there exists a constant $\Lambda > 0$ such that the corresponding operator $\tilde{G}_{\nu_k}$ fulfills $\lambda_2(\tilde{G}_{\nu_k}) \leq 1 - \Lambda$. Using now the fact, that ν_2 is a δ-approximate 2-design, we get

$$\lambda_2(G_{\nu_k}) \leq \lambda_2(\tilde{G}_{\nu_k}) + k\|G_{\nu_2} - G_{\mathrm{H}_2}\|_\infty \leq 1 + k\delta - \Lambda. \qquad (29)$$

For the right-hand side to be smaller than one, which is necessary and sufficient for an exponentially fast convergence, one has to choose $\delta = O(1/k)$. To realize this *local* approximate 2-design, we use a result from Ref. [16] which states that one needs to draw only $s = O(\log(1/\delta))$ gates from an arbitrary universal gate set to achieve this. Thus, we have a circuit with depth

$$n_\varepsilon = O(ns) = O(\log(1/\varepsilon)k \log k) \qquad (30)$$

where the random choice between M_e and M_o is not made in every step but in blocks of s steps which corresponds to the operator $(M_e^s + M_o^s)/2$ while s steps of the actual quantum circuit performed are described by $((M_e + M_o)/2)^s$. As they have the same fixed point and

$$\lambda_2\left(\left(\frac{M_e + M_o}{2}\right)^s\right) \leq \lambda_2\left(\frac{M_e^s + M_o^s}{2}\right), \qquad (31)$$

the convergence of the actual circuit cannot be slower and (30) holds. The last thing needed to obtain Observation 4 is to switch to open boundary conditions, i.e., remove the first tensor factor in Eq. (26). As this does not affect the fixed point and the operator norm difference is on the order of δ, only the prefactor is changed slightly. We note that the prefactor depends on the actual choice of ν_2.

The conditions for Observation 3 to apply are fulfilled if $1/\varepsilon = O(d^{5/2})$. Using this in Eq. (30), we get

$$n_{\mathrm{T}} = O(k^2 \log k). \qquad (32)$$

This means compressed sensing is possible with a single traceless observable and a parallel random quantum circuit with a depth given by Eq. (32).

We did not explicitly discuss the case of restricted tight frames because it can be, again treated by block-decomposing all matrices according to the spectral decomposition of the operator $\hat{N}$ describing the symmetry, c.f. Eq. (10). Thus, with a parallel random quantum circuit as in Observation 4 with unitaries which are universal for $SU_{\hat{N}}(d_l^2)$ with a depth as in Eq. (32), one can perform compressed sensing for states within some eigenspace of $\hat{N}$.

V. OPTICAL LATTICE SYSTEMS

Ultra-cold atoms in optical lattices form some of the cleanest quantum many-particle systems available for experiments and allow for the realization of well-known effects from condensed matter. For example, both the bosonic super-fluid Mott-insulator transition and the Mott state of fermions where observed by changing the intensity of the laser forming the lattice [23, 24]. Such systems also have the potential of functioning as quantum simulators which means, they allow to simulate systems from other branches of physics like the notoriously difficult quantum chromodynamics (QCD) [25].

A. Time-of-flight measurements

Even though measurements with spatial resolution have been demonstrated in recent experiments [15], the standard technique is still provided by time-of-flight absorption imaging. Here, the lattice and the confining trap are instantaneously switched off and, after some time during which the atoms expand approximately without interaction, an absorption image is taken [23, 27]. As the distance that the atoms fly during the expansion is proportional to the initial momentum, this procedure amounts to a measurement of the density in momentum space. We restrict ourselves to the bosonic case, while noting that fermions can be treated in a completely analogous way, and expand the field operators of the one-dimensional bosonic field as

$$\hat{\Psi}(x) = \sum_{j=1}^{\infty} \sum_{s=1}^{k} W^{(j)}(x - x_s)\hat{b}_s^{(j)} \qquad (33)$$

where $W^{(j)}$ is the Wannier function of the j-th band, x_s is the position, and $\hat{b}_s^{(j)}$ the corresponding annihilation operator at site s. If the lattice is sufficiently deep, all bands but the lowest one can be neglected, and we drop the upper index in (33). The momentum-space distribution, which is measured in the time-of-flight experiment, is given by $n(p) = |\tilde{W}(p)|^2 S(p)$ where $\tilde{W}$ is the Fourier transform of the Wannier function and the quasi-momentum distribution is given by

$$S(p) = \sum_{s,l=1}^{k} e^{ip(x_s - x_l)} \langle \hat{b}_s^\dagger \hat{b}_l \rangle. \qquad (34)$$

As we do not assume the state to be translationally invariant, Eq. (34) cannot be inverted to get the two-point correlation functions in real space but we can get for integer l

$$\sum_{s=1}^{k} \langle \hat{b}_s^\dagger \hat{b}_{s+l} \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} dp \, e^{ipl} S(p), \qquad (35)$$

where we have set the lattice spacing to one. If the atoms are bosons, the local Hilbert space is infinite-dimensional. However, as the interaction between the atoms must always be repulsive to ensure stability of the quantum gas, one can neglect state where more than a given cut-off number of atoms are

present on a single lattice site. This allows us to work with a finite-dimensional Hilbert space. We note that this does not even needs to be an approximation as one can set the maximal number of bosons per site N_S to their total number N. However, in any practical setting, one would use $N_S \ll N$ and still get a very good approximation. For an arbitrary i, we set

$$w_0^{(i)} \propto \sum_{j=1}^{k} (\hat{b}_i^\dagger \hat{b}_{i+j} + \hat{b}_{i+j}^\dagger \hat{b}_j) \qquad (36)$$

which is traceless. Due to the sum, which stems from the absence of translational invariance, Eq. (36) is not exactly of the form given by (23) but a sum of few, i.e. logarithmically many in the Hilbert space dimension d, terms of this form. This implies

$$\|w_0^{(i)}\|_\infty \leq C \sum_{j=1}^{k} \|\hat{b}_j^\dagger \hat{b}_{i+j} + \hat{b}_{i+j}^\dagger \hat{b}_j\|_\infty \leq \tilde{C}k \qquad (37)$$

where C and $\tilde{C}$ are constants. Because $k = \Theta(\log d)$, we can employ Observation 1 to show that a measurement of the momentum space distribution, together with an approximate 2-design, allows for efficient compressed sensing.

Although already a single choice of i yields an approximate tight frame, we can use the data corresponding to all $i = 1, \ldots, k$, as they are measured anyway, to reduce the necessary number of experiments.

B. Realization of the random circuit

We now discuss how a probability measure on the nearest-neighbor unitaries that is universal can be obtained. To be as specific and simple as possible, we use the single-band Bose Hubbard model with Hamiltonian [28]

$$\hat{H} = -\sum_{i=1}^{k} J_i(\hat{b}_i^\dagger \hat{b}_{i+1} + \hat{b}_{i+1}^\dagger \hat{b}_i) + \frac{U_i}{2}\hat{n}_i(\hat{n}_i - 1) + \Delta_i \hat{n}_i, \quad (38)$$

where $\hat{n}_i = \hat{b}_i^\dagger \hat{b}_i$ and $J_1 \ldots, J_k; \Delta_1, \ldots, \Delta_k; U_1, \ldots, U_k \in \mathbb{R}$. To realize the parallel random quantum circuit, we make use of the techniques of super-lattices [14, 15, 26] and speckle patterns [13]. As the total number of atoms is conserved and super-selection rules forbid the superposition of states with different particle numbers, we restrict ourselves to particle-number conserving operations.

By using an additional lattice for which the lattice constant is twice as large, one can change the height of the wells between alternating pairs of site. This mainly affects the hopping constants J_i and to much less extend the interaction parameter U_i, an effect which we neglect. By choosing the depth of the super-lattice large enough, we get $J_i = 0$ for the non-coupled pairs and $J_i = J$ for the coupled ones. Such a double-well structure has been used in Ref. [15] to probe correlation functions. A speckle pattern is created by illuminating an uneven surface with a laser and can be modelled by a spatially fluctuating Δ_i. The situation is sketched in Figure 1. For reasons of

simplicity, we only consider the regime for which the strength of the speckle potential is not correlated over the lattice sites, i.e., all Δ_i are independently distributed. Thus, we have a random circuit as in Observation 4 and we only need to show that the corresponding gates generate a dense set in $SU_N(d_l^2)$. The local gates are

$$
\begin{aligned}
U_{i,j}(\Delta_1, \Delta_2, t) = \exp\big(-it(-J(\hat{b}_1^\dagger \hat{b}_2 + \hat{b}_2^\dagger \hat{b}_1) \\
+ \frac{U}{2}(\hat{n}_1(\hat{n}_1 - 1) + \hat{n}_2(\hat{n}_2 - 1)) \\
+ \Delta_1 \hat{n}_1 + \Delta_2 \hat{n}_2)\big),
\end{aligned}
\tag{39}
$$

where $t \geq 0$ is the time after which the super-lattice is switched and a new realization of the laser speckle is created. We assume $\Delta_{1,2}$ to be Gaussian distributed, and we have neglected global phases. Now one can adopt an argument used in Ref. [29] for showing that the Gaussian operations together with a single non-Gaussian one allow for continuous-variable quantum computation. For universality to hold, one has to generate all operations where the corresponding Hamiltonian is a polynomial in the creation and annihilation operators where every monomial must contain an equal number of creation and annihilation operators, i.e., must be balanced, to ensure particle-number conservation. This is true as Eq. (39) contains all quadratic terms and a single quartic one. Since one can generate the entire algebra generated from the original set of Hamiltonian by commutation, one can approximate an arbitrary unitary [29]. Thus, by varying $\Delta_{1,2}$ we can approximate any gate to arbitrary accuracy which implies, by continuity of Eq. (39) universality as in Definition 5. By appropriately choosing the distribution from which t is chosen, the set can be made closed under Hermitian conjugation. Now, we can apply Observation 4 which shows together with Observation 1 that one can perform efficient compressed sensing by using a optical super-lattices, laser speckles, and time-of-flight imaging.

VI. MORE EFFICIENT TOMOGRAPHY SCHEME FOR MATRIX PRODUCT STATES AND OPERATORS

Even though compressed sensing notably reduces the number of necessary measurements, it still scales exponentially with the number of lattice sites. Without any further assumption on the state, this cannot be overcome. However, when the state is described by a generic matrix-product state (MPS) with a fixed bond dimension, tomography is possible with the number of measurements growing, in general, almost linearly with the system size and being constant for translationally invariant MPS. Ground states of gapped Hamiltonians which are a sum of terms acting only on a constant number of lattice sites are, generically, of this type [30].

A. Reconstructing reduced density matrices

The exponential reduction of the necessary number of measurements if the state is a MPS is due to the fact that such

states are completely determined by their reduced density matrices on all blocks of l consecutive sites where l only depends on the bond dimension [9, 30]. In Ref. [9], an efficient algorithm is given for finding the MPS matrices from these reduced density matrices. Note that this procedure only works if an upper bound to the bond-dimension, or, equivalently, to the locality size of the Hamiltonian is known. This is obvious as there could always be long range correlations which do not affect the l site reduced density matrices.

Let $d_B = d_l^l$ be the dimension of the subsystem under consideration and define the operator $\mathcal{T}_{R_q} : \mathcal{B}(\mathbb{C}^d) \mapsto \mathcal{B}(\mathbb{C}^{d_B})$ be the operator acting as $\mathcal{T}_{R_q}(\rho) = \mathrm{Tr}_{R_q} \rho$ and R_q denote all lattice sites but $q, \ldots, q + l - 1$. Then, the tight frame-condition of Definition 2 becomes

$$
\mathcal{T}_{R_q} \mathcal{W}_{d_B} = \mathcal{T}_{R_q}
\tag{40}
$$

If additional constraints apply, the condition reads

$$
\mathcal{T}_{R_q} \mathcal{W}_{d_{BN}} \mathcal{P}_N = \mathcal{T}_{R_q} \mathcal{P}_N
\tag{41}
$$

where d_{BN} is the dimension of the matrix block corresponding to the eigenvalue N of $\hat{N}$, restricted to the subsystem of l lattice sites. We concentrate on the former case as the second follows in an analogous way and show that a random circuit with a depth which does not depend on the system size can realize such a reduced tight frame: Assume w_0 to be a sum of traceless observables, each acting non-trivially only on some block of l lattice sites. For compressed sensing to be possible, it is sufficient for ν to induce, for every q, an approximate reduced tight frame with

$$
\|\mathcal{T}_{R_q}(\mathcal{W}_{d_B, \nu} - \mathbb{1})\| \leq \sqrt{d_B}(d_B^2 - 1)\varepsilon.
\tag{42}
$$

This is the case if it is an approximate reduced unitary 2-design with

$$
\forall q : \sup_{X, \|X\|_2 = 1} \left\| \mathrm{Tr}_{R_q}(\mathcal{G}_\nu - \mathcal{G}_{\mathrm{H}_l}^{(q)})(X) \right\|_2 \leq \varepsilon.
\tag{43}
$$

where $\mathcal{G}_{\mathrm{H}_l}^{(q)}$ denotes the channel which acts as $\mathcal{G}_\mathrm{H}$ on a block of l sites starting at q and as the identity on the rest of the system. To see that this is true, we calculate

$$
\begin{aligned}
\|\mathcal{T}_{R_q}(\mathcal{W}_{d_B, \nu} - \mathbb{1})\| = \\
\sup_{X, \|X\|_2 = 1} \|\mathrm{Tr}_{R_q}(\mathcal{W}_{d_B, \nu}(X) - \mathcal{W}_{d_B, \mathrm{H}}(X))\|_2,
\end{aligned}
\tag{44}
$$

which yields the desired result after inserting (17) and applying (43).

To show that obtaining such a tight frame is efficiently possible, we adapt Observation 4:

Observation 5 (Reduced tight frames by reduced unitary 2-designs)**.** *Let the parallel random circuit be as in Observation 4. There exists some constant C such that it forms an ε-approximate reduced l-site 2-design as defined in (42) after $n = C \log(1/\varepsilon) l \log l$ steps.*
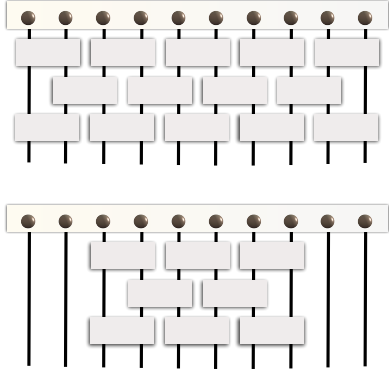
FIG. 2. Top: Parallel random circuit acting on the entire system. Bottom: Random circuit used in the proof of Observation 5 which acts only on the inner sites. The above circuit randomizes the states of the inner sites not less than the above one.

Proof: We show the fast convergence of our random circuit to a reduced unitary 2-design by comparing it with another random circuit which is easier to deal with, see Fig. 2. In analogy to Eq. (24) we denote by $\mathcal{G}_{\nu_l}^{(q)}$ the channel corresponding to an application of the parallel quantum circuit to a block consisting of l lattice sites starting from q. Observation 4 implies

$$\sup_{X, \|X\|_2 = 1} \|\mathrm{Tr}_{R_q}((\mathcal{G}_{\nu_l}^{(q)})^n - \mathcal{G}_{\mathrm{H}_l}^{(q)})(X)\|_2 \leq \varepsilon \qquad (45)$$

for $n = C\log(1/\varepsilon)l\log l$ where C is a constant. Using that $\lambda_2(G_{\nu_k}) \leq \lambda_2(G_{\nu_l})$, one obtains (43) concluding the proof.

B. Complexity of classical post-processing

Observation 5 implies that a random quantum circuit of constant depth is sufficient to perform tomography on a reduced density matrix of constant size. The number of operations, i.e. random unitary gates and measurements of expectation values of w_0, does only scale polynomially with the number of lattice sites. Therefore, we regard the quantum part of the protocol as efficient. However, this says nothing about the amount of post-processing needed because the reconstruction of the l-site reduced density matrices requires the knowledge of $\mathrm{Tr}_{R_q} w$ for all observables w obtained by the realizations of the random quantum circuit. If one needed to keep track of the evolution of observables on the entire Hilbert space this would require an exponential amount of computational resources. To see that this is not a problem in the present situation, we use the fact that w_0 is a sum of terms which act non-trivially only on blocks of constant size. Thus, the observables induced by the constant-depth random circuit can be written as

$$w(U) = \sum_i w_i(U) \otimes \mathbb{1}_{R_i}/(d_{R_i})^{1/2} \qquad (46)$$
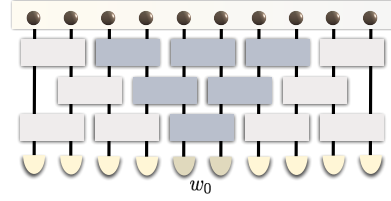

FIG. 3. Influence region for a local observables. Only the action of the darkly colored gates influences the measured observables.

where $w_i(U)$ acts non-trivially only on a block of constant size L starting with site i and where R_i denote all lattice sites but $i, \ldots, i + L - 1$, see Figure 3. From Eq. (46), we get

$$\mathrm{Tr}_{R_q}(w(U)) = \sum_i \sqrt{d_{R_i}} \mathrm{Tr}_{R_q}(w_i(U)), \qquad (47)$$

which means that one only needs to deal with observables on k Hilbert spaces which all have dimension d_l^L not depending on k making also the classical part of the protocol efficient. If the system is assume to be translationally invariant, all reduced density-matrices are equal, reducing the necessary number of measurements to a constant. Roughly speaking, the random circuit transforms a local observable to a reduced tight frame on l lattice sites with some influence on $L > l$ sites and none on the rest of the system.

C. Mixed states

Even though the method developed in this section is, in its present form, limited to pure states, it can be naturally extended to mixed states by using recent results on tomography for matrix-product *operators* (MPO) which are a natural generalization of MPS [31, 32]. In Ref. [33], it is shown that large classes of MPO states can be efficiently reconstructed from the reduced density matrices on a constant number of sites not depending on the system size. As we have presented a method of recovering these objects, one directly obtains a way of performing tomography on states in optical lattices which are described be MPO while requiring the same experimental techniques.

VII. CONCLUSION

In this article, we have presented a new route towards efficient quantum state tomography for quantum many-body systems, specifically for bosons in optical lattices. By using random circuits, which can be implemented by means of super-lattices and laser speckles, one can avoid the use of tomographically complete local measurements. These are very challenging, needless to say, but rely solely on time-of-flight imaging techniques, and seem conceivable with present technology. Without any further assumptions to the state, the number of necessary measurements is optimal up to constants and

logarithmic factors in the systems dimension. Restricting the set of possible states to matrix product states, both the number of measurements and the depth of the required random quantum circuit does not at all depend on the system size. This idea gives rise to the exciting perspective of actually measuring out the quantum state of a quantum many-body system in the laboratory.

There are a number of questions arising from this: For example, it would be very interesting to compare the performance of the present scheme, which is based on random circuits, with one where the quantum gates are chosen in an optimal way from some set of feasible operations. While there is not much room for improvements concerning the asymptotic behavior, the performance for small systems might differ notably. This will be studied extensively by numerical means in forthcoming work. We hope that this work stimulates further work – both of theoretical and especially experimental kind – in "quantum system identification", in order to innovate ways of rendering quantum state tomography feasible for large quantum systems.

[1] H. Haeffner, et al., Nature **438**, 643 (2005).
[2] J. B. Altepeter et al., Phys. Rev. Lett. **90**, 193601 (2003).
[3] G. M. D'Ariano and P. Lo Presti, Phys. Rev. Lett. **86**, 4195 (2001).
[4] E. J. Candès and T. Tao, IEEE Trans. Inf. Th. **56**, 2053 (2010).
[5] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Phys. Rev. Lett. **105**, 150401 (2010).
[6] D. Gross, IEEE Trans. on Inf. Th. **57**, 1548 (2011).
[7] A. Shabani et al., Phys. Rev. Lett. **106**, 100401 (2011).
[8] M. Ohliger, V. Nesme, D. Gross, Y.-K. Liu, and J. Eisert, arXiv:1111.0853.
[9] M. Cramer, M. B. Plenio, S. T. Flammia, D. Gross, S. D. Bartlett, R. Somma, O. Landon-Cardinal, Y.-K. Liu, and D. Poulin, Nat. Commun. **1**, 149 (2010).
[10] F. G. S. L. Brandão and M. Horodecki, arXiv:1010.3654.
[11] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, submission for QIP 2012.
[12] R. A. Low, PhD thesis, University of Bristol (2009).
[13] D. Clément, A. F. Varón, J. A. Retter, L. Sanchez-Palencia, A. Aspect, and P. Bouye, New J. Phys. **8**, 165 (2006).
[14] S. Trotzky, Y.-A. Chen, U. Schnorrberger, P. Cheinet, and I. Bloch, Phys. Rev. Lett. **105**, 265303 (2010); G.-B. Jo, J. Guzman, C. K. Thomas, P. Hosur, A. Vishwanath, and D. M. Stamper-Kurn, Phys. Rev. Lett. **108**, 045305 (2012).
[15] S. Trotzky, Y.-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwöck, J. Eisert, and I. Bloch, Nature Physics **8**, 325 (2012).
[16] A. W. Harrow and R. A. Low, Commun. Math. Phys. **291**, 257 (2009).
[17] C. Dankert, R. Cleve, J. Emerson, and E. Livine, arXiv:quant-ph/0606161.
[18] D. Gross, K. M. R. Audenaert, and J. Eisert, J. Math. Phys. **48**, 052104 (2007).
[19] E. J. Candès and Y. Plan, arXiv:1011.3854.
[20] S. T. Merkel, C. A. Riofrio, S. T. Flammia, and I. H. Deutsch, Phys. Rev. A **81**, 032126 (2010).
[21] D. Burgarth and K. Yuasa, Phys. Rev. Lett. **108**, 080502 (2012).
[22] B. Collins and P. Sniady, Comm. Math. Phys. **264**, 773 (2006).
[23] M. Greiner, O. Mandel, T. Esslinger, T. W. Hänsch, and I. Bloch, Nature **415**, 39 (2002).
[24] R. Jördens, N. Strohmaier, K. Günter, H. Moritz, and T. Esslinger, Nature **455**, 204 (2008).
[25] M. Lewenstein, A. Sanpera, V. Ahufinger, B. Damski, A. Sen, and U. Sen, Adv. Phys. **2**, 243 (2007); I. Bloch, J. Dalibard, and S. Nascimbene, Nature Physics **8**, 267 (2012).
[26] P. Barmettler, A. M. Rey, E. Demler, M. D. Lukin, I. Bloch, and V. Gritsev, Phys. Rev. A **78**, 012330 (2008); A. Flesch, M. Cramer, I. P. McCulloch, U. Schollwoeck, and J. Eisert, Phys. Rev. A **78**, 033608 (2008)
[27] A. Hoffmann and A. Pelster, Phys. Rev. A **79**, 053623 (2009).
[28] M. P. A. Fisher, P. B. Weichman, G. Grinstein, and D. S. Fisher, Phys. Rev. B. **40**, 546 (1989).
[29] S. Lloyd and S. L. Braunstein, Phys. Rev. Lett. **82**, 1784 (1999).
[30] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Quant. Inf. Comp. **7**, 401 (2007).
[31] F. Verstraete, J. J. Garcia-Ripoll, and J. I. Cirac, Phys. Rev. Lett. **93**, 207204 (2004).
[32] M. Zwolak and G. Vidal, Phys. Rev. Lett. **93**, 207205 (2004).
[33] T. Baumgratz et al., in preparation (2012).

Chapter 7

Solving frustration-free spin systems

PHYSICAL REVIEW LETTERS

Solving Frustration-Free Spin Systems

N. de Beaudrap,[1] M. Ohliger,[1] T. J. Osborne,[2] and J. Eisert[1,2]

[1]*Institute of Physics and Astronomy, University of Potsdam, 14476 Potsdam, Germany*
[2]*Institute for Advanced Study Berlin, 14193 Berlin, Germany*
(Received 22 May 2010; published 6 August 2010)

We identify a large class of quantum many-body systems that can be solved exactly: natural frustration-free spin-1/2 nearest-neighbor Hamiltonians on arbitrary lattices. We show that the entire ground-state manifold of such models can be found exactly by a tensor network of isometries acting on a space locally isomorphic to the symmetric subspace. Thus, for this wide class of models, real-space renormalization can be made exact. Our findings also imply that every such frustration-free spin model satisfies an area law for the entanglement entropy of the ground state, establishing a novel large class of models for which an area law is known. Finally, we show that our approach gives rise to an ansatz class useful for the simulation of almost frustration-free models in a simple fashion, outperforming mean-field theory.

PACS numbers: 03.67.Mn, 03.65.Ud, 05.50.+q, 89.70.Eg

Understanding the physics of quantum many-body systems is a central goal of modern physics, as they can exhibit exotic phenomena with no parallel in classical physics, including topological effects and quantum phase transitions at zero temperature. However, the very source of their rich physics also leads to a major roadblock in their study: The Hilbert space dimension of these systems scales exponentially with the number of particles. This means that brute-force numerical techniques fail even for systems of only a handful of particles.

A key insight in the study of local quantum many-body systems is that naturally occurring states occupy only a small subspace of the Hilbert space which, in principle, is available to them. Specifically, it has been realized that ground, thermal, and dynamically evolving states are only weakly entangled: The entanglement entropy satisfies what is referred to as an "area law" [1–3]. This insight is the basis of the density-matrix renormalization group approach and higher-dimensional analogues [4]. So successful are these methods in practice that one is tempted to boldly conjecture that all *physically* relevant systems will soon be tractable to one or another of the numerical tools we have at hand. Recent results give cause for caution, showing that general numerical methods cannot well approximate the physics of an *arbitrary* local quantum system, even in 1D; these include local glassy models where approximating the ground-state energy is NP-hard [5], suggesting that such systems would be intractable. However, these results do not give much reason for practical concern thus far, as at least in 1D they rely on rather baroque constructions (involving very large local dimensions).

In this work, we approach the issue of the complexity or "hardness" of finding ground states from the other direction: We establish a large class of models for which the task of finding the ground state is *easy*, in that the ground-state manifold can be described exactly and efficiently. This is the class of all natural frustration-free

spin-1/2 models with nearest-neighbor interaction on general lattices. (By the qualifier "natural," we mean that all two-spin interaction terms have excited states which are entangled, which might be taken as implicit in the semantics of an "interaction term.") Extending ideas of Ref. [6] on QUANTUM 2-SAT and going beyond 1D models as in Ref. [7], we find that the complete ground-state manifold of such Hamiltonians can be constructed by reduction to the symmetric subspace of a smaller system. In doing so, we find that the resulting ground-state manifold can be efficiently grasped in terms of tree-tensor networks. We discuss how this allows expectation values of local observables to be computed efficiently. What is more, the ground states satisfy an area law. Physically, we can view this work as describing a large class of models for which an instance of real-space renormalization provides an exact solution to the true genuine quantum many-body model. Finally, we see how this construction—a tree-tensor network with a symmetric subspace as an input—can serve as an ansatz class to simulate systems which are "close" to frustration-free models, outperforming mean-field approaches in a very simple fashion.

In our analysis, we allow for nearest-neighbor Hamiltonians on arbitrary lattices. This could be a cubic lattice of some dimension or, more generally, any graph, the vertex set of which we denote by V. On this lattice, the spin Hamiltonian H is represented as

$$H = \sum_{\{a,b\}} h_{a,b} \tag{1}$$

for terms $h_{a,b}$ acting on pairs of spins $\{a, b\} \subset V$. By rescaling, we may require that the ground energy of each term $h_{a,b}$ is zero. The ground-state manifold M of such a Hamiltonian may be degenerate: We identify the ground state ρ with the maximal mixture over M. (This is a pure state only if H is nondegenerate.) We describe properties of the ground state ρ and, more generally, the manifold M, given H as in Eq. (1). The Hamiltonian H is frustration-

free (or unfrustrated) if the ground-state vectors $|\Phi\rangle \in M$ correspond to ground states of individual coupling terms, that is, if $h_{a,b}|\Phi\rangle = 0$ holds for all $h_{a,b}$ and all $|\Phi\rangle \in M$; we say otherwise that H is frustrated.

We will call a spin Hamiltonian H natural if it contains no isolated subsystems, and each interaction term $h_{a,b}$ (considered as an operator on $\mathbb{C}^2 \otimes \mathbb{C}^2$) has at least one entangled excited state (i.e., an entangled state orthogonal to the ground-state manifold of $h_{a,b}$). In what follows we will consider only such natural Hamiltonians.

Frustration-free spin Hamiltonians.—Our main results concern the class of frustration-free spin Hamiltonians H. We show that the ground-state manifold M of such a spin Hamiltonian on N spins has dimension at most $N + 1$. What is more, it is the image of a space of low Schmidt measure [8] under a tree-tensor network.

First, we describe the needed components of Ref. [6], in the language of frustration-free models. Consider a Hamiltonian H_U containing terms $h_{u,v}$ of rank 2 or 3. If H_U is frustration-free, the reduced state $\rho_{u,v}$ of any state vector $|\Phi\rangle \in \ker(H)$ is in the kernel of $h_{u,v}$; we may then consider a subspace $S_{u,v} \subset \mathcal{H}_u \otimes \mathcal{H}_v$ of dimension 2 which contains $\mathrm{supp}(\rho_{u,v})$. By defining an isometry

$$R_{u:uv} : \mathcal{H}_2^{\otimes\{u\}} \to S_{u,v} \subset \mathcal{H}_2^{\otimes\{u,v\}}, \qquad (2)$$

we can reduce to a Hamiltonian on fewer spins: We let

$$H'_U = R_{u:uv}^\dagger H_U R_{u:uv} = \sum_{\{a,b\}} R_{u:uv}^\dagger h_{a,b} R_{u:uv}. \qquad (3)$$

Such a spin Hamiltonian H'_U is a sum of two-spin interactions (and possibly single-spin terms) of the form $h'_{a,b} = R_{u:uv}^\dagger h_{a,b} R_{u:uv}$. (If $h_{u,v}$ has rank 3, then $h'_{u,v}$ is a nonzero single-spin operator acting on u alone.) If H contains nonzero terms $h_{a,u}$ and $h_{a,v}$, we obtain two nonzero contributions $h'_{a,u} = R_{u:uv}^\dagger h_{a,u} R_{u:uv}$ and $h'_{a,v} = R_{u:uv}^\dagger h_{a,v} R_{u:uv}$ in the Hamiltonian H', each of which act on $\{u, a\}$. The sum gives a combined term $\bar{h}'_{a,u} = h'_{a,u} + h'_{a,v}$ in H'_U, possibly of higher rank than either $h'_{a,u}$ or $h'_{a,v}$ [9]. If the new Hamiltonian H'_U contains terms of rank 2 or 3, we may perform another such reduction, and so on. This reduction procedure has the following features:

a. Preservation of the kernel dimension.—By construction, we have $\ker(H_U) = R_{u:uv} \ker(H'_U)$. Thus, the kernels of H_U and H'_U have the same dimension. In particular, if H'_U has any terms of full rank acting on either one or two spins, then $\dim[\ker(H_U)] = 0$, in which case H_U is frustrated. If no full-rank terms are produced, each reduction leads to an operator acting on fewer spins, until we obtain a Hamiltonian having only terms of rank 1.

b. Arbitrariness of reduction order.—The dimension of the kernel is preserved by these reductions. We may perform such reductions until we obtain a Hamiltonian which either (i) contains only terms of rank 1 or (ii) contains a full-rank term. The latter cannot occur in the reduction of a frustration-free Hamiltonian; we discuss below the analy-

sis for Hamiltonians having only rank-1 terms. Thus, we may choose any convenient reduction sequence.

The above features allow us to reduce to the special case of a Hamiltonian H'_U (acting on a system V') which has only interaction terms of rank 1. Each two-spin Hamiltonian term $h'_{a,b} = |\beta_{a,b}\rangle\langle\beta_{a,b}|$ may be regarded as imposing constraints on the corresponding two-spin marginals $\rho_{a,b}$ of states $|\Phi\rangle \in \ker(H)$: We aim to obtain additional constraints on pairs of spins $u, v \in V'$ by combining the known constraints. To this end, Ref. [6] shows that a state $|\Phi\rangle$ which is in the kernel of two functionals $\langle\beta_{a,b}|$ and $\langle\beta_{b,c}|$ is also in the kernel of

$$\langle\beta'_{a,c}| = (\langle\beta_{a,b}| \otimes \langle\beta_{b,c}|)(\mathbb{1} \otimes |\Psi^-\rangle \otimes \mathbb{1}) \qquad (4)$$

acting on the spins $\{a, c\}$, where $|\Psi^-\rangle$ is the two-spin antisymmetric state vector. For each such "induced" constraint $\langle\beta'_{u,v}|$ on spins $\{u, v\}$, we may add the term $\tilde{h}_{u,v} = |\beta'_{u,v}\rangle\langle\beta'_{u,v}|$ to H'_U, resulting in a Hamiltonian $\tilde{H}_U$ which has the same kernel as H'_U. (If H'_U contains a term $h'_{u,v} \neq \tilde{h}_{u,v}$, it can be subsumed into a term $\bar{h}_{u,v}$ with rank at least 2, in which case we apply a reduction $R_{u:uv}$ as above.) One may induce further constraints from the terms of $\tilde{H}$, until we arrive at a "complete homogeneous" Hamiltonian H_c, having only terms of rank 1, for which the constraints $\langle\beta_{u,v}|$ are closed (up to scalars) under the constraint-induction procedure of Eq. (4).

Reference [6] shows that such a Hamiltonian H_c, acting on at least one spin and lacking single-spin operators [10], has a ground space containing product states. Thus, the above remarks essentially recap the following result.

Observation 1.—There is an efficient algorithm to determine whether a spin Hamiltonian is frustration-free.

We now extend the above results, to obtain a strong characterization of ground-state manifolds for natural frustration-free systems. We note the following three additional features of the isometric contraction scheme above:

c. Tree-tensor construction.—The complete network of isometric reductions T represents a tree-tensor network, a special case of the multiscale entanglement renormalization ansatz [11] which is related to real-space renormalization. Each isometry $R_{u:uv}$ has one free input tensor index and two free output indices, and the sequential nature of the reduction ensures that the network is directed and acyclic. Thus, any spin v introduced by an isometry $R_{u:uv}$ is a "daughter spin" of a unique parent u, leading to a treelike structure on the tensor network T. Note, however, that T has free input indices, corresponding to the roots of each tree: By construction, the ground space of H_U is the image $T|\Psi\rangle$ of states $|\Psi\rangle \in \ker(H_c)$.

d. Preservation of natural Hamiltonians.—Importantly, the isometric reductions above preserve the class of natural frustration-free spin Hamiltonians: That is, the mapping $h_{a,u} \mapsto R_{u:uv}^\dagger h_{a,u} R_{u:uv}$ does not decrease the rank of the interaction on $\{a, u\}$ and does not map the orthocomplement of the kernel to the space $|\phi\rangle \otimes \mathbb{C}^2$ for any $|\phi\rangle$.

e. Reduction to the symmetric subspace.—As a consequence of the previous feature, we may use the isometric reductions to map any natural frustration-free Hamiltonian to a complete homogeneous Hamiltonian H_c on a system V_c, in which the nonzero terms $h_{a,b} = |\beta_{a,b}\rangle\langle\beta_{a,b}|$ are supported on entangled states $|\beta_{a,b}\rangle$. We show that the kernel of such a Hamiltonian has small dimension and is spanned entirely by product states. For an arbitrary spin $a \in V_c$, we may let $L_a = \mathbb{1}$ and define a family of operators L_v satisfying

$$\langle\beta_{a,v}| \propto \langle\Psi^-|_{a,v}(\mathbb{1}_a \otimes L_v) \tag{5}$$

for spins $v \in V_c$ and operators $\langle\beta_{a,v}|$. One then finds that $C = \bigotimes_{v \in V_c} L_v$ is a linear isomorphism (not necessarily an isometry) from the subspace $\mathrm{Symm}(\mathcal{H}_2^{\otimes V_c})$ of symmetric states to the ground space of H_c. (This isomorphism is also noted in Ref. [12], Sec. III A, for generic Hamiltonians with rank-1 interactions.) As $\mathrm{Symm}(\mathcal{H}_2^{\otimes V_c})$ is spanned by uniform superpositions $|W_k\rangle$ of standard basis states having Hamming weight $0 \leq k \leq n_c = |V_c|$, we have $\dim[\mathrm{Symm}(\mathcal{H}_2^{\otimes n_c})] = n_c + 1$. This subspace may also be spanned by product state vectors $|\alpha_0\rangle^{\otimes n_c}, \ldots, |\alpha_{n_c}\rangle^{\otimes n_c}$ for any set of $n_c + 1$ pairwise independent state vectors $|\alpha_j\rangle \in \mathcal{H}_2$. Thus, any complete homogeneous (natural) Hamiltonian H_c has a ground space spanned by vectors

$$|\Psi_j\rangle = \bigotimes_{v \in V_c}(L_v|\alpha_j\rangle) = C|\alpha_j\rangle^{\otimes n_c}, \tag{6}$$

for some $|\alpha_j\rangle \in \mathcal{H}_2$ as above. Coupled with the tree-tensor structure of the isometric reductions, this characterization has the following consequences.

Observation 2.—For an unfrustrated spin Hamiltonian H_U, any constant k, and k-local operators A, $\langle A\rangle$ can be efficiently computed with respect to ground states of H.

Let H_c be a homogeneous Hamiltonian acting on n_c spins, obtained by isometric reduction of an unfrustrated Hamiltonian H_U. Consider a k-local operator $\tilde{A}$. As $\ker(H_c)$ is spanned by product vectors $|\Psi_0\rangle = C|\alpha_0\rangle^{\otimes n_c}, \ldots, |\Psi_{n_c}\rangle = C|\alpha_{n_c}\rangle^{\otimes n_c}$ as in Eq. (6), we can efficiently compute the restriction of $\tilde{A}$ to $\ker(H_c)$ by evaluating

$$W(\tilde{A}) = \sum_{j,k=0}^{n_c} |j\rangle\langle\Psi_j|\tilde{A}|\Psi_k\rangle\langle k| \tag{7}$$

followed by a suitable transformation. Specifically, consider the operator $B = W(\mathbb{1})$; we have $B = U\Delta U^\dagger$ for some U unitary and Δ positive and diagonal. We find

$$\Delta^{-1/2}U^\dagger \sum_{j=0}^{n_c} |j\rangle\langle\Psi_j| = \sum_{j=0}^{n_c} |j\rangle\langle\Phi_j|, \tag{8}$$

for some orthonormal basis $|\Phi_0\rangle, \ldots, |\Phi_{n_c}\rangle$ of $\ker(H_c)$; thus, the restriction of $\tilde{A}$ to $\ker(H_c)$ with respect to the basis of states $|\Phi_j\rangle$ may be computed as

$$\bar{A} = \Delta^{-1/2}U^\dagger W(\tilde{A})U\Delta^{-1/2}. \tag{9}$$

[For $\tilde{A}$ consisting of a single k-spin term, the inner products of Eq. (7) are products of constant-dimensional inner products; for $\tilde{A}$ a sum of multiple terms, we extend linearly.] Let $T: \ker(H_c) \to \ker(H_U)$ be the network of isometric reductions. Then, by considering operators $\tilde{A} = T^\dagger AT$, we may compute the restriction $\bar{A}$ of such operators A to the ground space of H_U. We may then efficiently compute expectation values by using such matrices.

Observation 3.—Ground states of frustration-free spin Hamiltonians H_U on lattices obey an entanglement area law.

For any contiguous subsystem A containing a spins, we may reduce the Hamiltonian $H_U^{(A)}$ acting internally on A—by a tree-tensor isometry T_A acting on A alone—to obtain a homogeneous Hamiltonian $H_c^{(A)}$, acting on at most a spins. The ground space of $H_c^{(A)}$ has dimension at most $a + 1$; as the ground space of $H_U^{(A)}$ is an isometric image of that of $H_c^{(A)}$, the same is true of $H_U^{(A)}$. As H_U is unfrustration, any ground state of H_U is also a ground state of $H_U^{(A)}$; it follows that the Schmidt measure [8] of $|\Psi\rangle$ with respect to the bipartition $V = A \cup (V \setminus A)$ is at most $\log(a + 1)$. For a spin lattice of any dimension, we obtain an area law for arbitrary subsystems by summing this logarithmic bound over the number of distinct connected components.

"Almost" frustration-free Hamiltonians.—We now leave the rigorous exact setting and turn to the observation that the above techniques can be used to grasp approximately frustration-free models. Observation 2, in particular, suggests a variational approach to estimating ground energies for Hamiltonians H for which

$$H = H_U + \lambda H_F \tag{10}$$

for some $\lambda \ll 1$, where H_U is frustration-free but the Hamiltonian H itself is frustrated. For such Hamiltonians H, if no phase transition is encountered, the eigenvalues and eigenstates may differ little from those of the unfrustrated Hamiltonian H_U (for related bounds on eigenvalues, see, e.g., Ref. [13]). If the lowest k eigenvectors (for some suitable $1 \leq k \leq n + 1$) have a sufficiently high overlap with the lowest k eigenvectors of H_U, we may approximately sample from the low-energy eigenvectors of H by restricting to the kernel of H_U, by using the efficient algorithm above. In particular, as this procedure is variational, estimates obtained in this way for the ground-state energy of H are guaranteed to be upper bounds.

We may consider improvements to this ansatz which retain more information about the "frustrating" component λH_F than in the tree-tensor renormalization procedure for H_U. To obtain better estimates—and to extend these techniques to the case where λ may be significantly large—we may consider partial reductions by tree-tensor networks, having many free input spins, and isometries depending on the terms of the perturbed Hamiltonian H (rather than those of H_U).

If the ground-state manifold of H_U is contained in a subspace $\mathcal{K}$ which is spanned by product states and has
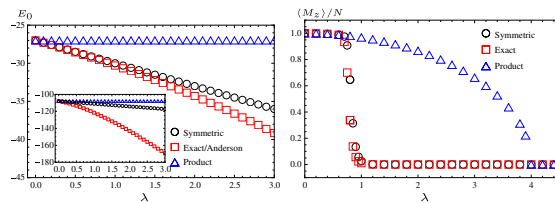
FIG. 1 (color online). Left: Ground-state energy for *XXZ* model on a trigonal lattice on a 3×3 torus, $h_{i,j} = -X_i X_j - Y_i Y_j - (1 - \lambda) Z_i Z_j$, by symmetric subspace estimate compared to product state ansatz and exact diagonalization. The inset shows the same model on a 6×6 torus where exact solution is not feasible and, therefore, is replaced by an Anderson lower bound. Right: Magnetization in z direction for Ising model in a transverse field on a 4×4 torus, $h_{i,j} = -Z_i Z_j$, $h_i = -\lambda X_i$, by symmetric subspace estimate compared to mean-field approximation (product state ansatz) and exact diagonalization.

"small" dimension (i.e., polynomial rather than exponential in the system size, as in the case of the symmetric subspace for ferromagnetic Ising or *XXX* models), we may indeed forgo isometric reductions entirely and estimate the ground energy of H by considering the restriction of H to $\mathcal{K}$ by using the techniques described for Observation 2. In contrast to a full tree-tensor contraction, this has the advantage of yielding exact results for the frustration-free case $\lambda = 0$ while retaining more information about the frustrating component H_F. The resulting estimate for the ground-state energy will be a linear function of λ, whose value and first derivative with respect to λ agree with that of the exact ground energy at $\lambda = 0$. For small λ and modest system sizes, this may yield a good estimate of the ground-state energy of H; see Fig. 1 [14].

To obtain estimates which account for spatially decaying correlations, we may perform a partial tree-tensor reduction with a small number of contraction layers and sample with respect to a subspace $\mathcal{K}$ as described above. In each layer, we may fix a collection of (nonintersecting) adjacent site pairs $\{a, b\}$ to contract and for each such pair $\{a, b\}$ apply some isometric contraction as described in Eq. (3). However, rather than apply the reductions which would be suggested by the frustration-free Hamiltonian H_U, we may use isometries

$$Q_{a:ab} = |\psi_{a,b}^0\rangle\langle 0| + |\psi_{a,b}^1\rangle\langle 1|, \tag{11}$$

where $|\psi_{a,b}^0\rangle$ and $|\psi_{a,b}^1\rangle$ are the lowest energy eigenvectors of the interaction term of H on sites a and b. Given a tensor network T_1 consisting of a product of such two-site operators $Q_{a:ab}$, we may then consider the spin model given by $H' = T_1^\dagger H T_1$ and estimate the ground energy of H' with respect to a low-dimension subspace or another isometric contraction; this yields an upper bound on the ground energy of the original Hamiltonian H.

Summary.—In this work, we have introduced a class of spin models that can be completely solved: The entire ground-state manifold can be explicitly given and parametrized by an entire symmetric subspace under a tensor network. This class of models is expected to provide a rich playground of exploring ideas on quantum lattice models, complementing work that exemplifies how computationally difficult it can be to approximate ground-state energies. The considered models can also be viewed as the parent Hamiltonian of the network, in a converse approach taken for tree-tensor networks in Ref. [15]. The models we consider obey area laws, establishing a class of models beyond free systems for which this holds. It is the hope that this work stimulates further research on models for which tensor networks arise not only as computational but as essentially analytical tools.

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

[2] M. Srednicki, Phys. Rev. Lett. **71**, 666 (1993).

[3] K. M. R. Audenaert, J. Eisert, M. B. Plenio, and R. F. Werner, Phys. Rev. A **66**, 042327 (2002); G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev, Phys. Rev. Lett. **90**, 227902 (2003); B.-Q. Jin and V. Korepin, J. Stat. Phys. **116**, 79 (2004); P. Calabrese and J. Cardy, J. Stat. Mech. (2004) P06002.M. B. Plenio, J. Eisert, J. Dreissig, and M. Cramer, Phys. Rev. Lett. **94**, 060503 (2005); M. B. Hastings, J. Stat. Mech. (2007) P08024.

[4] F. Verstraete, J. I. Cirac, and V. Murg, Adv. Phys. **57**, 143 (2008); U. Schollwoeck, Rev. Mod. Phys. **77**, 259 (2005).

[5] D. Aharonov, D. Gottesman, S. Irani, and J. Kempe, Commun. Math. Phys. **287**, 41 (2009).

[6] S. Bravyi, arXiv:quant-ph/0602108.

[7] R. Movassagh, E. Farhi, J. Goldstone, D. Nagaj, T. J. Osborne, and P. W. Shor, arXiv:1001.1006 [Phys. Rev. A (to be published)].

[8] J. Eisert and H. J. Briegel, Phys. Rev. A **64**, 022306 (2001).

[9] We similarly accumulate any single-spin contributions in H_U' which arise from single-spin contributions in H_U.

[10] While we effectively ignore single-spin terms in our analysis, such terms do impose constraints on the ground-state manifold and thus on whether the Hamiltonian is frustration-free. However, these may be accounted for with little difficulty.

[11] G. Vidal, Phys. Rev. Lett. **99**, 220405 (2007).

[12] C. R. Laumann, R. Moessner, A. Scardicchio, and S. L. Sondhi, arXiv:0903.1904.

[13] S. Bravyi, M. B. Hastings, and S. Michalakis, arXiv:1001.0344.

[14] For systems whose ground states have very localized correlations, estimating with respect to a subspace spanned by a small number of product states may scale poorly for large systems. When $\mathcal{K} = \text{Symm}(\mathbb{C}^N)$, the very symmetry of states in $\mathcal{K}$ entails that correlations do not decay spatially but are constant.

[15] P. Silvi, V. Giovannetti, S. Montangero, M. Rizzi, J. I. Cirac, and R. Fazio, Phys. Rev. A **81**, 062335 (2010).

Chapter 8

Discussion and conclusion

In the five articles, we have elaborated on various aspects of quantum systems which are relevant for quantum information where we have put an emphasis on measurement-based quantum computing and quantum state tomography. In the present chapter, we give a short summary of those articles and also mention what the contributions of the author were. Furthermore, we provide a brief conclusion and give some outlook to future research projects.

- **Limitations of quantum computing with Gaussian cluster states**a

  In this article, we have discussed the limitations of using Gaussian states as resources for measurement-based quantum computing. Gaussian cluster states are the natural continuous-variable generalizations of qubit cluster states, which were the first known resource for MBQC. Gaussian cluster states allow for universal MBQC where only Gaussian measurements and a single non-Gaussian one are necessary [37]. Unfortunately, in their ideal form, they need an infinite amount of energy to be prepared which renders them unphysical. We have shown that in any Gaussian state with finite energy, the entanglement localizable by Gaussian measurement decays exponentially in the distance between the modes. This is true on any graph and shows that every scalable MBQC scheme would need non-Gaussian measurements already to just transport a single qubit. Furthermore, if the modes are arranged on a one-dimensional lattice, transport is impossible even when allowing for completely arbitrary measurements. These results gravely limit the use of Gaussian states in creating a scalable measurement-based quantum computer.

  The author's contribution was to show the technical results and to write the manuscript while acknowledging support from KK and JE especially on section IV.C, V.B, and the appendix.

- **Efficient measurement-based quantum computing with continuous-variable systems**

  In this chapter, we have continued the investigation of continuous-variable MBQC and have both introduced a unifying framework allowing for the description of MBQC schemes both with qudits and CV systems. We have formalized the requirements for efficient quantum computing in this paradigm which helped us to find another important no-go result: In a system composed of one-dimensional quantum wires, scalable measurement-based quantum computing with Gaussian measurements only is impossible when the one-dimensional resource state can be described by a matrix product state with constant bond dimension. Thus, both non-Gaussian resource state and non-Gaussian measurements are necessary. As a second main result, we have provided the first scalable CV-MBQC scheme. It only uses coherent states as an input, beam-splitters and cross-Kerr cells as interactions, and displaced photon-counting as measurement. Even though the last two requirements are experimentally quite challenging, they have all already been demonstrated at least in proof-of-principle experiments [38].

  The main contribution of the author was to formulate and show the observations made in this article and to write the manuscript. He acknowledges contributions of JE to this article, especially concerning the presentation.

- **Continuous-variable quantum compressed sensing**

  We have introduced the concept of continuous-variable compressed sensing. We were able to provide sufficient conditions observables must fulfill to be of use for efficient tomography. These conditions are not merely of purely theoretical interest but they are fulfilled for the measurements most commonly used in continuous-variable quantum optics. Furthermore, we have presented a method to certify the success of the tomography protocol from the measured data. This is especially of interest when considering medium sized systems which are accessible to present day experiments. We have exemplified this with numerical studies showing the superior performance of compressed sensing.

  The main part of manuscript of this article has been written by the author who also showed, together with his co-authors, most of the technical results. This excludes sections 4 and 5.2 as well as the last appendix.

- **Efficient and feasible state tomography of quantum many-body systems**

  We have made use of the technique developed in the previous chapter to propose a scheme for full quantum state tomography of ultra-cold atoms in optical lattices which only relies on the techniques of speckle-patterns and super-lattices which have both already demonstrated in experiment. Under the assumption that the state is a generic ground state of a Hamiltonian, which is a sum of local terms, the required effort, both measurement procedure and post-processing, grows only polynomially in the system size. This is to be compared to a necessarily exponential growth in the general situation. The protocol is based on the observation that compressed sensing is possible by applying a random unitary to the quantum state in question followed by a measurement of a fixed observable and repeating this procedure to get the required amount of expectation values. We have shown that such random unitary matrices, called approximate unitary 2-designs, can be realized by suitable random circuits which, in turn, can be performed by means of the mentioned super-lattices and speckle-patterns. We have discussed the scaling of the necessary depth of the random circuit, i.e., the number of operations performed before the measurements, with the system size and have shown that it is, in the particular important case of matrix product state tomography, independent of it.

  The manuscript was written by the author. Contributions of VN and JE to the idea of this research projects, the mathematical foundations, and the presentation of the results are acknowledged.

- **Solving frustration-free spin systems**

  In the first four articles, we have investigated the particularities of continuous-variable systems with respect to measurement-based quantum computing and compressed sensing while in this chapter, we have considered quantum many-body systems consisting of qubits, in this context often called spins. We have presented a method to find the ground-state manifold of a large class of Hamiltonians, i.e., frustration free nearest-neighbor ones on any graph. Such Hamiltonians are sums of terms acting non-trivially only on two neighboring systems. In addition, the global ground state is simultaneously the ground state of all the terms contributing to the Hamiltonian. We have done this by adapting the method used in Ref. [39]. The key idea is to perform a sequence of isometric re-

ductions which leave the ground-state invariant. At the end, one obtains an isometric tree-tensor network which has the property that the image of the symmetric subspace under it is exactly the ground-state manifold. As the algorithm to find this tree-tensor network scales polynomially in the number of spins N and the dimension of the symmetric subspace is $N+1$, the solution fulfills the requirements of efficiency. The notable difference to other techniques based on tensor-networks is that we were able to provide an analytic, compared to a merely numerical, algorithm. Furthermore, our algorithm can be used as an approximative technique to find solutions of Hamiltonians which are close to being frustration free. We have applied this technique to the XXZ model and compared our findings to results obtained by exact diagonalization, variational methods, and Anderson bound techniques showing a favorable performance at least in part of the parameter range.

The main contributions of the author to this work were the numerical results and the part on the idea of using the presented method as a simulation technique for slightly frustrated systems.

The main aim of this thesis was to investigate and characterize continuous-variable quantum states where we have focused on two important aspects. The first two articles have discussed the possibility of using such states as resources for measurement-based quantum computing (MBQC). This was of particular interest in the view of the research program started with Ref. [37] in which Gaussian analogs of the cluster state, as used in the first MBQC protocol [6], were introduced and which inspired a substantial amount of research in this direction. We have contributed to this efforts by developing a framework to describe a very general class of MBQC schemes both discrete and continuous. This allows to clearly state the requirements any efficient MBQC scheme has to fulfill. Using this framework, we have shown serious limitations of the Gaussian cluster state approach on one hand and presented a measurement-based computing scheme not suffering from these drawbacks on the other hand. Besides the direct implications, these results highlight again the necessity of non-Gaussian elements in quantum information. While it was known before that non-Gaussian measurements are needed, we have shown that non-Gaussian resource states are required at the same time. This is a substantial difference to the situation of entanglement distillation where non-Gaussian resource states are sufficient [19–21]. This motivates a closer investigation of the resource character of non-Gaussian states as already started in Appendix A.

The second subfield of quantum information, this thesis has contributed to, is quantum state tomography. We have addressed this problem from three directions. On the theoretical side, we have generalized the technique of compressed sensing such that a large class of measurements can be used. This has allowed us to give detailed description how compressed sensing can be performed in experiments involving quantum optics or ultra-cold gases in optical lattices. To connect our theoretical findings, which are mainly concerned with asymptotic statements, i.e., the scaling of the effort with the system size, to experiments, we have also performed numerical simulations. They show that compressed sensing outperforms other methods of quantum state tomography not only in the asymptotic limit but also in medium size systems accessible in present day experiments. A next step in this research program is to apply these results to actual experimental data to show the superiority of this method also in real-life situations.

We hope the present work is a little puzzle piece in the quest for a better understanding of the properties of quantum mechanical systems and for realizing a quantum computer and other quantum information devices.

Bibliography

[1] A. Einstein, B. Podolsky, and N. Rosen, *Can quantum-mechanical description of physical reality be considered complete?*, Phys. Rev. 47, **777** (1935).

[2] J. S. Bell, *On the Einstein Podolsky Rosen Paradox*, Physics **1**, 195 (1964).

[3] A. Aspect et. al. *Experimental Tests of Realistic Local Theories via Bell's Theorem*, Phys. Rev. Lett. **47**, 460 (1981).

[4] P. W. Shor, *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*, arXiv:quant-ph/9508027.

[5] R. Ursin, et al., *Entanglement-based quantum communication over 144 km*, Nature Physics **3**, 481 (2007).

[6] R. Raussendorf, and H. J. Briegel, *A One-Way Quantum Computer*, Phys. Rev. Lett. **86**, 5188 (2001).

[7] D. Gross and J. Eisert, *Novel schemes for measurement-based quantum computation*, Phys. Rev. Lett. **98**, 220503 (2007).

[8] D. Gross, J. Eisert, N. Schuch, and D. Perez-Garcia, *Measurement-based quantum computation beyond the one-way model*, Phys. Rev. A **76**, 052315 (2007).

[9] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, *Quantum state tomography via compressed sensing*, Phys. Rev. Lett. **105**, 1504001 (2010).

[10] D. Gross, *Recovering low-rank matrices from few coefficients in any basis*, IEEE Trans. Inf. Theory, **57**, 1548 (2011).

[11] M. Greiner, et. al. *Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms*, Nature **415**, 39 (2002).

[12] M. Cramer et. al. *Efficient quantum state tomography*, Nature Comm. **1**, 149, (2010).

[13] E. Jané, G. Vidal, W. Dür, P. Zoller, and J. I. Cirac, *Simulation of quantum dynamics with quantum optical systems*, Quant. Inf. Comp. **3**, 15 (2003).

[14] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, *Matrix Product State Representations*, Quant. Inf. Comp. **7**, 401 (2007).

[15] U. Leonhardt, *Measuring the Quantum State of Light*, Cambridge University Press (1997).

[16] W. P. Schleich, *Quantum Optics in Phase Space*, Wiley-VCH (2001).

[17] A. Mari, K. Kieling, B. Melholt Nielsen, E. S. Polzik, and J. Eisert, *Directly estimating non-classicality*, Phys. Rev. Lett. **106**, 010403 (2011).

[18] S. D. Bartlett, B. C. Sanders, S. L. Braunstein, K. Nemoto, *Efficient Classical Simulation of Continuous Variable Quantum Information Processes*, Phys. Rev. Lett. **88**, 097904 (2002).

[19] J. Eisert, S. Scheel, and M. B. Plenio, *Distilling Gaussian States with Gaussian Operations is Impossible*, Phys. Rev. Lett. **89**, 137903 (2002).

[20] J. Fiurasek, *Gaussian transformations and distillation of entangled Gaussian states*, Phys. Rev. Lett. **89**, 137904 (2002).

[21] G. Giedke and J. I. Cirac, *The characterization of Gaussian operations and Distillation of Gaussian States*, Phys. Rev. A **66**, 032316 (2002).

[22] E. T. Campbell and J. Eisert, *Gaussification and entanglement distillation of continuous variable systems: a unifying picture*, Phys. Rev. Lett. **108**, 020501 (2012).

[23] M. S. Kim, E. Park, P. L. Knight, and H. Jeong, *Nonclassicality of a photon-subtracted Gaussian field*, Phys. Rev. A **71**, 043805 (2005).

[24] T. Kiesel, W. Vogel, M. Bellini, and A. Zavatta, *Nonclassicality Quasiprobability of Single-Photon Added Thermal States*, Phys. Rev. A **83**, 032116 (2011).

[25] A. M. Brańczyk and T. C. Ralph, *Teleportation using squeezed single photons*, Phys. Rev. A **78**, 052304 (2008).

[26] A. Ourjoumtsev, R. Tualle-Brouri, J. Laurat and P. Grangier, *Generating Optical Schrödinger Kittens for Quantum Information Processing*, Science **7**, 83 (2006).

[27] M. Kohl, K. Günter, T. Stöferle, H. Moritz, and T. Esslinger, *Strongly interacting atoms and molecules in a 3D optical lattice*, J. Phys. B **39**, 47 (2006).

[28] S. Ospelkaus et al., *Localization of Bosonic Atoms by Fermionic Impurities in a Three-Dimensional Optical Lattice*, Phys. Rev. Lett. **96**, 180403 (2006).

[29] S. Trotzky et al., *Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional Bose gas*, Nature Physics **8**, 325 (2012).

[30] M. P. A. Fisher, P. B. Weichman, G. Grinstein, and D. S. Fisher, *Bose localization and the superfluid-insulator transition* Phys. Rev. B **40**, 546 (1989).

[31] T. D. Kuehner, S R. White, and H. Monien, *The one-dimensional Bose-Hubbard Model with nearest-neighbor interaction* Phys. Rev. B **61**, 12474 (2000).

[32] A. Eckardt et al., *Frustrated quantum antiferromagnetism with ultracold bosons in a triangular lattice*, Europhys. Lett. **89**, 10010 (2010).

[33] D. Gross and J. Eisert, *Quantum computational webs*, Phys. Rev. A **82**, 040303(R) (2010).

[34] A. I. Lvovsky and M. G. Raymer, *Continuous-variable optical quantum-state tomography*, Rev. Mod. Phys. **81**, 299 (2009).

[35] E. J. Candes, T. Tao, *The Power of Convex Relaxation: Near-Optimal Matrix Completion*, IEEE Trans. Inf. Theory **56**, 2053 (2010).

[36] Y.-K. Liu, *Universal low-rank matrix recovery from Pauli measurements*, Advances Neural Inf. Proc. Systems **24**, 1638 (2011).

[37] N. C. Menicuicci et al., *Universal Quantum Computation with Continuous-Variable Cluster States*, Phys. Rev. Lett. **97**, 110501 (2006).

[38] K. Laiho, M. Avenhaus, K. N. Cassemiro and C. Silberhorn, *Direct probing of the Wigner function by time-multiplexed detection of photon statistics*. New. J. Phys. **11**, 043012 (2009).

[39] S. Bravy, *Efficient algorithm for a quantum analogue of 2-SAT*, arXiv:quant-ph/0602108.

[40] A. K. Ekert, *Quantum Cryptography Based on Bell's Theorem*, Phys. Rev. Lett. **67**, 661 (1991).

[41] M. G. Genoni, M. G. A. Paris, and K. Banaszek, *A measure of the non-Gaussian character of a quantum state*, Phys. Rev. A **76**, 042327 (2007).

[42] M. G. Genoni, M. G. A. Paris, and K. Banaszek, *Quantifying the non-Gaussian character of a quantum state by quantum relative entropy*, Phys. Rev. A **78**, 060303(R) (2008).

[43] J. Eisert and M. M. Wolf, *Gaussian Quantum Channels*, Quantum Information with Continous Variables of Atoms and Light, pages 23-42, Imperial College Press, (2007).

[44] A. Kenfack and K. Zyczkowski, *Negativity of the Wigner function as an indicator of non-classicality*, J. Opt. B **6**, 396 (2004).

[45] F. Nicacio, R. N. P. Maia, F. Toscano, R. O. Vallejos, *Phase Space Structure of Generalized Gaussian Cat States*, arXiv:1002.2248.

[46] M. Cramer, J. Eisert, and F. Illuminati, *Inhomogeneous Atomic Bose-Fermi Mixtures in Cubic Lattices*, Phys. Rev. Lett. **93**, 190405 (2004).

[47] W. Kohn, *Analytic Properties of Bloch Waves and Wannier Functions*, Phys. Rev. **115**, 809 (1959).

[48] D. S. Lühmann, K. Bongs, K. Sengstock, and D. Pfannkuche, *Self-Trapping of Bosons and Fermions in Optical Lattices* Phys. Rev. Lett. **101**, 050402 (2008).

[49] M. Ohliger and A. Pelster, *Green's Function Approach to the Bose-Hubbard Model*, arXiv:0810.4399.

Appendix A

Measures of non-Gaussianity

As discussed in Chapter 4, non-Gaussian quantum states are necessary for measurement-based quantum computing. Another important task in quantum information, where Gaussian states are required, is Gaussian entanglement distillation. Assume that two parties, called Alice and Bob, want to perform a quantum information protocol which requires them to share an entangled quantum optical state, e.g. entanglement-based quantum key distribution [40]. Now, Alice prepares such a state locally and then sends half of it to Bob. In reality the link connecting them, e.g. an optical fiber, will always be lossy. Thus, if the distance is too long, the state shared between Alice and Bob might not possess enough entanglement or might be too mixed. An entanglement-distillation protocol now allows to convert a number of weakly entangled, mixed states to a smaller number of highly entangled, purer states which can then be used to perform the required task. However, if Alice and Bob share only Gaussian states, entanglement distillation with Gaussian operations only is impossible [19–21]. Non-Gaussian states, on the other hand, allow to be distilled with Gaussian operations [19, 22].

This motivates the search for a quantification of non-Gaussianinity. The measure proposed first was based on the Hilbert-Schmidt distance [41], which, unfortunately lacks an operational meaning. A second one was based on the quantum relative entropy, which is connected to the distinguishability of two quantum states in the limit of many copies [42].

In this section, we define another non-Gaussianity measure based on the trace-distance between the state and a reference Gaussian state. As the trace-distance determines the probability to distinguish two quantum states in a single measurement, our measure has a operational meaning which is relevant for quantum information processing.

A.1 Non-Gaussianity based on trace-distance

For a quantum state ρ, we denote the Gaussian state with the same first and second moments by τ. We then define the non-Gaussianity by

$$\delta(\rho) = \frac{1}{2}\|\rho - \tau\|_1, \tag{A.1}$$

where $\|\cdot\|_1$ denotes the trace norm which is also called the Schatten 1-norm, and which is given by the sum of all singular values. From the properties of the trace norm, it is clear that δ is continuous, takes values only in the interval $[0, 1]$ and is zero if and only if ρ itself is a Gaussian state. Furthermore, it has the following properties:

Lemma 1. The measure δ is invariant under Gaussian unitary transformations.

Proof: The state ρ transforms as $\rho' = U\rho U^\dagger$. Its covariance matrix transform as $\sigma' = \Gamma\sigma\Gamma^T$ where Γ is the symplectic transformation corresponding to U while the vector of the first moments $\mathbf{d}$ transforms as $\mathbf{d}' = \Gamma\mathbf{d}$. The first and second moments of the Gaussian reference states undergo the same transformation. Because the Gaussian reference state is completely determined by its first and second moments, it must also transform as $\tau = U\tau U^\dagger$. Then, the invariance of δ follows from the unitary invariance of the trace norm.

Lemma 2. The measure δ stays invariant when adding a Gaussian state to the system.

Proof:

$$\begin{aligned}
\delta(\rho_A \otimes \tau_B) &= \frac{1}{2}\|(\rho_A \otimes \tau_B) - (\tau_A \otimes \tau_B)\|_1 \\
&= \frac{1}{2}\|(\rho_A - \tau_A) \otimes \tau_B\|_1 = \frac{1}{2}\|\rho_A - \tau_A\|_1\|\tau_B\|_1 = \delta(\rho_A).
\end{aligned} \tag{A.2}$$

Lemma 3. The measure δ does not increase under the partial trace-operation, i.e., $\delta(\mathrm{Tr}_B(\rho_{AB})) \leq \delta(\rho_{AB})$.

Proof: Because the second moments of the reduced state form a block in the covariance matrix, the Gaussian reference state of the reduced state is $\tau_A = \mathrm{Tr}_B(\tau_{AB})$. As the trace distance does not increase under the partial trace, δ must also be non-increasing.

Lemma 4. The measure δ does not increase under the application of a Gaussian channel.

Proof: Every Gaussian channel can be performed by the following procedure: Adding a Gaussian ancilla state, performing a Gaussian unitary operation, and tracing out the ancilla state [43]. Thus Lemma 4 follows directly from Lemmata 2, 1, and 3.

Lemma 5. The measure δ is subadditive under tensor products, i.e.,

$$\delta(\rho_A \otimes \rho_B) \leq \delta(\rho_A) + \delta(\rho_B). \tag{A.3}$$

Proof:

$$\begin{aligned}
\delta(\rho_A \otimes \rho_B) =& \frac{1}{2} \|(\rho_A \otimes \rho_B) - (\tau_A \otimes \tau_B)\|_1 \\
=& \frac{1}{2} \|\rho_A \otimes \rho_B - \tau_A \otimes \rho_B + \tau_A \otimes (\rho_B - \tau_B)\|_1 \\
\leq& \frac{1}{2} \|(\rho_A - \tau_A) \otimes \tau_B\|_1 + \|\tau_A \otimes (\rho_B - \tau_B)\|_1 \\
=& \delta(\rho_A) + \delta(\rho_B). \tag{A.4}
\end{aligned}$$

Single-mode continuous-variable quantum optical states are most conveniently described in terms of the Wigner function which is a real quasi-probability distribution of two phase-space coordinates and the most straightforward generalization of the classical phase-space distribution to the quantum regime [16].

Unfortunately, it is not possible to calculate δ directly from the Wigner function without first reconstruction the density matrix. However, there exists a lower bound to the non-Gaussianity which follows from the ordering of the different p-norms:

$$\begin{aligned}
\delta =& \frac{1}{2} \|\rho - \tau\|_1 \geq \frac{1}{2} \|\rho - \tau\|_2 = \frac{1}{2} \sqrt{\text{Tr}(\rho - \tau)^2} \\
=& \frac{1}{2} \sqrt{2\pi \int d\xi \, (W_\rho(\xi) - W_\tau(\xi))^2}, \tag{A.5}
\end{aligned}$$

where $\| \cdot \|_2$ denotes the Schatten 2-norm, i.e., the square root of the sum of the squares of the singular values.

A.2 Negativity of Wigner function

We note that the non-Gaussianity δ is not convex. Even more, a non-trivial convex combination of Gaussian states with different first moments can have a non-vanishing δ. This already points in the direction that δ may not be a suitable measure if one is interested in the properties of non-Gaussianity as a resource. For this reason, we examine a measure of non-classicality, which was introduced in Ref. [44] and which is given by the volume of the negative part of the Wigner function

$$\delta_N(\rho) = \int d\xi \, |W_\rho(\xi)| - 1, \tag{A.6}$$

where $W_\rho(\xi)$ is the Wigner function of the state ρ. One can show that it fulfills Lemmas 1, 2, 3, and 4. Furthermore, it is convex.

Lemma 6. The measure δ_N is convex, i.e., $\delta_N(p_1\rho_1 + p_2\rho_2) \leq \delta_N = p_1\delta_N(\rho_1) + p_2\delta_N(\rho_2)$ where $p_1, p_2 \geq 0$.

Proof: The proof directly follows from the definition (A.6) and the triangle inequality.

We now consider Gaussian measurements: As a Gaussian measurement can always be written as an x-quadrature measurement proceeded by a Gaussian unitary, we restrict ourselves to this case.

Lemma 7. The measure δ_N does not increase on average when performing a Gaussian measurement on a subsystem.

Proof: We state the proof for a two-mode system, however, it is valid for any number of modes. The Wigner function of the remaining system ($\rho_A(x_2)$) after performing a x-quadrature measurement on the second subsystem and obtaining the result x_2 is

$$\tilde{W}(x_1, p_1) = \frac{1}{p(x_2)} \int dp_2 \, W(x_1, p_1, x_2, p_2), \qquad (A.7)$$

where $p(x_2)$ is the probability distribution for the measurement results. The average Wigner function negativity is

$$\langle \delta_N(\rho_A) \rangle = \int dx_2 \, p(x_2) \delta_N(\rho_A(x_2)). \qquad (A.8)$$

Inserting (A.7) into (A.8) yields

$$\langle \delta_N(\rho_A) \rangle = \int dx_2 \int dx_1 \int dp_1 \left| \int dp_2 \, W(x_1, p_1, x_2, p_2) \right| - 1$$
$$\leq \int dx_2 \int dx_1 \int dp_1 \int dp_2 \, |W(x_1, p_1, x_2, p_2)| - 1 = \delta_N(\rho_{AB}) \quad (A.9)$$

which concludes the proof. Note that this statement is only true on average. For some measurement outcomes, δ_N does increase.

A.3 Distance to states with positive Wigner function

The measure of non-classicality δ_N defined above unfortunately lacks an operational meaning. To cure this problem, we define yet another non-classicality

measure by means of the trace distance to the set of states with positiv Wigner function:

$$N(\rho) = \min_{\sigma} \frac{1}{2}\|\rho - \sigma\|_1, \tag{A.10}$$

where the minimization runs over all states with positiv Wigner-function ($W_\rho > 0$). This measure has nice properties too:

Lemma 8. The negativity N is invariant under Gaussian unitary operations.

Proof: We observe that Gaussian unitaries conserve the positivity of the Wigner function. Denoting the optimal reference state for ρ by ρ^+, we get

$$N(U\rho U^\dagger) \leq \frac{1}{2}\|U\rho U^\dagger - U\rho^+ U^\dagger\|_1 = \frac{1}{2}\|\rho - \rho^+\|_1 \tag{A.11}$$

This shows that N cannot increase. However, if it could decrease, the inverse operation U^{-1} would increase the negativity of the state $\rho' = U\rho U^\dagger$. As this is not possible, N must stay invariant.

Lemma 9. The negativity N does not increase when tracing out a subsystem.

Proof: Here we only need the fact that the partial trace operation also preserves Wigner function positivity and get

$$N(\text{Tr}_A\rho_{AB}) \leq \frac{1}{2}\|\text{Tr}_A\rho_{AB} - \text{Tr}_A\rho_{AB}^+\|_1 \leq N(\rho_{AB}), \tag{A.12}$$

where we have also used that the trace distance is non-increasing under partial trace.

Lemma 10. The negativity N is convex.

Proof: We use the convexity of the set of states with positive Wigner function and get

$$N(p_1\rho_1 + p_2\rho_2) \leq \frac{1}{2}\|(p_1\rho_1 + p_2\rho_2) - (p_1\rho_1^+ + p_1\rho_2^+)\|_1$$
$$\leq \frac{1}{2}p_1\|\rho_1 - \rho_1^+\|_1 + \frac{1}{2}p_2\|\rho_2 - \rho_2^+\|_1 = p_1 N(\rho_1) + p_2 N(\rho_2). \tag{A.13}$$

Lemma 11. The negativity N does not increase when performing a tensor-product with a positive state.

Proof: We pick a possible reference state to get a lower bound

$$N(\rho_1 \otimes \rho_2^+) \leq \frac{1}{2}\|(\rho_1 \otimes \rho_2^+) - (\rho_1^+ \otimes \rho_2^+)\|_1 = \frac{1}{2}\|\rho_1 - \rho_1^+\|_1\|\rho_2^+\|_1 = N(\rho_1). \tag{A.14}$$

Note that we can only proof an inequality here because the optimal reference state of the tensor product does not need to be the tensor product of the optimal reference states of the factors.

Lemma 12. The negativity N is sub-additive under tensor products.

Proof: We take Lemma 11 and procede exactly as in the proof of Lemma 5.

Those lemmata show together that the negativity can only decrease under Gaussian channels, which can be also shown directly using the contractive property of the trace distance. In addition, we have:

Lemma 13. The negativity N does not increase on average under Gaussian measurements.

Proof: We denote the (non trace-preserving) quantum operation corresponding to the Gaussian measurement by $\mathcal{E}_k$ where k denotes the measurement outcomes. This operation preserves positivity, i.e., $W_{\mathcal{E}_k(\rho^+)} > 0$ and, therefore, $\mathcal{E}_k(\rho^+)$ is a valid possible reference state for remaining the system after the measurement. This yields

$$\sum_k N(\mathcal{E}_k(\rho)) \leq \sum_k \|\mathcal{E}_k(\rho) - \mathcal{E}_k(\rho^+)\|_1 = \sum_k \|\mathcal{E}_k(\rho - \rho^+)\|_1. \qquad (A.15)$$

The trace distance can not increase on average under any quantum operation. Thus, we obtain

$$\sum_k N(\mathcal{E}_k(\rho)) \leq N(\rho) \qquad (A.16)$$

which completes the proof. Again, the statement is only true on average as for some measurement outcomes, the negativity can increase.

A reason why negativity is a better description of the resource character of non-Gaussian states than the non-Gaussianities discussed in the literature and above is that convex mixtures of Gaussian states can be non-Gaussian. However, they are, for example, not useful for teleportation as the following lemma shows:

Lemma 14. Non-trivial entanglement distillation with states which are mixtures of Gaussian states is impossible under Gaussian operation.

Proof: Suppose the states to be distilled are of the form $\rho = \sum_k p_k \rho_k$, where the p_k are positive and sum to one and all ρ_k are Gaussian. An entanglement distillation protocol is a (non-trace preserving) quantum operation. The input state $\rho^{\otimes N}$ is also a convex mixtures of bipartite Gaussian states, i.e., $\rho^{\otimes N} = \sum_j q_j \sigma_j$ where q_j and σ_j can be easily obtained from p_j and ρ_j. Applying now the protocoll, we can use linearity to get

$$\tau := \mathcal{E}(\rho^{\otimes N}) = \sum_j q_j \mathcal{E}(\sigma_j). \qquad (A.17)$$

Now all σ_j are Gaussian states which are no more entangled than ρ_k due to convexity of entanglement measures. Using now the no-go theorem from Ref. [19], we get $E(\mathcal{E}(\sigma_j)) \leq E(\rho_k)$ where $E(\rho)$ is an entanglement measure. Using convexity again we finally get $E(\tau) \leq \max_k E(\rho_k)$. This means, it is not possible to distill a state which is more entangled than the most entangled state forming the convex mixture ρ. It is possible to get a state which is more entangled than the mixture ρ itself but this is not surprising. With the same technique, one can also show similar results for other quantum information tasks where Gaussian no-go theorems exist.

Not all states with positive Wigner functions are convex combinations of Gaussian states. One example is the single-photon state which has undergone a sufficient amount of photon loss, i.e., $\rho = (1 - \eta)|1\rangle\langle 1| + \eta|0\rangle\langle 0|$ for $\eta > 1/2$. The general question of deciding whether a state can be written as a mixture of Gaussian states can be translated to the level of the Wigner function. Denoting a Gaussian centered around ζ with the matrix M^{-1} as done in Ref. [45] by

$$G(\xi; M, \zeta) = \frac{\sqrt{\det M}}{\pi^n} \exp\left[-(\xi - \zeta)^T M (\xi - \zeta)\right], \qquad (A.18)$$

we ask if we can write

$$W_\rho(\xi) = \int d\zeta \, p(\zeta) G(\xi; M(\zeta), \zeta) \qquad (A.19)$$

with a positive normalized probability distribution $p(\zeta)$. For constant $M(\zeta)$ this is just a folding integral which can be converted to a product when making the transition to Fourier space. This renders it possible to calculate $p(\zeta)$ and to check whether it is positive. In the general case, however, on has to solve an integral equation for $p(\zeta)$ and $M(\zeta)$ which might be complicated.

It is important to note that negativity of the Wigner function is not necessary for entanglement distillation as the following example shows [19]: Let the density matrix elements be $\rho_{0,0,0,0} = 1/(1 + \epsilon^2)$, $\rho_{1,1,0,0} = \rho_{0,0,1,1} = \epsilon/(2 + 2\epsilon)$, $\rho_{1,1,1,1} = \epsilon^2/(1 + \epsilon^2)$ with zeros everywhere else. This state allows for entanglement distillation with Gaussian operation for all $\epsilon > 0$ but its Wigner function is positive for small enough ϵ. This means that to investigate the resource character of non-Gaussian state, one should define a measure which we will call the *essential* non-Gaussianity

$$N_G(\rho) = \min_\sigma \frac{1}{2} \|\rho - \sigma\|_1, \qquad (A.20)$$

where the minimization runs over all states which can be written as a convex combination of Gaussian states, i.e., whose Wigner functions are of the form

(A.19). This measure is bounded by the trace distance non-Gaussianity and the Wigner function negativity by

$$N(\rho) \leq 2N_G(\rho) \leq 2\delta(\rho). \qquad (A.21)$$

Furthermore, $N_G(\rho)$ fulfills all Lemmata shown for $N(\rho)$ as can be easily shown. The explicit calculation of $N_G(\rho)$ is difficult, but bounds can be calculated using (A.21).

A.4 Summary and Conclusion

In this appendix, we have introduced various measures which allow to quantify non-Gaussianity and non-classicality. By measuring the distance to a reference state by means of the trace norm, the non-Gaussianity δ obtains an operational meaning. However, states with large δ can be created by classical mixing of Gaussian states, and such states are not useful for entanglement distillation and other protocols where non-Gaussian states are necessary. For this reason, we have introduced the essential non-Gaussianity N_G which is defined as the minimal trace distance to the convex hull of Gaussian states. This quantity is a good candidate for measuring the resource character of non-Gaussian states but its use is limited as it is not easily calculable.

Regions where the Wigner function are negative are a clear indication for the non-classicality of the corresponding state [17]. Nevertheless, we have shown that negative Wigner functions are by no means necessary to circumvent the no-go theorems, e.g. for entanglement distillation with Gaussian operations. Thus, non-classicality measures which are based on the negativity of the Wigner function, as δ_N defined in (A.6) and N defined in (A.10) are not suited to measure the capability of a state in entanglement distillation with Gaussian operations and related protocols.

Appendix B

Ultra-cold atoms in optical lattices beyond the single-band approximation

In this appendix, we investigate the behavior of Bosons and Fermions in optical lattices. We start with solving the Schödinger equation for a single particle in a periodic potential. Then, we go the picture of second quantization and derive a Hamiltonian describing the mixture of interacting spinless Bosons and Fermions and numerically calculated the various hopping and interaction parameters characterizing the systems as well as the Wannier functions of the lowest few bands. This allows us to assess the single-band approximation commonly employed when discussing such systems [46]. Then, we solve the Hamiltonian, we have derived, numerically in the limit of a deep lattice, i.e., where hopping between the lattice sites is absent. We calculate the occupation of the higher bands depending on the interaction and the lattice depths.

B.1   Non interacting Bosons and Fermions in optical lattices

We consider a single particle in three dimensions with mass m moving in a potential $U(\mathbf{r})$ which is periodic with lattice constant a. The single-particle Hamiltonian reads

$$\hat{h} = -\frac{\hbar^2\nabla^2}{2m} + U(\mathbf{r}).\tag{B.1}$$

For mathematical convenience and for later numerical calculations, we consider a homogeneous finite D-dimensional cube $[-aN/2, aN/2)^D =: \mathcal{D}$ with periodic

boundary conditions. Furthermore, we introduce the lattice wave vector $k_L = \pi/a$ and the recoil energy $E_R = \hbar^2 k_L^2/(2m)$. The latter will be used as a unit of energy whenever it is possible. From Bloch's theorem we know that the solutions of the Schrödinger equation are of the form

$$\psi_{\mathbf{x},\mathbf{n}}(\mathbf{r}) = [1/(aN)^{D/2}]\mathrm{e}^{2ik_L\mathbf{x}\cdot\mathbf{r}/N}u_{\mathbf{x},\mathbf{n}}(\mathbf{r}), \qquad (B.2)$$

where $\mathbf{x} \in \{-N/2, \ldots, N/2 - 1\}^{\times D} =: \mathcal{B}$. The Bloch functions $u_{\mathbf{x},\mathbf{n}}(\mathbf{r})$ have period a and are labeled by the band index $\mathbf{n}$. The corresponding energy eigenvalues will be denoted by $E_{\mathbf{x},\mathbf{n}}$ and are shown in Fig. B.1.

Now, we can define the Wannier functions for $\mathbf{z} \in \mathcal{B}$ as

$$w_{\mathbf{z},\mathbf{n}}(\mathbf{r}) = \frac{1}{N^{D/2}} \sum_{\mathbf{x}\in\mathcal{B}} \mathrm{e}^{-2\pi i\mathbf{x}\cdot\mathbf{z}/N}\psi_{\mathbf{x},\mathbf{n}}(\mathbf{r}). \qquad (B.3)$$

There exists a notable freedom in the definition of the Wannier functions, i.e., the choice of the phases in (B.3). However, Ref. [47] provides a rule to choose the phases such that the resulting Wannier functions are real, symmetric (for even bands) or antisymmetric (for odd bands), and maximally localized. Because the lattice potential depends on the intensity of the light-field we take the one-dimensional potential $U(r) = V_0 \sin^2(k_L r)$. The one-dimensional Wannier functions for the first band is shown for different lattice depths in Figs. B.2. In the limit of an infinitely deep lattice, the Wannier functions converge to the eigenfunctions of the harmonic oscillator. For a simple cubic lattice, and this is the situation we want to consider in this appendix, the Wannier functions factorize and we can write

$$w_{\mathbf{z},\mathbf{n}}(\mathbf{r}) = w_{\mathbf{z},n_1}(x)_{\mathbf{z},n_2}(y)w_{\mathbf{z},n_3}(z). \qquad (B.4)$$

Thus, in a three-dimensional system all bands but the first one are degenerate. For example, there exist three first excited bands, i.e., $\mathbf{n}_x = (1,0,0)$, $\mathbf{n}_y = (0,1,0)$, and $\mathbf{n}_z = (0,0,1)$. However, because of the factoring property (B.4), it is in most of the cases enough to consider the one-dimensional situation as we will do in the following. The Wannier functions of the first four bands are shown in Fig. B.3.

B.2 Interacting atoms

In order to consider the interaction between the particles, we turn to the second-quantized formulation of the problem. The interaction between ultra-cold atomic gases is very well described by s-wave scattering with a delta-shaped pseudo-potential [11]. Therefore, all interactions are characterized
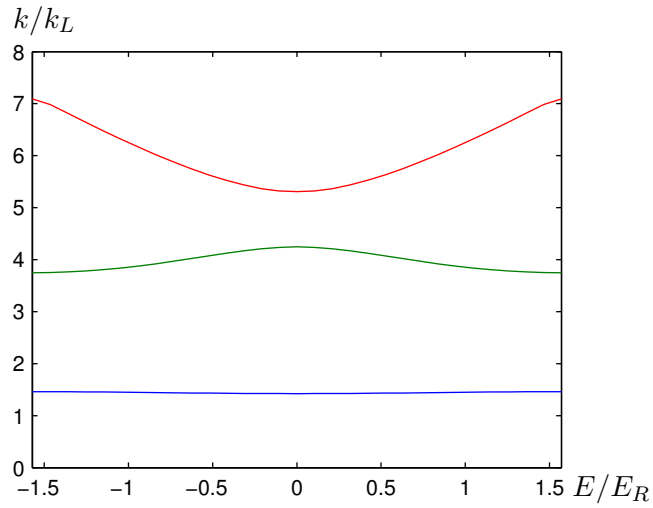
Figure B.1: Lowest three energy bands of a one-dimesional system for $V_0 = 10\,E_R$.
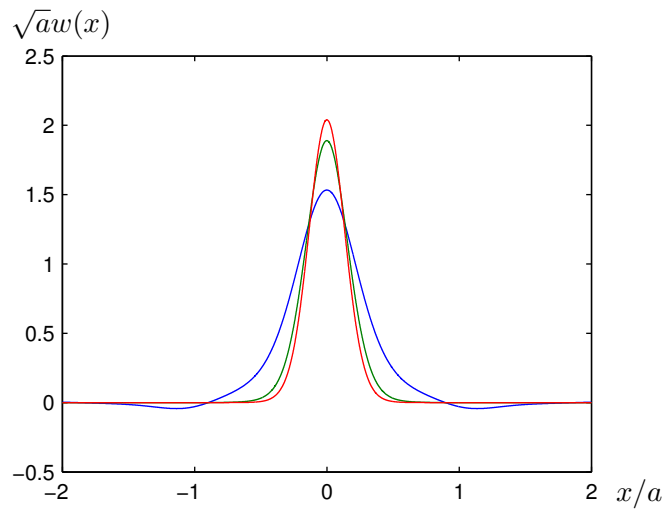

Figure B.2: One-dimensional Wannier function of the lowest band for $V_0 = 5\,E_R$ (blue), $V_0 = 20\,E_R$ (red), and $V_0 = 30\,E_R$ (green).
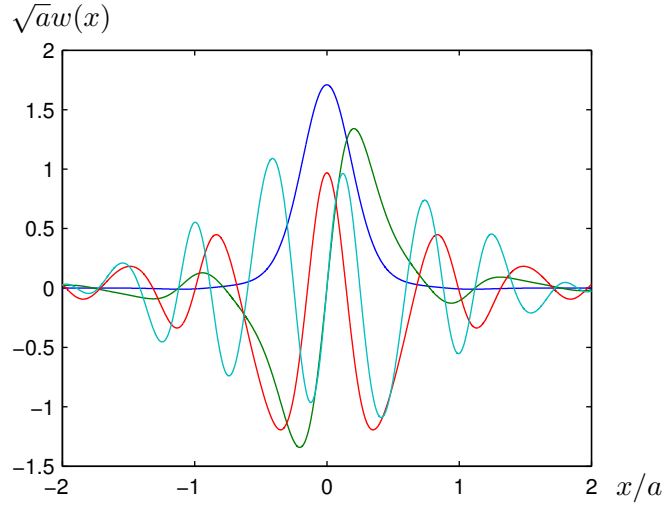
Chapter B - Ultra-cold atoms in optical lattices beyond the
114
single-band approximation

Figure B.3: One-dimensional Wannier function of the first four bands (green, red, blue, pink) for $V_0 = 19\,E_R$.

by only two parameters, i.e., the scattering length for Bose-Bose scattering (a_{BB}) and the one for Bose-Fermi scattering (a_{BF}). In a second quantized description, the Hamiltonian reads

$$\hat{H} = \hat{H}_B + \hat{H}_F + \hat{H}_{BB} + \hat{H}_{BF}, \tag{B.5}$$

with

$$\hat{H}_{B,F} = \int d\mathbf{r}\, \hat{\Psi}^\dagger_{B,F}(\mathbf{r})\hat{h}_{B,F}(\mathbf{r})\hat{\Psi}_{B,F}(\mathbf{r}), \tag{B.6}$$

$$\hat{H}_{BB} = \frac{g}{2} \int d\mathbf{r}\, \hat{\Psi}^\dagger_B(\mathbf{r})\hat{\Psi}^\dagger_B(\mathbf{r})\hat{\Psi}_B(\mathbf{r})\hat{\Psi}_B(\mathbf{r})\,, \tag{B.7}$$

$$\hat{H}_{BF} = f \int d\mathbf{r}\, \hat{\Psi}^\dagger_B(\mathbf{r})\hat{\Psi}^\dagger_F(\mathbf{r})\hat{\Psi}_B(\mathbf{r})\hat{\Psi}_F(\mathbf{r}), \tag{B.8}$$

where Ψ_B (Ψ_F) is the Bosonic (Fermionic) annihilation operator and $g = (4\pi^2\hbar^2 a_{BB})/m_B$, $f = 4\pi^2\hbar^2 a_{BF}(m_B + m_F)/(2m_B m_F)$.

We now expand the field operators into Wannier functions

$$\hat{\Psi}_{B,F} = \sum_{\mathbf{i},n} w_{\mathbf{i},\mathbf{n}}(\mathbf{r})\hat{b}_{\mathbf{i},\mathbf{n}} \tag{B.9}$$

which yields $(\hat{a} = \hat{b}, \hat{f})$

$$\hat{H}_{B,F} = \sum_{\mathbf{i},\mathbf{j},\mathbf{n}} J^{B,F}_{\mathbf{i},\mathbf{j},\mathbf{n}} \hat{a}^\dagger_{\mathbf{i},\mathbf{n}}\hat{a}_{\mathbf{j},\mathbf{n}}, \tag{B.10}$$

where

$$J_{\mathbf{i,j,n}}^{B,F} = \frac{1}{N^D} \sum_{\mathbf{x} \in \mathcal{B}} e^{2\pi i \mathbf{x}(\mathbf{i}-\mathbf{j})/N} E_{\mathbf{x,n}}^{B/F}, \tag{B.11}$$

i.e., different energy bands are not coupled through hopping. The intra-species interaction Hamiltonian reads (we use super-indices $(\mathbf{i}, \mathbf{n}) = \mathbf{i}$)

$$\hat{H}_{BB} = \sum_{\substack{\mathbf{i,j}, \\ \mathbf{k,l}}} \frac{U_{\mathbf{i,j,k,l}}}{2} \hat{b}_{\mathbf{i}}^{\dagger} \hat{b}_{\mathbf{j}}^{\dagger} \hat{b}_{\mathbf{k}} \hat{b}_{\mathbf{l}}, \tag{B.12}$$

where $U_{\mathbf{i,j,k,l}} = gW_{\mathbf{i,j,k,l}}$ with

$$W_{\mathbf{i,j,k,l}} = g \int_{\mathcal{D}} d\mathbf{r}\, w_{\mathbf{i}}^*(\mathbf{r}) w_{\mathbf{j}}^*(\mathbf{r}) w_{\mathbf{k}}(\mathbf{r}) w_{\mathbf{l}}(\mathbf{r}). \tag{B.13}$$

We note that because we have chosen the Wannier functions to be real, a permutation of the indices does not chance the value of the interaction coefficients. Because of the factoring property (B.4), all these integrals factor into products of one-dimensional integrals over the respective one-dimensional Wannier functions. For all Wannier functions centered around the same lattice site, i.e., the ones occurring in the terms describing local interactions, they read

$$W_{i,j,k,l} = \int_{-\infty}^{\infty} dfr\, w_i^*(r) w_j^*(r) w_k(r) w_l(r), \tag{B.14}$$

where i, j, k, l are band indices. Because the even (odd) Wannier functions are even (odd), most of these coefficients, e.g. $W_{0,0,0,1}$, vanish. The inter-species interaction Hamiltonian reads in complete analogy

$$\hat{H}_{BF} = \sum_{\mathbf{i,j,k,l}} V_{\mathbf{i,j,k,l}} \hat{b}_{\mathbf{i}}^{\dagger} \hat{b}_{\mathbf{j}} \hat{f}_{\mathbf{k}}^{\dagger} \hat{f}_{\mathbf{l}}, \tag{B.15}$$

where $V_{\mathbf{i,j,k,l}} = fW_{\mathbf{i,j,k,l}}$. In Fig. B.6 coefficients corresponding to intra-band interaction of the Bosons are shown. They are larger in the lower bands because these bands are more localized and, therefore, the overlap of the wavefunction with itself is larger. The coefficients which couple two bands, i.e., the one associated with inter-band interaction and with transfer of the atoms from one band to another, are also smaller for higher bands. This fact, together with the band-gap, c.f. Fig. B.7, which imposes an energy penalty when occupying higher bands, justifies to consider only the lowest two or three bands. In addition, we note that the interaction between particles on neighboring sites is much smaller then every other interaction coefficient and will be neglected in the remainder of this appendix. The hopping parameters defined in (B.11) are larger in the higher bands, because they are less localized. Furthermore, comparing Figs. B.4 and B.5 one sees that the effect of next-nearest neighbor hopping is very weak and therefore not considered in the following calculations.
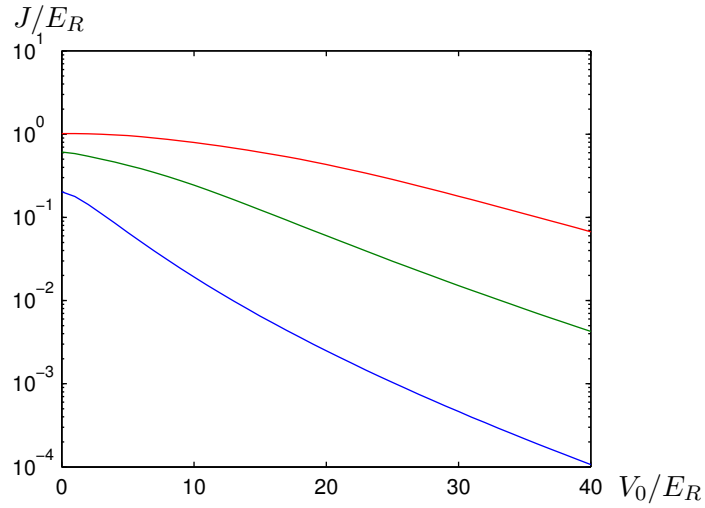
Figure B.4: Nearest-neighbor hopping parameter for the lowest (blue), first (red), and second (green) Bloch band.
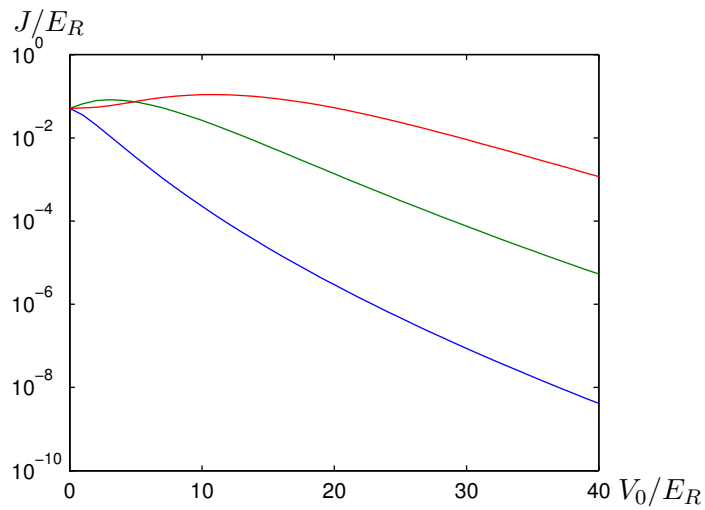

Figure B.5: Next-nearest neighbor hopping parameter for the lowest (blue), first (red), and second (green) Bloch band.
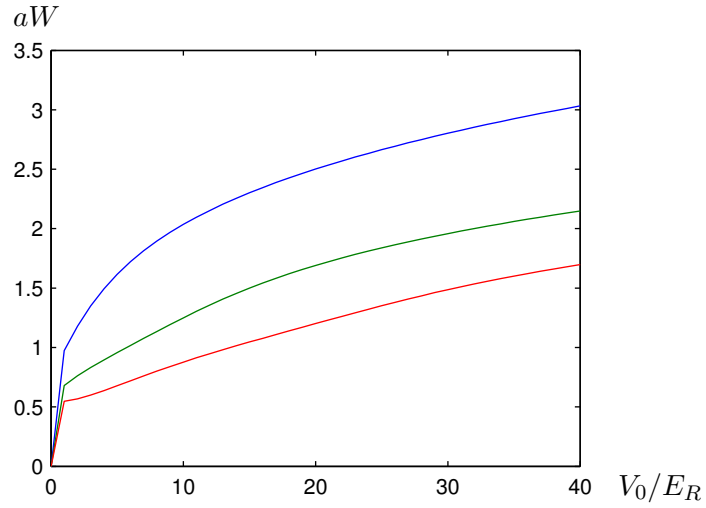
Figure B.6: On-site intra-band interaction coefficients $W_{0,0,0,0}$ (blue), $W_{1,1,1,1}$ (red), $W_{1,1,1,1}$(green).
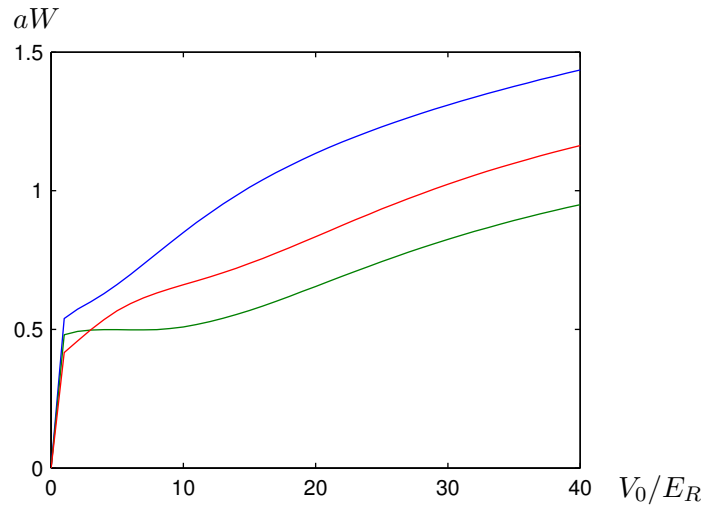

Figure B.7: On-site inter-band coupling coefficients $W_{0,0,1,1}$ (blue), $W_{0,0,2,2}$ (red), $W_{1,1,2,2}$(green).

B.3   Single site system

We consider a single lattice-site which corresponds to setting all hopping parameters to zero. Due to the presence of the Boson-Fermion interaction and especially because of the multiple bloch bands, already the single-site physics is highly non-trivial.

   We concentrate on the situation of one Fermion and n Bosons on a single site and restrict ourselves to the one-dimensional case, which is very similar to the three-dimensional one due to the factorization properties of the Wannier functions. We diagonalize the resulting Hamiltonian numerically exact. As the simplest non-trivial example, we take the Hamiltonian truncated to the lowest and one first excited band which are seperated by the gap $\Delta_{B/F}$. With the abbreviations $U_p = U_{p,p,p,p}$, $V_p = V_{p,p,p,p}$ and $\mu_0 = \mu$, $\mu_1 = \mu + \Delta_B$, $\nu_0 = \nu$, $\nu_1 = \nu + \Delta_F$, it reads

$$\hat{H}^{(0)} = \sum_{p=0}^{1} \hat{H}_{i,p}^{(0)} + \hat{H}_{i,p}^{(0)} \tag{B.16}$$

with

$$\hat{H}_{i,p}^{(0)} = \frac{U_p}{2}\hat{n}_p(\hat{n}_p - 1) - \mu_p\hat{n} + V\hat{n}_p\hat{m}_p - \nu_p\hat{m}, \tag{B.17}$$

while the one describing inter-band interaction and coupling reads with $U_{01} = U_{0,0,1,1}$ and $V_{01} = V_{0,0,1,1}$,

$$
\begin{aligned}
\hat{H}_{i,01}^{(0)} = {}& 2U_{01}\hat{n}_0\hat{n}_1 + V_{01}\left(\hat{n}_0\hat{m}_1 + \hat{n}_1\hat{m}_0\right) \\
& + U_{01}\left(\hat{b}_0^\dagger\hat{b}_0^\dagger\hat{b}_1\hat{b}_1 + \hat{b}_1^\dagger\hat{b}_1^\dagger\hat{b}_0\hat{b}_0\right) + V_{01}\left(\hat{b}_0^\dagger\hat{b}_1\hat{f}_0^\dagger\hat{f}_1\right. \\
& \left. + \hat{b}_0^\dagger\hat{b}_1\hat{f}_1^\dagger\hat{f}_0 + \hat{b}_1^\dagger\hat{b}_0\hat{f}_0^\dagger\hat{f}_1 + \hat{b}_1^\dagger\hat{b}_0\hat{f}_1^\dagger\hat{f}_0\right).
\end{aligned} \tag{B.18}
$$

The dimension of the Hilbert space for n Bosons and m Fermions is $d_H = n + 1$ for ($m = 0$ or $m = 1$) and $d_H = 2(n+1)$ for $m = 1$ which is easily diagonalized for all reasonable occupation numbers. Its eigenstates and energies are denoted by

$$\hat{H}^{(0)}|n, m; l\rangle = \epsilon_{n,m;l}|n, m; l\rangle. \tag{B.19}$$

Now, we calculate the expectation values of both the Bosons and the Fermions in the different bands, i.e., $\langle\hat{n}_i\rangle$ and $\langle\hat{m}_i\rangle$ for $i = 0, \ldots, k$. Before we do this, we shortly discuss the physical meaning of the contributions to (B.18) which are proportional to U_{01}. Because, as already mentioned, $W_{0,0,0,1} = 0$, it is not possible for a single Boson to jump into the second band, which means that the influence of the second band is absent in the important special case of one Boson per lattice site. In contrast, a process where a single particle jumps to
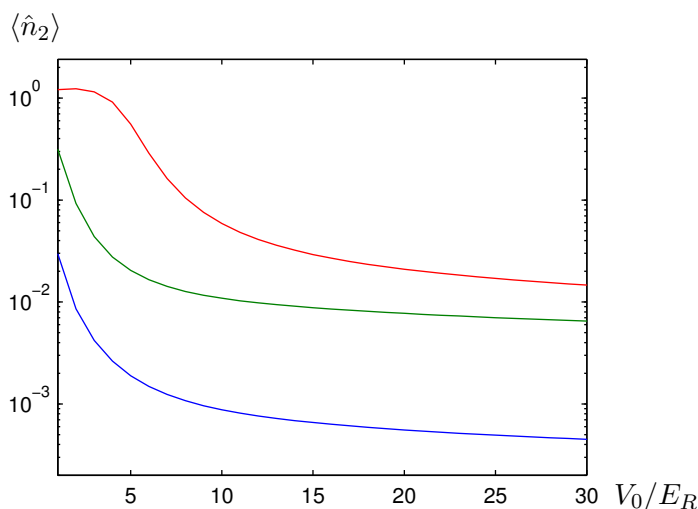
$\langle \hat{n}_2 \rangle$

Figure B.8: Occupation of the second band $\langle \hat{n}_2 \rangle$ for two Bosons and one Fermion on a single lattice site in one dimension without Bose-Fermi ineraction (blue), $a_{BF} = -400\,a_0$ (red), and $a_{BF} = +400\,a_0$ (green).

the *third* band, described by $\hat{b}_3^\dagger \hat{b}_1^\dagger \hat{b}_1 \hat{b}_1$ is possible. Even though its contribution is in generally not neglectable compared to the one of the second band, we, nevertheless, consider a two-band model for simplicity reason because we aim at understanding multiband effects with the simplest possible model.

From Fig. B.8, it is clearly visible that the occupation of the second band increases with increasing Bose-Fermi interaction and is larger for repulsive interaction. In order to be specific, we consider from now on a mixture of ^{87}Rb and ^{40}K in a lattice produced by a laser with a wave-length of 755nm [48]. In this experiment, the Bose-Bose scattering length was $a_{BB} = 100a_0$ while the Bose-Fermi scattering length could be changed form -800 a_0 to +800 a_0, where a_0 is the Bohr radius. Furthermore, the detuning of the lasers is chosen in such a way that the lattice potential for the two species are the same in their respective recoil energy, i.e. $\tilde{V} := V_0^B/E_R^B = V_0^F/E_R^F$.

The Wannier functions are a set of basis functiosn derived from a single particle wave function. Of course, they form a complete set of basis functions also for interacting particles, but in this situation they are not optimal anymore in that sense that only one them is necessary to fully describe the system. Instead of taking higher Bloch bands into account, one can also consider Wannier-functions which are modified by the interaction and then work with an effective single-band model with renormalized parameters. The key idea is the following: The presence of the Fermions changes the effective lattice

potential seen by the Bosons which yields, when approximating the Fermionic particle number operator by the Fermionic density,

$$V_{\text{eff}}^B(x) = V^B(x) + f\rho^F(x). \tag{B.20}$$

The same arguing is also true vise versa and the effective lattice potential for the Fermions reads

$$V_{\text{eff}}^F(x) = V^F(x) + nf\rho^B(x). \tag{B.21}$$

For an attractive Bose-Fermi interaction, the effective potentials are deeper than the bare ones which means that the hopping gets weaker and the on-site interaction stronger. Therefore, this effect, which was also found in [48] with more sophisticated methods, is often called "self-trapping". Starting from the non-interacting wave function, we can calculate the Wannier functions in a self-consistent way which converges after a few steps. From these Wannier functions we determine renormalized hopping ($\tilde{J}$) and interaction parameters ($\tilde{U}$) for the Bosons. Then, the Hamiltonian, restricted to the lowest band, is the well-known and extensively studied Bose-Hubbard Hamiltonian

$$\hat{H} = \frac{\tilde{U}}{2}\sum_i \hat{n}(\hat{n}-1) - \tilde{J}\sum_{<i,j>} \hat{b}_i^\dagger \hat{b}_j. \tag{B.22}$$

The whole effect of the Fermions is now only incorporated in the modification of $\tilde{J}$ and $\tilde{U}$. Because this ratio becomes lower, as seen in Fig. B.9, the superfluid-Mott insulator transition is shifted to lower values of the lattice depth which agrees with experimental findings. However, for repulsive Bose-Fermi interaction, the effective potential is shallower than the bare one which leads to anti-trapping and a shift of the quantum phase transition into the other direction. This contradicts the experimental results and, therefore, other explanations are needed in this parameter regime

B.4   Summary and Outlook

We have derived the Bose-Fermi Hubbard Hamiltonians for a contact interaction in the most general form. We have seen that even though the occupation of the higher bands is small for the interaction strengths realized in present day experiments, they still can play an important role for the physics of such systems. This is due to the fact that the atoms in these higher bands are less localized resulting in a higher hopping constant, i.e., their mobility is larger.
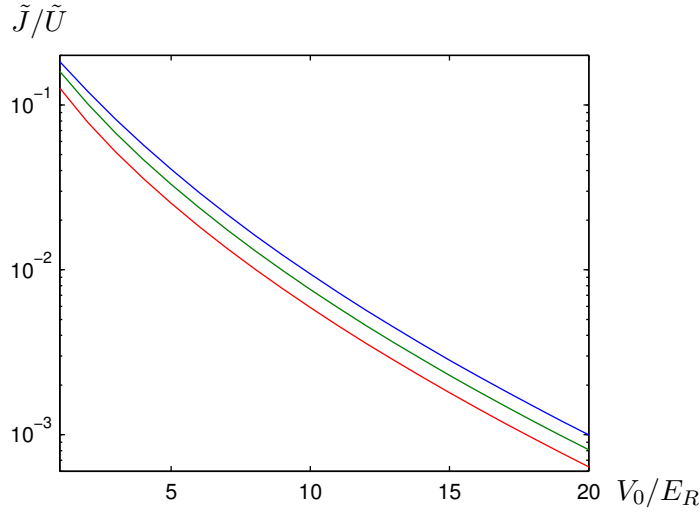
$\tilde{J}/\tilde{U}$

Figure B.9: Ratio of renormalized (see text) ratio of hopping and interaction coefficient in the lowest band for one Fermion and two Bosons per site without Bose-Fermi interaction (blue) for $a_{BF} = -400\,a_0$ (red), and $a_{BF} = +400\,a_0$ (green).

A next step should be to use the solution of the single-site systems as a starting point to address the full many-body system. One possible route is to use perturbative techniques, as used, e.g. in Ref. [49] for the single-band Bosonic model. A second promising idea is to treat the Bosons numerically by a Gutzwiller ansatz while solving the Fermionic subsystem exactly which is possible as spin-less Fermions do not interact with each other which is other due to the Pauli principle. This approach has also already been demonstrated for the single-band Bose-Fermi Hubbard model [46].