Freie Universität Berlin
Dahlem Center for Complex Quantum Systems

Complex quantum systems from a signal processing perspective

Dissertation

zur Erlangung des Grades eines Doktors der Naturwissenschaften
am Fachbereich Physik der Freien Universität Berlin

vorgelegt von

Adrian Steffens

Berlin, 2017

Supervisor	Erstgutachter	Prof. Dr. Jens Eisert, Freie Universität Berlin
Second reviewer	Zweitgutachter	Prof. Dr. Christoph Karrasch, Freie Universität Berlin
Date of defense	Tag der Disputation	11.03.2019

Meinen Eltern

Juliane und Sigurd Steffens

CONTENTS

SUMMARY OF PUBLICATIONS

The present dissertation is cumulative according to the doctorate rules of the department of physics of the Free University of Berlin and is based on the following first-authored publications.

Closely related to these are the following coauthored publications, which are provided in the appendix.

Ref. [3] was developed in collaboration with the Schmiedmayer group at the Vienna University of Technology, Austria, Ref. [1] in collaboration with the Quantum Engineering Technology Labs at the University of Bristol, UK, and Refs. [4, 6] were established during the author's research visit at the Research Laboratory of Electronics at the Massachusetts Institute of Technology, Cambridge, MA, USA. Not yet published is project [7], see p. 141, which resulted from the author's work as part of the theory group of the Fritz Haber Institute of the Max Planck Society, Berlin.

1 INTRODUCTION

Since its inception at the beginning of the 19th century, quantum mechanics has been fundamental in understanding the physical principles of nature, explaining phenomena from black body radiation and the photoelectric effect to superconductivity. This has made it possible to use quantum effects in technical applications such as solar cells or lasers. With the technical progress of the past years, the ability to actively control and manipulate complex quantum systems has come into reach, leading to a *second quantum revolution* [8] with entirely new technological prospects: *Quantum computers* are expected to substantially impact high-end computing, *quantum cryptography* protocols may provide intrinsically secure data transmission, and *quantum metrology* could allow for ultra-precise sensors. Commercial interest is growing, and a multi-billion Euro market is predicted to emerge.[9]

Quantum systems, however, are very fragile and require thorough analysis if they are to be effectively harnessed for quantum technologies. Accurately modeling quantum systems is necessary for explaining and predicting their behavior as building blocks of quantum devices. For a complete model describing a quantum system—a *quantum state*—the system has to be prepared multiple times and each time a certain *quantum measurement* has to be performed, resulting in data that can be processed to infer the model. This procedure is called *quantum state tomography* and will be a fundamental topic in the sections 2 and 3. With increasing size and complexity of the quantum system, this task soon becomes infeasible in the general case. This is due to an exponential increase of the number of modeling parameters and necessary measurements with the size of the quantum system, making it hard to find appropriate parameters ("curse of dimensionality"[10]). To cope with this problem, elaborate *signal processing* techniques have to be developed and employed.

Signal processing comprises the transmission, manipulation, and representation of *signals* and their underlying *information*. It is fundamental to a wide range of fields such as audio, image, and video processing, telecommunication, or seismology. Signals can be continuous functions in time or space such as voltages, magnetic fields, sound waves, angles, forces, or temperatures. For processing, especially digital signal processing, such signals are typically discretized, i.e., sampled at discrete points. Higher sampling rates generally result in higher accuracy in describing the input and the output quantities of interest, but then require more computational resources, especially computational time and memory. With finite resources, this calls for carefully weighing effort and quality of the output.

In the context of complex quantum systems, the basic input signals consist of quantum measurements to estimate the parameters that constitute the corresponding quantum state. In general, it is desirable to restrict the parameter set in such a way that the properties of the quantum system are essentially preserved while keeping the processing effort manageable. Finding such "compressed models" is not trivial and requires a certain intrinsic compressibility of the parameter space. And yet it is not uncommon: Often in the history of physics it was vital to discard unnecessary properties

of a system to properly model particular parts of nature, such as observing the free fall in a vacuum without drag forces. In the case of complex quantum systems, it turns out that a large part of the entire parameter space is occupied by highly *entangled* quantum states, which are unlikely to appear in large parts of nature. In many cases, the remaining states, the ones of actual interest, can be efficiently parametrized by use of *tensor network states*. They constitute a seminal class of quantum states with a plethora of groundbreaking applications, allowing the development of potent quantum state tomography protocols (see the publications [2] and [3] in section 3). For this, a crucial reconstruction step relies on a class of signal processing algorithms for *spectral estimation*, which provide frequency spectra for certain signals with high accuracy. Beyond quantum state tomography, in this thesis *tensor network states* are employed for succinctly describing physical processes such as quantum transport experiments (see publication [5] in appendix A). The established protocol facilitates the study of the short-time dynamics of these quantum systems by providing statistics that are not directly accessible otherwise.

If, on the other hand, one is interested in a very general and robust procedure with very few assumptions about the quantum state to be determined, which comes at the price of less total compressibility of the parameter space, one can make use of another famous signal processing paradigm: *compressed sensing*, which allows for the extraction of *sparse* solutions out of a higher-dimensional parameter space using only a comparably small number of measurements/samples (section 2). With this, signals can be recorded in a compressed way, i.e. substantially fewer input samples are required to recover the underlying model. *Compressed sensing* has vastly spread in recent years, making huge progress both in establishing provable theoretical recovery guarantees and practical applications in all kinds of fields ranging from machine learning to radar communication. E.g., in medical imaging, the resolution of the reconstructed images could be considerably increased while keeping scanning times fixed.

When applying *compressed sensing* to practical *quantum state tomography* with noisy measurement data, external tuning parameters arise that strongly influence the size of the model of the system and one has to trade off model complexity with fitting quality. A maximally accurate fit of the measurement realization at hand might seem like the best choice, however, this could also lead to fitting features of the input signal that are just due to noise, resulting in models that are not general enough to provide sound predictions for future measurements. Using the example of a photonic quantum experiment, we explore different methods for overcoming this issue, providing a toolbox for reliable quantum state reconstruction from compressed measurement data sets (see publication [1]). In principle, employing *compressed sensing* methods does not exclude the use of *tensor network* methods—in fact, it can be desirable to profit from both if the structure at hand permits it.

The power of *compressed sensing* to establish sparse representations can also be of advantage in quantum chemistry: In *density functional theory*, the electronic structure of complex quantum systems such as molecules or crystals is expressed in terms of sets of certain basis functions. It is often not clear, however, if a smaller basis *subset* still preserves the essential physics. Using *compressed sensing* methods, compressed basis sets and hence more compressed models of the quantum systems

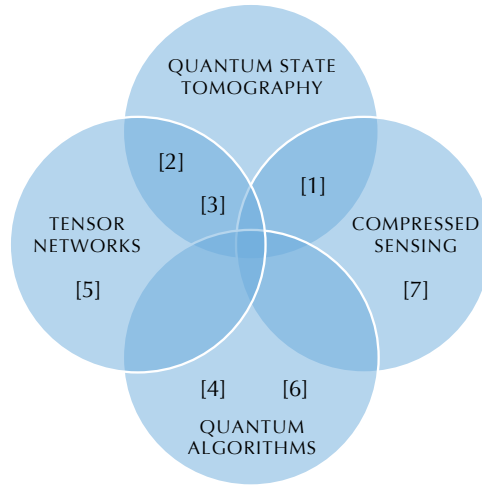

Figure 1: *Complex quantum systems from a signal processing perspective: Covered topics of the publications "Experimentally exploring compressed sensing quantum tomography"[1], "Quantum field tomography"[2], "Towards experimental quantum-field tomography with ultracold atoms"[3], "An efficient quantum algorithm for spectral estimation"[4], "Continuous matrix product state tomography of quantum transport experiments"[5], "Quantum singular value decomposition of non-sparse low-rank matrices"[6], and the project "Compressive density functional theory"[7].*

can be constructed, leading to reduced computational effort (see appendix A.3).

Signal processing methods can be used to productively analyze complex quantum systems. Interestingly, this can also be turned around by using a quantum computer—essentially a complex quantum system—to massively, in various instances even exponentially, accelerate signal processing algorithms, widening the range of problems that can be handled efficiently both for quantum and classical applications (section 4). An exponential acceleration would mean that classical computations that would take years to be carried out—even on a current supercomputer—could take only minutes on a quantum computer. Instead of bits, quantum computers operate on quantum bits (qubits), which can be in a superposition of both "0" and "1" and can be entangled with other qubits. The complexity of the arising system, which, as noted above, is a curse for quantum state reconstruction, can be a blessing for quantum computing. Working with n qubits essentially allows to simultaneously manipulate 2^n numbers, giving rise to a possibly exponential speedup for algorithms. The downside is that in general retrieving the information, i.e., estimating the output quantum state, is forbiddingly time-consuming, potentially undoing the achieved speedup. Quantum algorithms therefore have to be carefully designed such that the output quantities are accessible in an efficient manner. Also for this reason, the extent of the computational advantage of quantum computers still needs to be determined.

Considering that classical semiconductor-based computer technologies are reaching their physical limits, quantum computing promises great benefits for data processing. Huge efforts are being made to realize physical implementations ranging from trapped ions and superconducting electronic circuits to photonic systems that are both robust towards noise and scalable in the number of qubits. Due to the exponential acceleration, already a double-digit number of qubits could provide instances of a tangible quantum advantage.

In publication [4] in section 4, a quantum algorithm for the spectral estimation of signals is presented, providing an exponential speedup in comparison with classical algorithms. A key component for this algorithm is the development of a quantum version of the *singular value decomposition* (see publication [6] in appendix A), an important linear-algebra method with numerous applications.

Figure 1 provides a graphical overview of the topics that are covered by the publications constituting this dissertation and combining signal processing with complex quantum systems. A brief introduction to quantum state tomography is provided in section 2.1, an introduction to compressed sensing in section 2.2, to tensor networks in section 3.1, and to quantum algorithms in section 4. The publications [1–6] are presented in the respective sections and highlighted with gray frames.

Beyond signal processing and the thematic intersections of quantum tomography with tensor networks and compressed sensing, there is further overlap among the respective projects: The same algorithm that gains a quantum speedup in publication [4] constitutes an important reconstruction step in publication [2]; the publications [1], [2], [3], and [5] are concerned with *inverse problems*[II]—inferring the underlying parameters that determine a given signal. Publication [5] also extends the reconstruction protocol in [2] to a new class of input signals.

2 COMPRESSED SENSING QUANTUM TOMOGRAPHY

The *compressed sensing* paradigm comprises several signal processing techniques that are essential for estimating quantum states—i.e., performing *quantum state tomography*—of larger quantum systems, where it is difficult or virtually impossible to obtain sufficiently large sets of measurement data that would be necessary for ordinary methods to work. In the following, we provide the reader with a brief introduction to *quantum state tomography* with an emphasis on its scaling behavior for increasing size of the quantum system. This calls for signal processing methods from *compressed sensing* and *spectral estimation* as well as incorporation of a priori physical information via *tensor network* methods to manage the arising amount of data. *Tensor networks* will be introduced in section 3.1. In the subsequent section 2.2, the principles of *compressed sensing* are summarized and in section 2.3, publication [1][1] is presented, which focuses on the application of *compressed sensing* on an experimentally realized photonic quantum system The concepts in section 2.2 are also used in the project in appendix A.3.

2.1 Quantum state tomography and the curse of dimensionality

Quantum state tomography is concerned with the appropriate modeling of a quantum system by assigning a *quantum state* to it—commonly in the form of a Hilbert space vector or wave function, a density matrix/operator, or a *Wigner* quasiprobability distribution. An intrinsic challenge is that in general, together with their immanent uncertainty, quantum states cannot be determined with just one measurement, and the measurement changes the state itself: Measuring the position of one quantum particle may yield one position at a time, but not its underlying wave function, which describes the particle as well as where and when it will be found with which probability. We rather have to consider a set of *identically prepared* particles and subsequently perform a measurement on each particle. The information from all measurement outcomes is used to produce the quantum state—like two-dimensional image *slices* (ancient Greek: τομή) are combined to create a three-dimensional tomogram of a spatially extended object as in medical imaging. Hence the name *quantum state tomography*. The quantum state provides the necessary information to make statistical predictions about the behavior of the particle. The origins of quantum state tomography are found in quantum optics, estimating the state of a photonic system[12]. See Refs. [13, 14] for a comprehensive overview.

Quantum *state* tomography is closely related to and the starting point for quantum *process* tomography[15], the estimation of processes that represent the evolution in time of an initial quantum state and output the evolved state. For this, the process is applied to a set of known trial states, and

[1] Adrian Steffens, Carlos A. Riofrío, Will McCutcheon, Ingo Roth, Bryn A. Bell, Alex McMillan, Mark S. Tame, John G. Rarity, and Jens Eisert, "Experimentally exploring compressed sensing quantum tomography", *Quantum Science and Technology* 2:025005, 2017 (DOI:10.1088/2058-9565/aa6ae2). Published under a Creative Commons Attribution 3.0 License (creativecommons.org/licenses/by/3.0), © 2017 IOP Publishing. Reproduced with permission. All rights reserved.

the corresponding output states are estimated via state tomography. The quantum process parameters can be inferred from the change in the state parameters.

Quantum tomography is essential not only for a thorough understanding of many quantum experiments[16–19], but also for the fundamental operations and components of future quantum computers (see the quantum algorithm section 4) and other quantum devices such as quantum simulators[20,21]. These devices are described in terms of quantum states and processes, which makes the estimation of the latter indispensable for proper operation.

A *pure* quantum state $|\varphi\rangle$ of finite dimension d can be represented by a vector v in $\mathbb{C}^d$. If the state is *mixed*—i.e., it is not completely known in which pure state the system is—we describe it by use of a *density matrix* ϱ. One can think of it as a mixture of pure state vectors $v_1, \ldots, v_r$ appearing with respective classical probabilities $0 \leq w_1, \ldots, w_r \leq 1$ and summed up as weighted outer products: $\varrho = \sum_{j=1}^{r} w_j \, v_j v_j^\dagger$.[2] Note, however, that the representation in terms of pure states is not unique—see, e.g., Ref. [22]. By construction, the density matrix is Hermitian—$\varrho \in \mathscr{H}_d \subset \mathbb{C}^{d\times d}$—and because of the values w_j being probabilities adding up to one, it is positive semidefinite ($\varrho \succeq 0$, all eigenvalues are greater than or equal to zero) and has unit trace. These properties are sufficient to characterize the set $\mathscr{S}_d$ of density matrices of dimension d:

$$\mathscr{S}_d = \{\varrho \in \mathscr{H}_d : \varrho \succeq 0, \mathrm{tr}(\varrho) = 1\}. \tag{1}$$

This is a convex set, with the pure states, matrices of rank $r = 1$, constituting its boundary. This will be important for *convex optimization* methods, which are used in compressed sensing, as will be discussed in the next section 2.2.

Quantum measurements can be modeled by a set of measurement operators $A_1, \ldots, A_m \in \mathscr{H}_d$, where each measurement operator A_j corresponds to a measurement outcome that happens with probability

$$p_j = \mathrm{tr}(A_j^\dagger A_j \varrho) \tag{2}$$

and results in a post-measurement state proportional to $A_j \varrho A_j^\dagger$. Since the probabilities add up to one, the measurement operators have to fulfill the completeness relation[3]

$$\sum_{j=1}^{m} A_j^\dagger A_j = \mathbb{1}_d. \tag{3}$$

In order to perform tomography, the quantum system is N times repeatedly prepared and subsequently measured, the frequency N_j of each measurement outcome j is recorded ($N = \sum_j N_j$),

[2] With $(\cdot)^\dagger$, we denote the conjugate transpose of a matrix or a vector.

[3] With $\mathbb{1}_d$, we denote the d-dimensional unit matrix. The positive operators $A_j A_j^\dagger$ fulfilling the completeness relation are also called *positive-operator valued measure (POVM) elements*. In publication[I] in the following section 2.3, they correspond to projectors out of eigenvectors of Hermitian matrices.

yielding an estimate $\hat{p}_j = N_j/N$ for the probability p_j. The most straightforward way to obtain a state estimate is to plug the probability estimates into Eq. (2), resulting in a linear equation system

$$\mathrm{tr}(A_j^\dagger A_j \varrho) = \hat{p}_j,\ j = 1,\ldots,m, \tag{4}$$

which can be rewritten as

$$E\vec{\varrho} = \hat{p}, \tag{5}$$

where $\vec{\varrho} \in \mathbb{C}^{d^2\times 1}$ is the vectorization of ϱ and the j-th row of the matrix $E \in \mathbb{C}^{m\times d^2}$ consists of the vectorization of $A_j^\dagger A_j$. Inverting this equation system requires the number of measurement outcomes m to be greater than or equal to the dimension of the density matrix of the state d^2 (*tomographic completeness*), and N substantially larger than m to obtain sufficiently precise p_j-estimates according to the law of large numbers. Namely, determining ϱ up to a *trace-distance* error[4] ε requires a *sample complexity* scaling with $O(d^4/\varepsilon^2)$ many copies to be measured.[5]

Due to the errors from finite counting statistics, the solution $\hat{\varrho}$ from inverting Eq. (5) will in general neither be positive semidefinite nor have unit trace, i.e., it will not correspond to a physical quantum state. This is addressed by *maximum likelihood estimation* (MLE) methods of finding the state that "most likely" resulted in the observed measurement data.[25,26] To this end, a *likelihood function* consisting of the product of the conditional probabilities $\mathbb{P}(N_j|\varrho)$ is optimized over all ϱ in $\mathscr{S}_d$. There exist efficient iterative implementations of this procedure (see, e.g., Ref. [27]).

In instances of few available measurement outcomes, an MLE fit of the state might be too restricted to the specific features of the available data, leading to issues with predicting future measurements and underestimation of errors.[28] Alternatively, the state together with error regions can be obtained with *Bayesian* methods,[29,30] which, however, require a priori knowledge about the *prior* distribution of possible quantum states. More general procedures for obtaining reliable error bounds are discussed in Refs. [31, 32].

Since superposition can occur in quantum systems, a composite quantum system ϱ_c has to be represented by the tensor product of its constituent systems $\varrho_1 \otimes \varrho_2 \otimes \ldots$ (see, e.g., Ref. [22, p. 94]). This results in an exponential increase of the required parameters with the system size—if ϱ_c represents a chain of n spins, its state space is in $\bigotimes_{j=1}^{n} \mathbb{C}^2$ with dimension $d = 2^n$—making it infeasible to process or even measure the required amount of data for more complex quantum systems: the *curse of dimensionality*.

[4] The two most commonly used distance measures for two quantum states ϱ and σ are the *trace distance* $T(\varrho,\sigma) := 1/2\,||\varrho-\sigma||_*$ and the *fidelity* $F(\varrho,\sigma) := ||\sqrt{\varrho}\sqrt{\sigma}||_* = \mathrm{tr}\sqrt{\sqrt{\varrho}\,\sigma\sqrt{\varrho}}$, which are related via $1-F \le T \le \sqrt{1-F^2}$ [23]. With $||A||_* := \mathrm{tr}(\sqrt{A^\dagger A})$, we denote the nuclear norm or trace norm of a matrix A, with $\sqrt{\sigma}$ denoting the unique matrix square root of a positive semidefinite matrix σ such that $(\sqrt{\sigma})^2 = \sigma$.

[5] Here and in the following, we make use of the Landau Big O/Omega notation[24] for describing the growth rate of a function: $f(x) = O(g(x))$ is equivalent to $\limsup_{x\to\infty} |f(x)/g(x)| < \infty$; $f(x) = \Omega(g(x))$ is equivalent to $\liminf_{x\to\infty} |f(x)/g(x)| > 0$. For example, $O(d)$ means a growth at most linear in d, $O(\log d)$ at most logarithmic, $O(\mathrm{poly}\, d)$ at most polynomial, and $\Omega(\exp d)$ at least exponential in d. If $f(x) = O(g(x))$ and $g(x) = O(f(x))$, we write $f(x) = \Theta(g(x))$.

There are different complementary strategies to address this issue. Often, one is not interested in a complete characterization of the system, but rather in certain parts or properties. To *certify* that the state ϱ at hand is close to a certain target state σ, one can estimate the *fidelity*[4] by performing those types of measurements that (on average) most likely allow for the observation of deviations of ϱ and σ—the more important, the more likely the measurement will be chosen.[33,34] For many systems of interest, this scheme requires only $O(\text{poly}\log d)$ many measurements. The same applies to certain *entanglement detection* procedures.[35] If only the predictions from a state for the *most probable* observables from a distribution of observables—not all of them—are of interest, one can employ a logarithmically scaling *computational learning* algorithm.[36] In the same direction goes *shadow tomography*, using 2-outcome measurements.[37] An efficient alternative to process tomography of quantum gates is *randomized benchmarking*, which provides an estimate of the error that gates make by constructing random circuits from them and recording how much the output states deviate from the input states.[38,39]

If, however, a complete characterization of a complex quantum system still is essential, more extensive schemes to effectively model a state with substantially fewer parameters than with an exponentially large density matrix have to be utilized. Pure states can be characterized using only $O(d)$ instead of d^2 parameters, and these can be obtained with compressed sensing–based methods[40]. Although not evading the exponential parameter blowup, these methods are highly relevant for many systems that would be inaccessible with ordinary tomographic methods.[41] Moreover, apart from purity, no further assumptions about the state have to be made, resulting in very generally applicable methods. The sampling complexity is reduced to $N = O(d\,r^2/\varepsilon^2)$ for determining a state ϱ of rank r up to a trace-distance error ε.[42,43] Using entangled measurements, this can be further improved to $O(d\,r/\varepsilon^2 \log(d/\varepsilon))$.[44] With a priori information about the state such as internal symmetries, the parameter space can be drastically reduced to $O(\text{poly}\log d)$.[45]

More generally, at least for low spatial dimension, most quantum states that we encounter in nature—mixed or pure—are much less entangled and hence confined to a very small part of the space of theoretically possible states, the "physical corner of Hilbert space", which can be efficiently parametrized by *tensor network states*.[46,47] These allow for a description with $O(\text{poly}\log d)$ parameters and efficient tomography protocols.[48] *Tensor networks* can also be employed for continuous systems with in principle infinitely many degrees of freedom such as quantum fields. The efficient tomography of quantum fields is the focus of the publications [2] and [3] in section 3.1.

2.2 Compressed sensing

As pointed out in the previous section, it is imperative to find "compressed" models of physical systems for a feasible description. For this, the system needs to be representable in a compressed way ("compressible"). For instance, a given signal vector $y \in \mathbb{C}^m$ could be composed out of very few (s many) vector components $a_j \in \mathbb{C}^m$ out of a huge set of components $\{a_1, \ldots, a_n\}$: $y = \sum_j x_j a_j$, where many of the coefficients $x_j \in \mathbb{C}$ are zero, but it is not known a priori which ones. Think, e.g., of a signal consisting of few sinusoid functions, but with unknown frequencies (*line*

spectral estimation), or of a small set of factors that linearly contribute to an effect and the signal vector comprising different observations: The model space $(a_1, \ldots, a_n)$ is large, but much less data is needed for a complete description of the signal—just the nonzero coefficients x_j together with their corresponding vectors a_j.

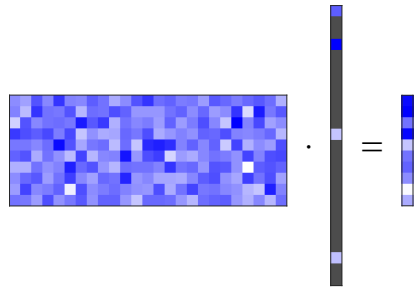

Figure 2: *Inverting a linear equation system with compressed sensing: Due to the sparsity of the parameter vector x (its zero-entries marked with gray), only a few columns of the "fat matrix" A (with few rows and many columns) contribute to the entries of the signal y. Although there are fewer rows than columns, the equation system has a unique sparse solution, which can be recovered with compressed sensing methods. This way, much shorter signals (fewer rows) can be used to determine the high-dimensional parameter vector.*

The task at hand can be rewritten in the following way (cf. Fig. 2): Given a linear equation system

$$Ax = y \tag{6}$$

with the matrix $A := (a_1|\ldots|a_n) \in \mathbb{C}^{m\times n}$, $m \ll n$, and $x \in \mathbb{C}^m$ being s-sparse ($s \ll n$), i.e., only s entries of x are nonzero: Find x.

The apparent problem is that this equation system is vastly underdetermined and the space of (not necessarily sparse) vectors that solve the equation system large. At first sight, it is far from clear how to find x within this space; applying the Moore-Penrose pseudoinverse to y just yields the solution of Eq. (6) with minimum ℓ_2-norm[6], which is non-sparse in general (cf. Fig. 3). It turns out that for general A and y it is even *NP-hard*[49] to solve the equivalent (non-convex) optimization problem of minimizing the support of x subject to Eq. (6),

$$\min_x ||x||_0 \quad \text{s.t.} \quad Ax = y, \tag{7}$$

which essentially amounts to trying out all possible combinations of indices. The optimization

[6] For $p \geq 1$, the ℓ_p-norm $||x||_p$ of a vector $x \in \mathbb{C}^n$ is defined as $||x||_p = \left(\sum_{j=1}^n |x_j|^p\right)^{1/p}$. For $p = 2$, we obtain the Euclidean norm, for $p = 1$ the *Manhattan* norm, and for $p = \infty$ the maximum norm. We also use the map $||\cdot||_0$, which counts the nonzero indices of a vector x, $||x||_0 := \mathrm{card}(\{j : x_j \neq 0\})$, and is not an actual norm, however, it holds that $||x||_0 = \lim_{p\to 0} ||x||_p^p$.

problem Eq. (7) can, however, be well approximated by minimizing the ℓ_1-norm of x,

$$\min_x ||x||_1 \quad \text{s.t.} \quad Ax = y, \tag{8}$$

cf. Fig. 3. This is a *convex optimization problem* due to the convexity of the target function $x \mapsto ||x||_1$ and the convexity of the constraints. Convex optimization problems are a very well-behaved type of

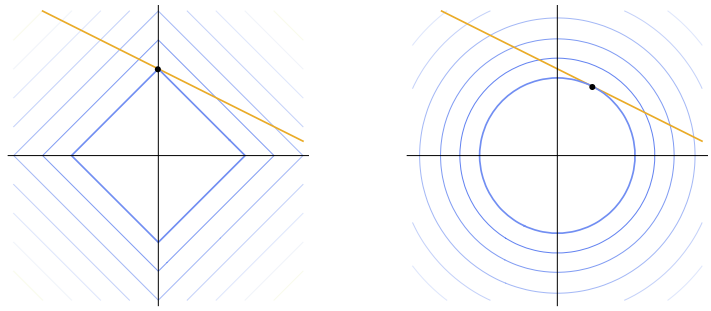

Figure 3: *ℓ_1-norm- vs. ℓ_2-norm-minimization: impact on sparsity. Left: Minimization of the ℓ_1-norm (2D norm balls $\{x : ||x||_1 = a\}$ of different size a in blue) subject to a 1D linear constraint (orange). The optimizing point—where the constraint intersects with the smallest feasible norm ball—is marked in black. Right: Minimization of the ℓ_2-norm subject to the same constraint. The extreme points of the ℓ_1-balls are located at the component axes—and this is where the optimizing points are to be found (except for the special constraints parallel to the lines of equal norm size) and only one component is nonzero. This carries over to higher dimensions, where optimizers with few nonzero entries are favored. In contrast, the solutions of the ℓ_2-optimization problems will in general have full support. The set $\{x : ||x||_0 = 1\}$ coincides with the x_1 and x_2 axis—minimizing $||x||_0$ subject to the constraint leads to the same optimizer as the ℓ_1-minimization.*

optimization problems—e.g., any local minimum is also a global minimum—with a vast body of optimization techniques.[50,51] In fact, providing provable and robust reconstruction guarantees for obtaining x by solving Eq. (8) with certain assumptions about x and A—having small *coherence*[7] or satisfying a *restricted isometry property* (RIP)[8] —is, what triggered the field of *compressed sensing*.[55,56] Specifically, it can be shown[57] that x in Eq. (8) can be *uniquely reconstructed* if x has at most s nonzero entries, the matrix A satisfies the RIP of order s, and its number of rows scales as

$$m = \Omega(s \log(n/s)). \tag{9}$$

[7] For a matrix $A \in \mathbb{C}^{m \times n}$ with normalized columns a_j, $||a_j||_2 = 1$ for $j = 1, \ldots, n$, its coherence[52] μ is defined as $\mu(A) = \max_{1 \le i \ne j \le n} |\langle a_i, a_j \rangle|$.

[8] If for an $(m \times n)$-matrix A and an integer $1 \le s \le n$ there exists a constant $\delta_s \in (0, 1)$ such that for every $m \times s$ submatrix A_s of A and for every s-sparse vector x holds $(1 - \delta_s)||X||_2^2 \le ||A_s||_2^2 \le (1 + \delta_s)||x||_2^2$, then A satisfies the s-restricted isometry property[53] with restricted isometry constant δ_s. An RIP is, e.g., fulfilled for Gaussian matrices, i.e., matrices with independent identically Gauss distributed entries.[54]

This means that much fewer observations (the rows of A, growing only linearly in the sparsity and logarithmically in the size of the solution vector) than contributing factors (columns of A) are required, allowing for obtaining a compressed model while "sensing" only few samples, hence the term "compressed sensing" or "compressive sensing". The results still apply when considering noisy signals and relaxing the constraint in Eq. (8) to a (still convex) least squares constraint

$$\min_x ||x||_1 \quad \text{s.t.} \quad ||Ax - y||_2 < \varepsilon, \tag{10}$$

yielding sparse approximate solutions.

Many problems in different fields can be rewritten as compressed sensing optimization problems of the type Eq. (8), resulting in a plethora of applications in fields such as magnetic resonance imaging (MRI)[58], fast photography[59], prediction of molecular vibrations[60], face recognition[61], machine learning[62], error correction[53], or radar and wireless communication[63]. For a concise introduction, see, e.g., Ref. [64]; for a comprehensive, more mathematical overview, see Ref. [65].

The heuristic use of the ℓ_1-norm as a penalty or regularizing term[9] already dates back to the 1960s and 1970s, where it was used in spectral estimation[68] and geophysics[69]. A method for obtaining the sparse spectral support for a signal was already developed in the 18th century[70] and is also employed in the reconstruction procedures beyond compressed sensing that are discussed in the quantum field tomography section 3.

Compressed sensing is not only limited to the recovery of vectors x from linear equation systems: The same paradigm can also be applied to the reconstruction of *low-rank matrices* $B \in \mathbb{C}^{p\times n}$, given only partial knowledge of it—such as a certain number of its entry values or functional values (see Fig. 4). This is called *matrix completion*[71,72], which has applications from collaborative filtering[73], phase retrieval[74], and system identification[75] to machine learning[76]. Low-rank matrices comprise a lot of internal structure due to the linear dependencies in their rows and columns so that in the compressed sensing spirit only a few matrix entries suffice to estimate the remaining ones. The underlying compressed model is found in the *singular value decomposition* (SVD)[10]

$$B = \sum_{j=1}^{r} s_j u_j v_j^\dagger \tag{12}$$

[9] An equivalent formulation of Eq. (10) with regularizing term $\lambda||\cdot||_1$ for suitable λ is the *basis pursuit denoising* problem[66]

$$\min_x ||Ax - y||_2 + \lambda||x||_1, \tag{11}$$

very often also called *least absolute shrinkage and selection operator* (LASSO)[67].

[10] For any matrix $B \in \mathbb{C}^{m\times n}$, there exists a (not necessarily unique) factorization $B = USV^\dagger$, with $S \in {\mathbb{R}_+}^{m\times n}$ diagonal and $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ unitary—the SVD. It can be computed in a numerically robust manner and plays a central role in many linear algebra applications. For positive semidefinite matrices, it coincides with the eigendecomposition of the matrix. The diagonal elements of S are called singular values, the orthonormal columns of U left singular vectors, the columns of V right singular vectors. The rank of B is equal to the number of non-zero singular values. Singular value decompositions are, e.g., used for computing the pseudo-inverse, the range, and the null space of a matrix. See also the section *quantum singular value decomposition* (appendix A.2).

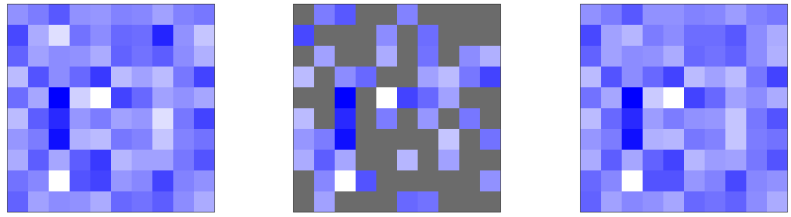

Figure 4: *Matrix completion: A low-rank matrix (left, $n = 10, r = 2$) is known only up to a few entries (middle, screened values marked with gray). The low rank of the matrix corresponds to linear dependencies and visible correlations between different columns and rows (compare, e.g., the coloring of the second and the last column of the matrix), providing sufficient information to a close reconstruction of the matrix (right). We can also see that a certain spreading of the screened values is important. If for instance an entire row or column is screened there is no way of reconstructing the full matrix.*

with $s_j \in \mathbb{R}_+$, $u_j \in \mathbb{C}^m$, and $v_j \in \mathbb{C}^n$ for $j = 1, \ldots, r$, amounting to $O(r \max\{m,n\}) \leq m\,n$ parameters—near-quadratic, low-rank matrices essentially require *quadratically fewer* parameters. The optimization is formally very close to the one in Eq. (8), namely

$$\min_{B \in \mathbb{C}^{p \times n}} ||B||_* \quad \text{s.t.} \quad \mathscr{A}(B) = y, \tag{13}$$

with the linear map $\mathscr{A} : \mathbb{C}^{m \times n} \to \mathbb{C}^p$ and the nuclear norm $||\cdot||_*$.[4] The analogy becomes even more apparent when considering that $||B||_*$ is also equal to the sum of the singular values of B, i.e., $||B||_* = ||(s_j)||_1$ and that the linear map $\mathscr{A}$ has a matrix representation. Moreover, Eq. (13) can be considered to be the convex relaxation of the optimization problem $\min_B \operatorname{rank}(B)$ s.t. $\mathscr{A}(B) = y$.[77] So, while minimizing the ℓ_1-norm yields a sparse vector, minimizing the nuclear norm yields a matrix with sparse singular spectrum, which is nothing else than a low-rank matrix. The big difference lies in the unitary degrees of freedom constituted by the singular vectors in the matrix completion setting. This also results in a different order of magnitude of the required number p of samples of B for matrix recovery algorithms: If $\mathscr{A}$ fulfills a matrix restricted isometry property[11] and

$$p \geq c\, r\, \max\{m,n\}, \tag{14}$$

with a numerical constant c, then B can be reconstructed with high probability,[78,79] which is also consistent with the minimum number of required model parameters. A similar scaling, picking up another logarithmic factor in n, exists for coherence assumptions.[72,80] Eq. (13) is a *semidefinite program*[81] and can be solved very efficiently.[82]

[11] Analogous to the vector case, a linear map fulfills the RIP at rank r if for all matrices B of rank at most r holds $(1-\delta_r)||B||_F^2 \leq ||\mathscr{A}(B)||_2^2 \leq (1-\delta_r)||B||_F^2$ with sufficiently bounded constant δ_r and entry-wise matrix ℓ_2-norm F (*Frobenius* norm).

Pure and low-rank quantum states constitute a compressible subset of the mixed quantum states and the benefits of compressed sensing can be harnessed for quantum state tomography[40] by restricting the feasible set in Eq. (13) to Hermitian positive semidefinite matrices (which encompass the set of quantum density matrices, Eq. (1)) and providing the constraint map $\mathscr{A}$ that models quantum measurements as in Eq. (5). For positive semidefinite matrices, the singular values and spaces coincide with the respective—positive—eigenvalues and -spaces; hence, also nuclear norm and trace are the same. Furthermore, experimental measurements generally entail noisy data, suggesting to relax the tight equality constraint to a least squares constraint. We therefore arrive at the following optimization problem to obtain a d-dimensional close to pure (small rank r) quantum state ϱ from roughly quadratically fewer measurement recordings y than full tomography:

$$\min_{\varrho \succeq 0} \operatorname{tr} \varrho \quad \text{s.t.} \quad ||E\,\vec{\varrho} - y||_2^2 < \varepsilon. \tag{15}$$

The measurements are modeled by the measurement matrix E as in Eq. (5). The positive semidefinite matrices form a convex cone within the set of Hermitian matrices and optimization on this set is still efficient. The requirement $\operatorname{tr} \varrho = 1$ is not explicitly part of Eq. (15)—and would also contrast with the goal of the objective function to minimize the trace—but it is implicitly included in the constraints. For noisier data y, the trace of the optimizing ϱ might differ more substantially from one, which is handled by dividing it by its trace.

It is still challenging to provide theoretical recovery guarantees for quantum states with practically relevant measurement maps as opposed to what is achievable in practice with numerical data, as will also be seen in the following publication [1]. If the matrix E consists of tensor products (or rather Kronecker products) of Pauli matrices, with $m = O(d\,r \operatorname{poly} \log d)$ it was shown that it satisfies the rank-r matrix restricted isometry property.[83] This, however, does not apply to E with rows consisting of eigenprojectors (outer products of eigenvectors) of Pauli matrices, each projector corresponding to a measurement outcome. If ϱ has rank one and the rows of E are built from outer products of Gaussian vectors, $O(d \log d)$ measurement outcomes suffice[74]—Gaussian vectors, however, are hard to implement experimentally. Closer towards experimental realization, yet random, are guarantees for outer products of rows of Haar distributed unitary matrices.[84] The randomness can be addressed by taking unitary t-designs[85] (*partial derandomization*). The rows of unitary designs in turn are closely related to eigenvectors of Pauli operators.[86]

Compressed sensing has successfully been applied for quantum state estimation in different experiments[41,87,88] as well for quantum process tomography[89] and is robust to noise together with theoretical error bounds[42,90]. Since its parameter reduction essentially boils down to a square root factor, it can only mitigate the effect of the exponential increase of the Hilbert space size. This makes compressed sensing the ideal tool for reconstructing intermediate-sized quantum systems. For even larger systems, additional assumptions have to be made, as will be the discussed in the subsequent section 3.

2.3 Experimental application and model selection

For practical quantum state tomography with experimental data, which in general is subject to noise, the noise parameter ε has to be taken in to account. Its magnitude has a large impact on the minimizer of Eq. (15): Large values for ε allow the algorithm to find a state with very low rank, while small values for ε restrict the feasible set of states compatible with the constraints to such an extent that only states with high rank (and enough parameters) solve the problem. This means that the same data could in extreme cases be modeled both by a pure (rank one) and a highly mixed (full-rank) state. It is therefore essential to carefully choose the appropriate model parameter ε such that the resulting model, the quantum state, predicts the future behavior of the quantum system best. This is an instance of a *model selection* problem. In general, various statistical models could be used to explain the outcome of an experiment. While some models may fit the specific measurement data realization very well, they might lack the generality to predict future instances accurately. This happens in particular in instances of *overfitting*, i.e., when properties of the data set at hand (due to noise etc.) that only appear in this specific realization of the experiment are incorporated into the model. In contrast, an *underfitting* model ignores features that are characteristic for the experiment and would repeatedly appear in different measurement instances.

A plethora of methods has been developed to tackle this problem; see Ref. [91] for an introduction and overview of model selection. Important tools are for example the *Akaike information criterion* (AIC)[92] and the *Bayesian information criterion* (BIC)[93], which balance the data fitting extent with the complexity of the model by introducing penalty terms into a maximum likelihood optimization problem that are proportional to the number of parameters of the model. The value ε directly affects the model complexity as well since the number of parameters that determine the model ϱ is proportional to its rank. Model selection techniques in the general quantum tomography context have, e.g., been employed in Ref. [94] and, using AIC and BIC, in Refs. [95, 96]. The latter methods, building on maximum likelihood estimation, however, are not directly applicable to the compressed sensing setting. A very robust method for model selection is *cross validation*, where the data set is split into training sets—used to create different models—and independent testing/validation sets, on which the models are evaluated how well they predict the data beyond the training set.

In the following publication [1], performing quantum state tomography via compressed sensing in a noisy data regime is systematically analyzed with model selection methods on the basis of a photonic experiment, preparing a four-qubit quantum state that could be employed as an element of a *one-way quantum computer* [97]. The experimentally realized quantum system was still small enough to supply sufficient data with reliable statistics for a complete state reconstruction with conventional methods, providing reliable reference information for estimating and certifying the performance of different model selection parameters as well as a large playground for different methods. The impact of the degree of incomplete information on the quality of the state reconstruction depending on the choice of the model was determined, providing a prescription for practical compressed sensing quantum state tomography.

FIRST-AUTHORED PUBLICATIONS

[1] Experimentally exploring compressed sensing quantum tomography
Adrian Steffens, Carlos A. Riofrío, Will McCutcheon, Ingo Roth, Bryn A. Bell, Alex McMillan, Mark S. Tame, John G. Rarity, and Jens Eisert, *Quantum Science and Technology* 2:025005, 2017.

[2] Quantum field tomography
Adrian Steffens, Carlos A. Riofrío, Robert Hübener, and Jens Eisert, *New Journal of Physics* 16:123010, 2014.

[3] Towards experimental quantum-field tomography with ultracold atoms
Adrian Steffens, Mathis Friesdorf, Tim Langen, Bernhard Rauer, Thomas Schweigler, Robert Hübener, Jörg Schmiedmayer, Carlos A. Riofrío, and Jens Eisert, *Nature Communications* 6:7663, 2015.

[4] An efficient quantum algorithm for spectral estimation
Adrian Steffens, Patrick Rebentrost, Iman Marvian, Jens Eisert, and Seth Lloyd *New Journal of Physics* 19:033005, 2017.

IOP Publishing Quantum Sci. Technol. 2 (2017) 025005 https://doi.org/10.1088/2058-9565/aa6ae2

Quantum Science and Technology

OPEN ACCESS

RECEIVED
5 December 2016

REVISED
15 March 2017

ACCEPTED FOR PUBLICATION
3 April 2017

PUBLISHED
11 May 2017

PAPER

Experimentally exploring compressed sensing quantum tomography

A Steffens[1,5], C A Riofrío[1], W McCutcheon[2], I Roth[1], B A Bell[3], A McMillan[2], M S Tame[4], J G Rarity[2] and J Eisert[1]

[1] Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, D-14195 Berlin, Germany
[2] Quantum Engineering Technology Laboratory, Department of Electrical and Electronic Engineering, University of Bristol, Merchant Venturers Building, Woodland Road, Bristol, BS8 1UB, United Kingdom
[3] Faculty of Science, University of Sydney NSW 2006, Australia
[4] School of Chemistry and Physics, University of KwaZulu-Natal, Durban 4001, South Africa
[5] Author to whom any correspondence should be addressed.

E-mail: adrian.steffens@fu-berlin.de

Keywords: quantum tomography, compressed sensing, quantum photonics

Abstract

In the light of the progress in quantum technologies, the task of verifying the correct functioning of processes and obtaining accurate tomographic information about quantum states becomes increasingly important. Compressed sensing, a machinery derived from the theory of signal processing, has emerged as a feasible tool to perform robust and significantly more resource-economical quantum state tomography for intermediate-sized quantum systems. In this work, we provide a comprehensive analysis of compressed sensing tomography in the regime in which tomographically complete data is available with reliable statistics from experimental observations of a multi-mode photonic architecture. Due to the fact that the data is known with high statistical significance, we are in a position to systematically explore the quality of reconstruction depending on the number of employed measurement settings, randomly selected from the complete set of data, and on different model assumptions. We present and test a complete prescription to perform efficient compressed sensing and are able to reliably use notions of model selection and cross validation to account for experimental imperfections and finite counting statistics. Thus, we establish compressed sensing as an effective tool for quantum state tomography, specifically suited for photonic systems.

Introduction

Quantum technologies have seen an enormous progress in recent years. Photonic architectures have matured from basic proof-of-principle schemes to intermediate scale quantum devices [1], while the robustness offered by integrated optical devices is poised to push these systems yet further [2, 3]. Similarly, systems of two-digit trapped ions [4] and other condensed-matter type systems such as superconducting devices are catching up at a remarkable pace [5]. Building upon this technological development, important primitives of quantum information science are being experimentally realised [6–10]. In light of these systems, it has become increasingly important to establish a toolbox for tomographic reconstruction that can keep up with this rapid development: The ironic situation that is emerging is that by now the state of large quantum systems can be manipulated with a high degree of control, but not easily reconstructed. Clearly, these technologies and the community require further advancement of their tools for state reconstruction. In this work, we discuss an explicit method to achieve such a reconstruction, thus contributing to this long-term goal. Specifically, we demonstrate a comprehensive exploration of the performance of state reconstruction in the photonic setting as one varies both the number of measurements and the noise model.

The framework of compressed sensing, a set of techniques originating from the context of classical signal processing [11, 12], has emerged as a key protagonist in closing the gap between technology and diagnostics [13–15]. The idea behind its functioning is rooted in the fact that a substantial amount of data encountered in realistic situations are structured and can be characterised by significantly fewer parameters than with ad hoc schemes. Approximately low-rank matrices are at the centre of the paradigm of matrix completion in

compressed sensing and correspond precisely to approximately pure quantum states. Since pure quantum states are widely regarded as the key resource for quantum information processing, such methods for reconstructing low-rank states are especially relevant. For even larger systems, tomographic tools based on basic variational sets are conceivable, with matrix product states [16, 17], their continuous analogues [18], and permutationally invariant states [19] providing prominent examples. The theory of such novel tools of reconstruction is progressing quickly. This applies, e.g., to new insights to the assignment of fair and rigorous confidence regions [20–23] as well as economical ways of performing instances of quantum process tomography [14, 24, 25].

Exciting steps toward using compressed sensing in experimental settings have been made [19, 24, 26, 27] in the regime in which one assumes knowledge about the basis in which sparsity is expected [24], assumes additional structure [19] or is in the highly tomographically incomplete regime [27]. In this work, we complement the picture for experimental tomography for medium-sized quantum systems. In its simplest formulation, compressed sensing tomography is based on a few random expectation values of suitable observables, from which approximately low-rank states can be accurately reconstructed [13]. This is suited for the situation in which expectation values can be obtained with good statistical significance, although acquiring many of them may be expensive. Still a missing piece in this picture, however, is the exploration of model selection techniques that have to be considered in the realm of experimental imperfections and finite counting statistics in order to make compressed sensing tomography a practical tool. Model selection allows to prevent over- and underfitting by controlling the dimensionality of the model of the system—in our case, the rank of the density matrix.

Here, we present a comprehensive analysis of experimental data from a multi-photon, multi-mode GHZ state source using tools of compressed sensing. Instead of working with expectation values of observables—as it is commonly done in this context, but may amount to information loss—our experimental setup allows us to obtain information on the individual projector level from the respective outcomes of each measurement setting. In contrast to complementing recent work [27], we are not tied to the regime of tomographically incomplete knowledge. This allows us to study the behaviour of the reconstruction for the entire range of measurement settings. We quantitatively explore model selection via cross validation and compare it to the model suggested by the anticipated noise statistics. With these tools, we provide a more systematic way to choose the appropriate parameters for compressed sensing quantum tomography. The results then provide the reader with the toolkit and understanding to effectively implement these methods for future quantum state tomography (QST) in general, and specifically for photonic systems.

This work is structured as follows. We start by reviewing concepts of quantum state tomography and discuss the specifics of compressed sensing in QST. We subsequently present our experimental setup consisting of a four-qubit photonic system, which is used as a test bed for our tomographical approach. We continue by discussing concepts of model selection in the context of QST and determine the appropriate model from the experimental data. With this, we perform compressed QST and study the performance of the reconstruction depending on the amount of collected data as well as the robustness of our method with respect to model mismatches.

Elements of quantum state tomography

Quantum state tomography is the most commonly used method to diagnose quantum information processing tasks. It is used to estimate the unknown quantum state of a system from data produced by measuring an ensemble of identically prepared systems. By fixing a basis, a general finite-dimensional quantum state can be identified with a positive semi-definite, unit-trace matrix, the density matrix

$$\varrho \in \mathcal{S}_d = \{\chi \in \mathcal{H}_d : \chi \succcurlyeq 0,\ \mathrm{tr}(\chi) = 1\}. \tag{1}$$

Here, $\mathcal{H}_d \subset \mathbb{C}^{d\times d}$ denotes the set of Hermitian matrices, and $\chi \succcurlyeq 0$ stands for a positive semi-definite matrix.

To determine the density matrix ϱ of a quantum system, we need to prepare sufficiently many copies of the state from identical preparations, perform a measurement on each copy using one out of m different measurement settings—corresponding to different observables, i.e. Hermitian matrices $A^{(j)}, j = 1,\ldots,m$—and count the respective number of measurement outcomes. Ideal measurements are associated with unit rank projectors $\Pi_k^{(j)} = v_k^{(j)} v_k^{(j)\dagger}$, where $v_k^{(j)}$ is the kth normalised eigenvector of $A^{(j)}$. For each measurement setting j the specific outcome $k = 1,\ldots, d$ occurs with probability

$$p_{j,k} \coloneqq \mathrm{tr}(\Pi_k^{(j)} \varrho). \tag{2}$$

Completeness, i.e. the property that the projectors sum up to unity,

$$\sum_{k=1}^{d} \Pi_k^{(j)} = \mathbb{1}, \tag{3}$$

ensures normalisation for each measurement setting j, so that $\sum_{k=1}^{d} p_{j,k} = 1$. For each measurement setting j, the outcome k corresponds to a random variable $Y_{j,k}$. Repeated measurements are independent from each other,

and are performed on N_j copies of the state for each measurement setting j, yielding the respective integer-valued realisation $y_{j,k}$ as observed frequency with $\sum_{k=1}^{d} y_{j,k} = N_j$. Hence, for each measurement setting j, the probability of the random variables $(Y_{j,1}, \ldots, Y_{j,d})$ to take the configuration of measurement outcomes $(y_{j,1}, \ldots, y_{j,d})$ is given by

$$\frac{N_j!}{y_{j,1}! \cdots y_{j,d}!} p_{j,1}^{y_{j,1}} \cdots p_{j,d}^{y_{j,d}}, \tag{4}$$

following a multinomial distribution $\mathcal{M}(N_j, (p_{j,1}, \ldots, p_{j,d}))$. Accordingly, we will obtain the kth outcome $N_j\, p_{j,k}$ times in expectation. We formalise the measurement process by introducing the linear operator

$$\mathcal{A}: \ \varrho \mapsto (N_j \mathrm{tr}(\Pi_k^{(j)} \varrho))_{j,k}, \tag{5}$$

which maps density matrices in $\mathcal{S}_d$ to matrices in $\mathbb{R}_+^{m \times d}$, corresponding to measurement outcomes $k = 1, \ldots, d$ for different measurement settings $j = 1, \ldots, m$. We emphasise that $\mathcal{A}(\varrho)$ is not an experimental data matrix itself; according to the law of large numbers, the frequencies in each measurement realisation $\mathcal{Y} := (y_{j,k}) \in \mathbb{N}^{m \times d}$ from the experiment will converge to $\mathcal{A}(\varrho)$ with growing number of measurements N_j, i.e. the expectation value $\mathbb{E}(Y_{j,k})$ of the random variable $Y_{j,k}$ is given by

$$\mathbb{E}(Y_{j,k}) = N_j \mathrm{tr}(\Pi_k^{(j)} \varrho) \tag{6}$$

for each $j,\ k$. Apart from additional systematic sources of error, e.g. due to experimental imperfections, the difference between $\mathcal{Y}$ and $\mathcal{A}(\varrho)$ is due to finite counting statistics, and in many settings, this is the largest contribution to the error.

The most straightforward approach to determine ϱ from $\mathcal{Y}$ would be to attempt to invert the linear system of equations

$$\mathcal{A}(\varrho) = \mathcal{Y}. \tag{7}$$

In general, however, noise on the data $\mathcal{Y}$ would render the *reconstructed* density matrix $\hat{\varrho}$ unphysical ($\hat{\varrho} \not\geqslant 0$). A generic (full rank) density matrix in $\mathcal{S}_d$ is determined by $d^2 - 1$ independent real parameters. Hence, in general, one requires at least $d^2 - 1$ linearly independent equations in order to solve equation (7). This is also called *tomographic completeness* and corresponds to *informational completeness* with sufficient information to in principle capture full rank states. For further notions about informational completeness under prior information (e.g. the rank of the state is assumed to not be maximal) see [28, 29]. When dealing with significantly less information, specialised reconstruction techniques are important with compressed sensing being a natural choice, which we will discuss in the next section.

In our system, we will be concerned with local Pauli measurements on each subsystem of a multi-partite state. We measure an n-qubit system ($d = 2^n$) using m different measurement settings, each of which corresponds to an n-qubit Pauli operator

$$A^{(j)} = \bigotimes_{i=1}^{n} \sigma_i^{(j)}, \tag{8}$$

$j = 1, \ldots, m$, with $\sigma_i^{(j)} \in \{\sigma_x, \sigma_y, \sigma_z\}$, where $\sigma_x, \sigma_y, \sigma_z$ are the Pauli matrices. This is often referred to as *Pauli basis measurement*. The projectors of the two-qubit operator $A^{(1)} := \sigma_z \otimes \sigma_z$, for example, are $\Pi_1^{(1)} = |0,\ 0\rangle\langle 0,\ 0|$, $\Pi_2^{(1)} = |0,\ 1\rangle\langle 0,\ 1|$, $\Pi_3^{(1)} = |1,\ 0\rangle\langle 1,\ 0|$, and $\Pi_4^{(1)} = |1,\ 1\rangle\langle 1,\ 1|$. The identity matrix can be excluded for each qubit since it has the same eigenvectors and hence corresponding projectors as σ_z and does not provide any additional information about the state. Note, that in a Pauli basis measurement, one obtains 2^n outcomes per measurements setting, as opposed to the *Pauli expectation value measurements*, in which we only use one expectation value per measurement setting. Pauli expectation value measurements (including those containing identity matrices) can easily be obtained from Pauli basis measurements by simply computing the expectation values from the projection data for each measurement setting. For n qubits, there exist $m_{\max} := 3^n$ different Pauli words in total, each with 2^n eigenvectors, which corresponds to a maximum of $3^n \cdot 2^n$ equations in equation (7). Each set of Pauli projectors $\{\Pi_k^{(j')}\}_{k=1}^{d}$ for fixed setting j' contains a subset of elements that is linearly independent from the projectors for all other settings. Hence, any number of smaller than $m_{\max}$ measurement settings will lead to the loss of tomographic completeness. When performing QST on large systems, however, it is of practical necessity to employ as few measurement settings as possible (and often also only few repetitions per measurement setting). The key question arising in this context, therefore, is whether it is feasible to reconstruct an unknown state ϱ with not only $m < m_{\max}$ measurement settings, but a significantly smaller subset. The need for minimising the number of measurement settings is particularly pressing in architectures such as linear optical ones, since high repetition rates and good statistics are available, while it can be tedious or costly to alter the measurement setting. This is indeed the case in many practically relevant situations using compressed sensing schemes, which will be discussed in the next section.

Compressed sensing for quantum state tomography

By parameter counting, a state with rank $r < d$ can be completely characterised by fewer than d^2 parameters, that is $\sim rd$ and informational completeness could in principle be achieved using correspondingly fewer measurement settings. However, it is far from obvious how to acquire these parameters using fewer measurement settings and how to do so in a robust fashion—this is the starting point for compressed sensing [11, 30]. Originally conceived for reconstructing sparse vectors, the concept was extended to the recovery of low-rank matrices [31, 32] and adapted to the problem of QST [13, 33]. Here, one again considers structured problems in which one can exploit the fact that in many useful settings approximately low-rank states are of interest. This is a reasonable assumption, since most quantum information experiments aim at preparing pure states.

In order to obtain a general complex-valued low-rank matrix from measurements $\mathcal{A}$, naïvely, one would search within the set of low-rank matrices for the one that matches the measurement constraint, solving

$$\min_{\chi\in\mathbb{C}^{d\times d}} \operatorname{rank}(\chi) \quad \text{s.t.} \quad \mathcal{A}(\chi) = \mathcal{Y}. \tag{9}$$

The key idea for compressed sensing in matrix recovery is to relax this NP-hard problem [34] into the closest convex optimisation problem [35]

$$\min_{\chi\in\mathbb{C}^{d\times d}} \|\chi\|_* \quad \text{s.t.} \quad \mathcal{A}(\chi) = \mathcal{Y}. \tag{10}$$

We denote the nuclear norm (better known as the trace norm in the context of reconstructions in quantum mechanics) of a matrix χ by $\|\chi\|_* := \operatorname{tr}(\sqrt{\chi^\dagger\chi})$. Such problems are well known to be efficiently solvable [36].

The crucial question in compressed sensing is how many measurements are required to satisfactorily reconstruct the sought-after matrix. Many proofs rely on randomised measurements schemes: In [37], it has been shown that for a general map $\mathcal{A}:\mathbb{R}^{d\times d}\to\mathbb{R}^M$ with Gaussian entries, $M \gtrsim 3r(2d-r)$ copies of ϱ are provably sufficient for the recovery of ϱ. Building on this and closer to our situation is the recovery guarantee presented in [38], in which $M \geqslant crd$ copies are needed with some constant $c > 0$, for $\mathcal{A}:\mathcal{S}_d\to\mathbb{R}^M$, $\varrho\mapsto(\operatorname{tr}(\Pi^{(j)}\varrho))_{j=1,\ldots,M}$, mapping density matrices from $\mathcal{S}_d$ to vectors in $\mathbb{R}^M$, with $\Pi^{(j)} = v^{(j)}v^{(j)\dagger}$, and $v^{(j)}$ a Gaussian vector for each j. In practice, numerical computations outperform these theoretical bounds. However, there is a fundamental lower bound for the number of copies, $M = 4r(d-r) - 1$, using a theoretically optimal POVM with M elements [39]. Note that—in the mindset of measurement settings and outcomes—the number of outcomes k per measurement setting j scales with the dimension of the Hilbert space d. Since M corresponds to $m\ d$, the number of measurement settings scales just with the rank, i.e. $m = c\ r$.

It is in general harder to prove comparable results for deterministic measurements—in our setting with $v^{(j)}$ being eigenvectors of Pauli operators. To bridge this gap, notions of partial derandomisation have been introduced, where $v^{(j)}$ are not Gaussian, but drawn from spherical designs—certain finite subsets of the d-dimensional complex sphere—leading to similar statements [38]. Spherical designs, in turn, can be related to eigenvectors of n-qubit Pauli operators [40]. Apart from results on the level of expectation values [41], less has been proven for products of single-qubit eigenvectors, the setting at hand—strikingly in contrast to the great success of the procedure in practice. These results remain stable when taking noise into account.

The measured data can be written as

$$\mathcal{Y} = \mathcal{A}(\varrho) + \mathcal{N}(\varrho) = (N_j\operatorname{tr}(\Pi_k^{(j)}\varrho))_{j,k} + (\eta_{j,k})_{j,k}, \tag{11}$$

with $\mathcal{N}$ and $\eta_{j,k}$ representing the noise due to finite counting statistics. For positive semi-definite matrices such as quantum states, the nuclear norm of a matrix reduces to the trace of the matrix. Consequently, relaxing the equality constraint in equation (10) and including the positivity constraint, we arrive at the semi-definite programme (SDP) [34]

$$\min_{\chi\succcurlyeq 0} \operatorname{tr}\chi \quad \text{s.t.} \quad \|\mathcal{A}(\chi) - \mathcal{Y}\|_2^2 < \varepsilon, \tag{12}$$

for some yet-to-be-determined $\varepsilon > 0$ and $\|\cdot\|_2$ representing the entrywise two-norm. This is exactly the problem we aim to solve in order to achieve efficient QST. SDPs, being convex programmes, feature a rich theory, and numerical implementation is easily achievable [42, 36]. Note that the procedure minimises the trace, which at first sight might seem contradictory to the requirement for density matrices to have unit trace. However, the unit trace requirement is implicitly included in the data constraint since the probabilities in the map $\mathcal{A}$ are normalised. Perfect data would lead to an optimiser with trace exactly equal to one. In turn, a relaxation of this constraint leads to a relaxation of the unit trace requirement. As a result, generically for not too small ε, the optimal χ, denoted by $\hat{\chi}$, will be subnormalised, due to its location on the part of the boundary of the ε-ball with the lowest trace. In order to obtain a physically meaningful reconstruction $\hat{\varrho}\in\mathcal{S}_d$, we find in our simulations that renormalising via

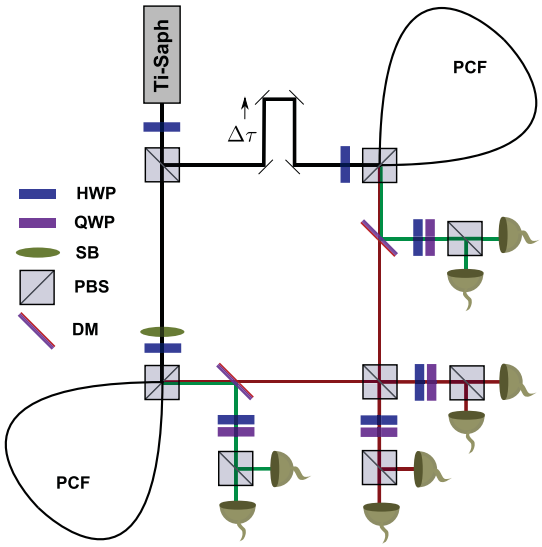

Figure 1. Experimental setup for generating the four-photon polarisation entangled states $|\psi_{\mathrm{GHZ}}\rangle$, consisting of photonic crystal fibre (PCF) sources, half-waveplates (HWPs), quarter-waveplates (QWPs), a Soleil-Babinet (SB), polarising beam splitters (PBSs) and dichroic mirrors (DMs). The 80 MHz Ti-Saph laser is split onto two PCF sources in twisted Sagnac-loop interferometer configurations generating polarisation entangled Bell pairs. The signal and idler photons from each source are separated by DMs and the signal photons interfere on a PBS with relative time between paths $\Delta\tau \approx 0$, which on post-selecting a single photon in each output port performs a fusion operation. The SB is set to match the phase between the $|H, H, H, H\rangle$ and $|V, V, V, V\rangle$ components to zero. Each mode is measured by single-qubit rotations consisting of a HWP and QWP, and is projected in the $\{|H\rangle, |V\rangle\}$-basis by PBSs and avalanche photodiode detectors.

$$\hat{\chi} \mapsto \hat{\varrho} := \frac{\hat{\chi}}{\mathrm{tr}(\hat{\chi})} \tag{13}$$

produces the highest fidelity results. To carry out the optimisation procedure, we employ the convex optimisation solver SDPT3 4.0 [43] together with CVX [44]. For higher Hilbert space dimensions, methods like singular value thresholding [45] come into play, which typically are faster, but less accurate.

Experimental setup

The experiment is designed to prepare the four-qubit GHZ state associated with the state vector

$$|\psi_{\mathrm{GHZ}}\rangle = \frac{1}{\sqrt{2}}(|H, H, H, H\rangle + |V, V, V, V\rangle) \tag{14}$$

with the qubits encoded in the polarisation degree of freedom of four photons. Here, $|H\rangle$ and $|V\rangle$ represent horizontally and vertically polarised photons, respectively, hence effectively spanning a two-dimensional Hilbert space. The experimental setup, building upon the one outlined in [46], is shown in figure 1 and consists of two Bell pair sources which undergo a parity check or post-selected fusion [8, 47–52] to probabilistically generate the GHZ state. Both the photon pairs, generated by spontaneous four-wave-mixing in microstructured fibres, and the fusion operation are successful only probabilistically, but in a heralded fashion, i.e. a classical signal is available signifying success of the preparation. Successful generation of the state is determined by post-selecting only four-photon coincident events which occur at a rate of approximately 1–2 Hz. The post-selected data is effectively free from dark counts—noise generated by single-photon detectors firing erroneously in the absence of a photon. This is due to the fact that the rate at which dark counts in n modes occur in the coincidence window decreases exponentially with n, i.e. four simultaneous dark counts are negligibly rare. Due to additional experimental imperfections, however, the prepared state is non-ideal. The main cause of deviation between the actually prepared state and the target state arises from the distinguishability of photons partaking in the fusion operation and inherent mixedness from the parasitic effects in the pair generation [53]. These tend to cause the generated state to resemble a partially dephased GHZ state [8]. Measurements on the state then proceed using single-qubit rotations (waveplates) and projections (polarising beam splitters and single-photon detection with avalanche photo-diodes) using well-characterised bulk-optical elements allowing high-fidelity measurements to be performed.

As stated above, in order to achieve a tomographically complete basis for n qubits, one requires $m_{\max} = 3^n$ measurement settings. In our system of four qubits, $n = 4$, we have measured a tomographically complete set of 81 local Pauli operators. For each measurement setting, around 650 four-coincident events are accumulated within an integration time of six minutes. Evidently, given the exponential scaling of the tomographically complete set of measurement settings, achieving such reliable statistics for larger states ($n > 4$) is increasingly demanding on resources and quickly becomes infeasible.

Model selection

The starting point for carrying out compressed sensing quantum tomography is the question of determining an appropriate value for ε in the optimisation procedure equation (12). Essentially, larger values of ε result in greater relaxation of the data fitting constraint, leading to lower-rank estimates $\hat{\varrho}$; while smaller ε values will yield $\hat{\varrho}$ matrices with larger rank, which better fit the particular data set. Depending on the underlying state and the particular instance of noise in the data, the choice of ε might result in underfitting with too coarse a model, or in overfitting—i.e. including parts of the noise into the model of the state. Both extremes in general lead to states that fail to correctly predict future data. In the most severe cases, it could happen that using the same measurement prescription $\mathcal{A}$ and the same data $\mathcal{Y}$, the optimisation procedure in equation (12) yields a full rank or a rank-one matrix, depending on the choice of ε. Worse still, too small a value of ε can make the optimisation procedure unfeasible, whereby there is no feasible state that would result in data sufficiently close to that measured. The task of determining the appropriate model—in our case, the value of ε—that is statistically faithful to the data via an appropriate choice of the respective external parameters is called *model selection* (see e.g. [54]). Several ideas of model selection have a rigorous mathematical underpinning: Particularly well known is the Akaike information criterion (AIC) [55], providing a measure of the relative quality of statistical models for a given set of data. For a collection of models compatible with a given data set, this criterion gives an estimate for the relative quality of each model. Similarly frequently employed is the Bayesian information criterion (BIC) [56]. Direct application of AIC and BIC to quantum tomography—an approach followed in [15]—is problematic for larger systems since it requires rank-restricted maximum likelihood estimation, leading to non-convex optimisation, which scales unfavourably with the system size. This is due to the fact that these techniques are discrete in the sense that they explicitly restrict the rank of the density matrix. In the compressed sensing mindset, the parameter that controls the rank in a continuous fashion is ε. As we mentioned above, this is at the centre of our discussion.

For sufficiently small noise, a promising ansatz for identifying a suitable ε is to use the data to compute the estimate $\hat{\varepsilon}(\mathcal{Y})$ according to the expectation value of

$$\|\mathcal{A}(\chi) - \mathcal{Y}\|_2^2 = \|\mathcal{N}(\varrho)\|_2^2. \tag{15}$$

Assuming the noise is solely due to finite counting statistics, i.e. the deviations from measurement outcomes from the expected variance of the multinomial distribution, we obtain

$$\mathbb{E}(\|\mathcal{N}(\varrho)\|_2^2) = \sum_{j,k}\mathbb{E}(\eta_{j,k}^2) = \sum_{j,k}\mathbb{V}(\eta_{j,k}) = \sum_{j,k}N_j p_{j,k}(1 - p_{j,k}), \tag{16}$$

with variance $\mathbb{V}$. The second step follows from $\mathbb{E}(\eta_{j,k}) = 0$ for each j and k. In order to compute $\hat{\varepsilon}$ from the data, we need to approximate $p_{j,k}$ as $y_{j,k}/N_j$, which is reasonable for sufficiently large N_j according to the law of large numbers. By equation (16), we obtain the estimate

$$\hat{\varepsilon}(\mathcal{Y}) \coloneqq \sum_{j=1}^{m}\sum_{k=1}^{d} y_{j,k}(1 - y_{j,k}/N_j). \tag{17}$$

This choice of $\hat{\varepsilon} = \hat{\varepsilon}(\mathcal{Y})$ scales linearly with m, the number of measurements in the data set $\mathcal{Y}$. Note that $\hat{\varepsilon}$ depends on the noise model, which in several cases may not be sufficiently established. In our case, however, the noise model is known to a high degree, which allows us to study and compare different methods for estimating the parameter ε.

Complementarily, we employ a straightforward, well-established model selection technique based on cross validation (see e.g. [57]), which is more scalable than the use of AIC or BIC in our case. Because of its generality and independence from the noise model, it can be used in a variety of scenarios. The method works as follows: The data is partitioned into independent training and testing sets. Different models, i.e. different values for ε, are built from the training data and used to predict the testing data. The sought-after parameters—in our case ε—then result from the model corresponding to the smallest error with respect to the testing data.

Specifically, we randomly draw $m = 10, 15, 20, 40, 60, 80$ out of the $m_{\max} = 81$ measurement settings without replacement, corresponding to different levels of limited experimental knowledge. The respective data sets $\mathcal{Y}(m) \in \mathbb{R}_+^{m\times d}$ are then partitioned into five subsets $\mathcal{Z}^{(1)}(m), \ldots, \mathcal{Z}^{(5)}(m) \in \mathbb{R}_+^{m/5\times d}$. The optimisation in equation (12) is performed with respect to every possible union of four subsets $\bigcup_{i=1,i\neq q}^{5}\mathcal{Z}^{(i)}(m)$, $q = 1, \ldots, 5$,

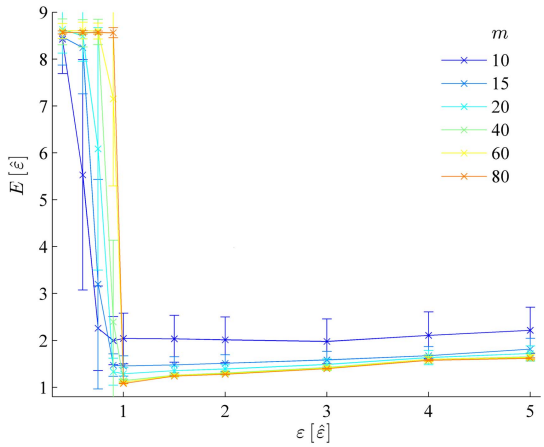

Figure 2. Cross validation results. Prediction errors $E(m, \varepsilon) = 1/5 \sum_{q=1}^{5} \|\mathcal{A}_{m,q}(\hat{\varrho}(m, q, \varepsilon)) - \mathcal{Z}^{(q)}(m)\|_2$ in units of $\hat{\varepsilon}$ depending on the model parameter ε and on the number of measurement settings m. The standard deviation is bigger for fewer measurement settings and for smaller ε. The latter is due to the increasing chance of the optimisation to be infeasible for smaller ε. For ε close to $\hat{\varepsilon}$ and sufficient many measurement settings, the error is only slightly bigger than the deviation due to the multinomial distribution of the measurement outcomes.

and different ε parameters. Each reconstruction yields an estimate $\hat{\varrho}(m, q, \varepsilon)$ and the remaining subset $\mathcal{Z}^{(q)}(m)$ is used as a testing set. The state estimate $\hat{\varrho}(m, q, \varepsilon)$ is used to compute the *predicted* measurement data $\mathcal{A}_{m,q}(\hat{\varrho}(m, q, \varepsilon))$ and compare these with the corresponding subset of the *experimental* measurement data $\mathcal{Z}^{(q)}(m)$ ($\mathcal{A}_{m,q}: \mathcal{S}_d \to \mathbb{R}_+^{m/5 \times d}$ being the reduction of the operator $\mathcal{A}$ to the subsets of measurement settings corresponding to m and q). The resulting distance $\|\mathcal{A}_{m,q}(\hat{\varrho}(m, q, \varepsilon)) - \mathcal{Z}^{(q)}(m)\|_2$, between the *predicted* and *measured* data, also known as the prediction error or predicted risk, is averaged over q (fivefold cross validation), yielding an estimate for the averaged prediction error (testing set error)

$$E(m, \varepsilon) = \frac{1}{5} \sum_{q=1}^{5} \|\mathcal{A}_{m,q}(\hat{\varrho}(m, q, \varepsilon)) - \mathcal{Z}^{(q)}(m)\|_2. \tag{18}$$

If the corresponding optimisation problem is infeasible for a certain combination of ε, m, and q (i.e. the set of density matrices that satisfy the constraint in equation (12) is empty), the prediction error is set to $\|\mathcal{Z}^{(q)}(m)\|_2$. For averaging, each point (m, ε) is sampled 50 times.

The mean values and standard deviations of the prediction error depending on the model parameter are depicted in figure 2. We see that for values of ε around $\hat{\varepsilon}$ the error is smallest, which is consistent with our ansatz and allows us to gain confidence in the assumption that the measurement data can be effectively modelled by a multinomial distribution. The more measurement settings are considered, the clearer the choice of the optimal ε becomes, with both the prediction error and its variance attaining their minima close to $\varepsilon = \hat{\varepsilon}$. For those values of ε close to $\hat{\varepsilon}$ and sufficiently many measurement settings, the prediction error $E(m, \varepsilon)$ is only slightly bigger than the error estimate for the data ε. Here, the error arises primarily from raw multinomial noise, ε, present in the testing set itself and cannot be overcome with improved reconstruction methods. Where fewer measurement settings are considered, less information about the state is available, resulting in large testing set errors as well as greater variance of the state estimates, although the smallest prediction errors are still seen for ε close to $\hat{\varepsilon}$. As ε decreases below $\hat{\varepsilon}$, the chance of the optimisation being infeasible increases, causing the prediction errors to effectively increase with a greater spread attributed to different optimisation runs. As ε increases above $\hat{\varepsilon}$, the data fitting constraint is weakened, resulting in too coarse model fits and a gradually increasing prediction error.

Using equation (17) instead of cross validation has the advantage of much less computational effort and is useful in a scenario with good statistics for each measurement setting. Moreover, cross validation relies on partially discarding data, which could aggravate the issues of having too little data, yielding poorer estimates for ε. However, equation (16) relies on the assumption of a well identified error model—in our case, multinomial noise, as verified by cross validation. In cases in which the error model is not known, cross validation can provide a more robust estimate of ε.

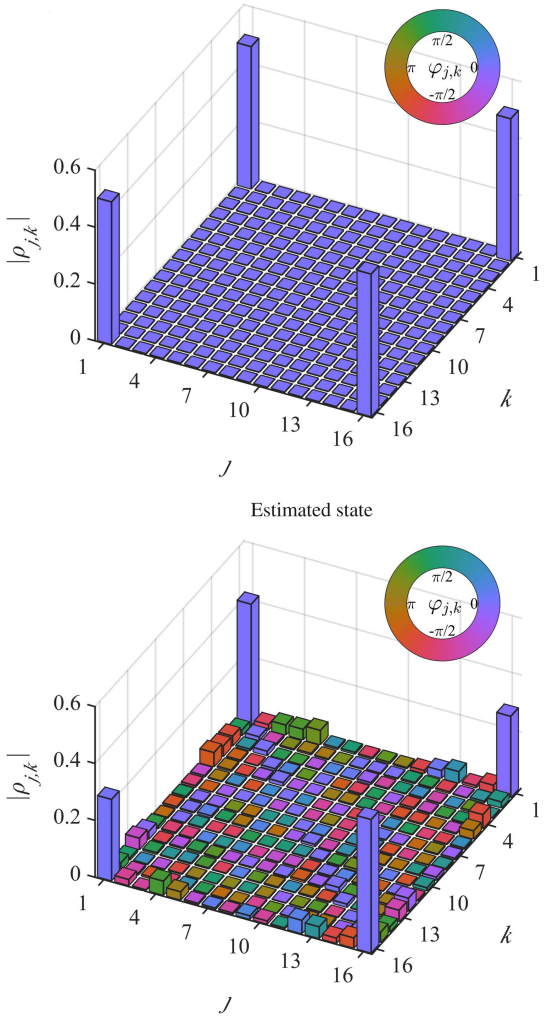

Figure 3. Bar plot of the density matrix of the target (GHZ) state ρ_{GHZ} and its optimal compressed sensing estimate $\hat{\rho}_{\mathrm{CS}}$. The basis is fixed to the tensor products of one-particle vectors in the order $|H, H, H, H\rangle$, $|H, H, H, V\rangle, \ldots, |V, V, V, V\rangle$. The height of each bar corresponds to the size of the absolute value of the respective density matrix entry $\varrho_{j,k} = |\varrho_{j,k}|e^{\mathrm{i}\varphi_{j,k}}$ and the colour to its complex phase $\varphi_{j,k} \in (-\pi, \pi]$. The colourmap is chosen to account for the periodicity of the phase. The fidelity of the estimate with respect to the GHZ state is 0.855 ± 0.006 and its purity $\mathrm{tr}(\hat{\varrho}_{\mathrm{CS}}^2) = 0.60 \pm 0.01$, representing an expected mixedness due to experimental imperfections.

Compressed sensing tomography of the GHZ state

Having verified that the optimal value for ε is close to that computed from equation (17), we use it as input for the compressed sensing tomography of the experimental state and compute the optimal estimate $\hat{\varrho}_{\mathrm{CS}} := \hat{\varrho}\,(m_{\max}, \hat{\varepsilon})$ of the *a priori* unknown experimentally prepared state ϱ. The good statistics available in our experiment allow us to estimate ϱ with comparably high accuracy. In general, due to experimental imperfections, ϱ (and hence $\hat{\varrho}$) will deviate from the target state $\varrho_{\mathrm{GHZ}} := |\psi_{\mathrm{GHZ}}\rangle\langle\psi_{\mathrm{GHZ}}|$, see figure 3 for a pictorial representation. There, we show a comparison between the density matrices of the target state and the optimal compressed sensing estimate using bar plots.

The standard figure of merit to determine the performance of tomography is the quantum fidelity F of two states χ and σ, which is defined as $F(\chi, \sigma) = \mathrm{tr}((\sqrt{\chi}\sigma\sqrt{\chi})^{1/2})$ [6]. We find that the fidelity between the GHZ state ϱ_{GHZ} and the estimated state $\hat{\varrho}_{\mathrm{CS}}$ is

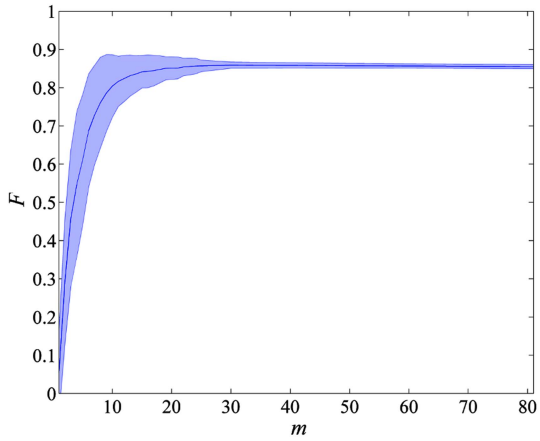

Figure 4. Fidelity $F(\varrho_{\mathrm{GHZ}}, \hat{\varrho}(m, \hat{\varepsilon}))$ as a function of the number of measurement settings m with uncertainty (shading) from bootstrapping for $\varepsilon = \hat{\varepsilon}$. For large m, F approaches the fidelity of ϱ_{GHZ} and $\hat{\varrho}$, $F(\varrho_{\mathrm{GHZ}}, \hat{\varrho}_{\mathrm{CS}}) = 0.855$, getting very close already for comparably few measurement settings, and the standard deviation becomes smaller.

$$F(\varrho_{\mathrm{GHZ}}, \hat{\varrho}_{\mathrm{CS}}) = 0.855 \pm 0.006. \tag{19}$$

The uncertainty of the fidelity is determined by using the optimal compressed sensing estimate, $\hat{\varrho}$, as input for the generation of simulated data—parametric bootstrapping [57]—and taking the empirical standard deviation of the fidelity values. This uncertainty determines the robustness of the method. Obtaining a closed expression for proper error bounds from the data with respect to positivity constraints is hard [23, 58], while bootstrapping and taking the empirical standard deviation gives a good estimate of uncertainty [57].

To build confidence, we also computed the maximum likelihood estimate [59], $\hat{\varrho}_{\mathrm{MLE}}$, using the same data to obtain a fidelity with respect to the target state of $F(\varrho_{\mathrm{GHZ}}, \hat{\varrho}_{\mathrm{MLE}}) = 0.843 \pm 0.004$, which shows that the estimators yield similar results; as will other estimators such as least squares with positivity constraint. Additionally, since we have measured a tomographically complete set of observables and the statistical properties of the measured data are sufficiently understood, we are able to provide an estimate of the fidelity with respect to the target state directly from the measured data without the need of performing tomography and an estimate of the corresponding error bound, see the appendix for details. With this, we obtain a fidelity of 0.845 ± 0.005, which again is in good agreement with the results computed from the compressed sensing estimate. We note that the standard technique for estimating the fidelity of a state with respect to a specific target state requires estimating only the expectation values of a set of operators that form a decomposition of the target state. For a four-qubit GHZ state, this requires a minimum of nine specific Pauli basis measurements, as explained in the appendix. In contrast, using compressed sensing tomography, even a random set of measurement settings produces fidelity estimates with respect to the GHZ state, which quickly approach the maximum at around 25 measurement settings. Furthermore, these measurement settings suffice to compute the fidelities with respect to arbitrary states, since they allow for the estimation of the entire state.

Compressed sensing is about employing provably fewer measurement settings than with standard methods, while still producing satisfactory reconstructions, i.e. to effectively sense in a compressive way. Along these lines, we explore how varying the number of measurement settings m affects the fidelity. This is shown in figure 4. In order to make the results independent from specific measurement settings, we randomly draw without replacement m out of m_{max} different settings 200 times and average over the resulting fidelities, thus providing a value for a typically expected fidelity for each m. As one would expect intuitively, we can see that the value of the fidelity increases monotonically with the number of measurement settings and converges to the fidelity of the estimate from tomographically complete data. The shaded region represents the uncertainty ($\pm$ standard deviation) in the fidelity computed via bootstrapping and displays the decreasing uncertainty with increasing numbers of measurement settings. The fidelity already falls within the error bars of its final value for comparably small m.

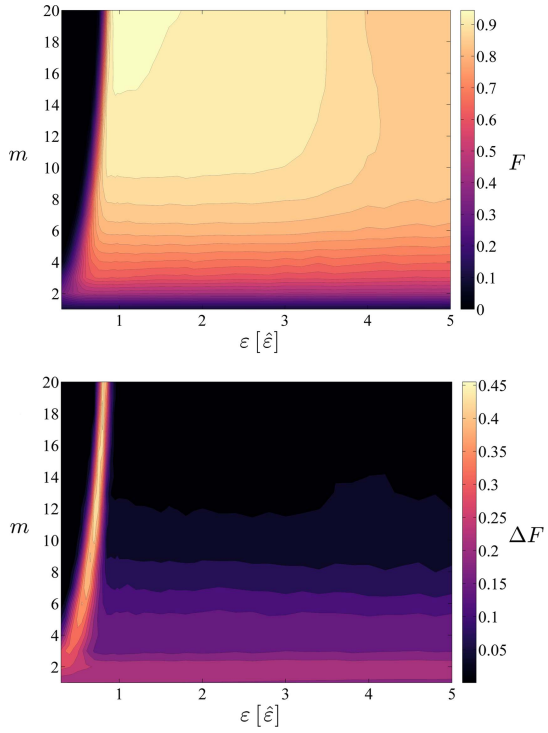

Figure 5. Fidelity $F(\hat{\varrho}_{\mathrm{CS}}, \hat{\varrho}(m, \varepsilon))$ depending on the number of measurement settings m and the model parameter ε (top) and corresponding standard deviation ΔF (bottom) obtained via bootstrapping. Since in compressed sensing we are more interested in the regime of few measurement settings and the fidelities do not change significantly for larger m, we restrict ourselves to the region with $m \leqslant 20$. The data are generated randomly from $\hat{\varrho}_{\mathrm{CS}}$ and the measurement settings per m are drawn randomly as well. The fidelities are averaged over different data realisations and measurement settings. The highest fidelities are achieved for $\varepsilon \approx \hat{\varepsilon}$ with rapid decrease for $\varepsilon < \hat{\varepsilon}$ where the fraction of infeasible optimisations increases. Note that the higher the fidelity, the lower the standard deviation.

Deviations from the optimal parameter

In this section, we study the effect that misestimating ε has in the performance of the reconstruction of the state. We carry out this task by numerical simulation: Using the compressed sensing state estimate $\hat{\varrho}_{\mathrm{CS}}$, we simulate measurement data, which we subsequently input to our compressed sensing reconstruction procedure, varying both ε, m and randomly drawing measurement settings without replacement. If the corresponding optimisation problem is infeasible and yields no estimate, the fidelity F is set to zero. The fidelities $F(\hat{\varrho}_{\mathrm{CS}}, \hat{\varrho}(m, \varepsilon))$ are averaged over data and measurement settings (500 different data sets and different measurement settings per m and ε).

The results for varying m and ε in units of $\hat{\varepsilon}$ are shown in figure 5. We compare the reconstructed states to $\hat{\varrho}_{\mathrm{CS}}$, which we used to generate the simulated data. We see that as m increases, the fidelity converges to unity at $\varepsilon = \hat{\varepsilon}$ (where $\hat{\varrho}_{\mathrm{CS}}$ is defined). We are interested in how quickly our reconstructed state approaches the optimal $\hat{\varrho}_{\mathrm{CS}}$ with fewer measurement settings, particularly if ε is misestimated. For instance, we see that we can obtain average fidelities of more than 0.8 for only 6 measurement settings. Figure 5 (top) again illustrates that $\varepsilon = \hat{\varepsilon}$ is the best choice as the fidelities around this region (and away from pathologically small numbers of measurement settings, $m > 3$) are the highest. Moreover, we also see that with increasing m, the standard deviation ΔF of the fidelity becomes smaller for $\varepsilon \geqslant \hat{\varepsilon}$. For $\varepsilon < \hat{\varepsilon}$, infeasibilities of the optimisation equation (12) that appear for certain choices of measurement settings lead to large standard deviations, which can be seen by the ridge in the area left of $\varepsilon = \hat{\varepsilon}$ in figure 5 (bottom). The ridge as well as the region of infeasibility gets close to $\varepsilon = \hat{\varepsilon}$ for large m, which is reasonable since more information (i.e. more constraints) puts greater restrictions on the optimisation problems. If fewer measurement settings are considered, as in the highly tomographically incomplete regime, overestimation of ε is less detrimental and state estimates still perform well, i.e. the fidelity is relatively constant for $\hat{\varepsilon} \lesssim \varepsilon \lesssim 3\,\hat{\varepsilon}$. However, as m increases, the reconstruction becomes more strongly dependant on the choice of ε. Generally, we see that the higher the fidelity, the lower the standard deviation.

Discussion

In this work, we have experimentally explored the compressed sensing paradigm for quantum state tomography as applied to the photonic setting. We have explicitly laid out a method for applying these techniques and reconstructed the state of a four-photon system with tomographically complete data available, observing a high fidelity of the reconstructed state with respect to the target state. The presence of noise in the data requires that one carefully chooses appropriate constraints on the optimisation. In current applications, these parameters are usually obtained in an ad hoc way. We have provided a prescription to establish the parameters in a more systematic way by modelling the noise and performing cross validation, which is a general method for model selection. The quality of the data, being afflicted with noise predominantly attributed to finite counting statistics, allows us to model the noise via a multinomial distribution. This is a situation commonly expected for photonic experiments with post-selected data. In fact, we observe a great agreement between estimating the model parameter from theoretical noise modelling and cross validation.

Having established the appropriate model, we have been able to perform state reconstruction with tomographically incomplete data, which rapidly converges to the highest fidelity estimate as the number of measurement settings increases. As a validity check, we have also run different estimators on the full data and obtained similar results, showing that our compressed sensing procedure yields reasonable estimates. As is predicted by the mathematical theory of compressed sensing, we have found that the number of measurement settings needed for a satisfactory estimate of the underlying state is much smaller than the number of measurements necessary for tomographic completeness. We have also carried out a comprehensive bootstrapping analysis to build confidence in the robustness of our method. In fact, we have observed that the uncertainty in the fidelity quickly decreases with increasing number of measurement settings.

Furthermore, we have studied the robustness of our method with respect to improper model selection and the effects on the reconstruction. We have found that for several choices of models and different numbers of measurement settings, the performance of the reconstruction can vary dramatically. For small numbers of measurement settings, our method depends less strongly on the model. In contrast, for large numbers of measurement settings, it is imperative to determine the appropriate model for optimal performance.

Our results confirm that compressed sensing in conjunction with suitable model selection gives rise to reliable procedures for state reconstruction leading to effective tomography with tomographically incomplete data. These techniques can be applied to a wide range of experimental settings and provide a means to identify and verify appropriate models thereby paving the way for the future of practical quantum state tomography. With this, we contribute to establishing compressed sensing as a practical tool for quantum state tomography in the low-information regime.

Acknowledgments

We thank the DFG (EI 519/7-1, EI 519/9-1 within SPP 1798 CoSIP), the Templeton Foundation, the EU (AQuS, SIQS, RAQUEL), EU FP7 grant 600838 QWAD, U.S. Army Research Office (ARO) grant W911NF-14-1-0133, the BMBF (Q.com), the Freie Universität Berlin within the Excellence Initiative of the German Research Foundation, the South African Research Foundation, the Fritz Haber Institute of the Max Planck Society, and the German National Academic Foundation (Studienstiftung des Deutschen Volkes) for support.

Appendix: fidelity estimation with error bound

In this section, we provide more detail to the fidelity estimation with an error bound without the need of resorting to quantum state tomography. In the Pauli operator basis

$$\{O_l : \ O_l \in \bigotimes_{j=1}^{n} \{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}\}, \tag{A1}$$

we can estimate from the measured probabilities $\hat{p}_{j,k} = y_{j,k}/N_j$ the expansion coefficients

$$\xi_\varrho^l = \mathrm{tr}(\varrho \ O_l/\sqrt{d}) \tag{A2}$$

of the prepared state ϱ by a linear transformation Ω,

$$\boldsymbol{\xi}_\varrho = \Omega \, \hat{\boldsymbol{p}}. \tag{A3}$$

For convenience, we denote by $\hat{\boldsymbol{p}}$ the row-vectorisation of the matrix with entries $\hat{p}_{j,k}$. The fidelity with respect to a pure target state ϱ_T can be written in terms of the expansion coeffcients as

$$F^2(\varrho_{\mathrm{T}}, \varrho) = \sum_l \xi^l_{\varrho_{\mathrm{T}}} \xi^l_{\varrho} = \boldsymbol{\xi}^T_{\varrho_{\mathrm{T}}} \Omega\, \hat{\boldsymbol{p}}. \tag{A4}$$

The frequency of the d different outcomes for the jth measurement setting is described by a multinomial distribution. The covariance matrix is given for each multinomial distribution by

$$\mathrm{Cov}(Y_{j,k}, Y_{j,l}) = N_j (p_{j,k}\, \delta_{i,j} - p_{j,k}\, p_{j,l}). \tag{A5}$$

Since different measurement settings correspond to mutually orthonormal operators, the frequencies of different measurement settings are uncorrelated, i.e. $\mathrm{Cov}(Y_{i,k}, Y_{j,l}) = 0$ for $i \neq j$. Therefore the covariance matrix for the probabilities $\hat{\boldsymbol{p}}$ can be estimated from the data as

$$\mathrm{Cov}(\hat{p}_{j,k}, \hat{p}_{j,l}) = N_j^{-1} (\hat{p}_{j,k} \delta_{i,j} - \hat{p}_{j,k} \hat{p}_{j,l}). \tag{A6}$$

By means of linear error propagation, the variance of the fidelity is given by

$$\mathrm{Var}(F^2) = \boldsymbol{\xi}^T_{\varrho_{\mathrm{T}}}\, \Omega\, \mathrm{Cov}(\hat{p}, \hat{p}) \Omega^T\, \boldsymbol{\xi}_{\varrho_{\mathrm{T}}}, \tag{A7}$$

which yields an estimate of the statistical error of the fidelity estimate from the data

$$\Delta F^2(\varrho, \varrho_{\mathrm{T}}) = \sqrt{\mathrm{Var}(F^2)}. \tag{A8}$$

In particular, in order to estimate the fidelity with respect to the GHZ state, only nine Pauli basis measurements contribute. This can be seen from the expansion of the GHZ density matrix in the Pauli operator basis

$$\varrho_{\mathrm{GHZ}} = \frac{1}{16}\left[\sum_{\sigma\in\{\mathbb{1},\sigma_x,\sigma_y,\sigma_z\}} \sigma^{\otimes 4} + \sum_{\mathrm{Perm.}} \mathbb{1}\otimes\mathbb{1}\otimes\sigma_z\otimes\sigma_z + \sum_{\mathrm{Perm.}} \sigma_x\otimes\sigma_x\otimes\sigma_y\otimes\sigma_y\right], \tag{A9}$$

where the last two sums run over all six distinct orders of the four factors of the tensor product.

To estimate the fidelity (A4), only the 16 Pauli coefficients of the prepared state are required that correspond to the operators of the expansion (A9). From the measurement outcomes of the measurement setting $\sigma_z^{\otimes 4}$, all coefficients of operators containing only the identity $\mathbb{1}$ and σ_z can be estimated. Thus, only nine Pauli basis measurements are necessary to estimate the fidelity.

Note that it is also possible to employ the measurement outcomes of all other measurement settings in the estimation of coefficients of terms that include the identity in equation (A9). In principle, it is thereby possible to further reduce the statistical error of the estimate of those coefficients. However, for the data set considered in this work, using more than nine measurement settings does not significantly alter the fidelity estimate.

References

[1] O'Brien J L, Furusawa A and Vucovic J 2009 Photonic quantum technologies *Nature Phot.* **3** 687–95
[2] Metcalf B J *et al* 2013 Multi-photon quantum interference in a multi-port integrated photonic device *Nature Comm.* **4** 1356
[3] Peruzzo A, Laing A, Politi A, Rudolph T and O'Brien J L 2011 Multimode quantum interference of photons in multiport integrated devices *Nature Comm.* **2** 224
[4] Monz T, Schindler P, Barreiro J T, Chwalla M, Nigg D, Coish W A, Harlander M, Haensel W, Hennrich M and Blatt R 2011 14-qubit entanglement: creation and coherence *Phys. Rev. Lett.* **106** 130506
[5] Barends R *et al* 2014 Logic gates at the surface code threshold: superconducting qubits poised for fault-tolerant quantum computing *Nature* **508** 500–3
[6] Nielsen M A and Chuang I L 2000 *Quantum Computation and Quantum Information* (Cambridge, UK: Cambridge University Press)
[7] Pirandola S, Eisert J, Weedbrook C, Furusawa A and Braunstein S L 2015 Advances in quantum teleportation *Nature Photon.* **9** 641–52
[8] Bell B, Clark A S, Tame M S, Halder M, Fulconis J, Wadsworth W J and Rarity J G 2012 Experimental characterization of photonic fusion using fiber sources *New J. Phys.* **14** 023021
[9] Bell B A, Herrera-Martí D A, Tame M S, Markham D, Wadsworth W J and Rarity J G 2014 Experimental demonstration of a graph state quantum error-correction code *Nature Comm.* **5** 3658
[10] Bell B A, Markham D, Herrera-Martí D A, Marin A, Wadsworth W J, Rarity J G and Tame M S 2014 Experimental demonstration of graph-state quantum secret sharing *Nat. Commun.* **5** 5480
[11] Candès E J, Romberg J and Tao T 2006 Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information *IEEE Trans. Inf. Theor.* **52** 489–509
[12] Foucart S and Rauhut H 2013 *A Mathematical Introduction to Compressive Sensing* (Heidelberg: Springer)
[13] Gross D, Liu Y-K, Flammia S T, Becker S and Eisert J 2010 Quantum state tomography via compressed sensing *Phys. Rev. Lett.* **105** 150401
[14] Flammia S T, Gross D, Liu Y-K and Eisert J 2012 Quantum tomography via compressed sensing: error bounds, sample complexity, and efficient estimators *New J. Phys.* **14** 095022
[15] Guta M, Kypraios T and Dryden I 2012 Rank-based model selection for multiple ions quantum tomography *New J. Phys.* **14** 105002
[16] Cramer M, Plenio M B, Flammia S T, Somma R, Gross D, Bartlett S D, Landon-Cardinal O, Poulin D and Liu Y-K 2010 Efficient quantum state tomography *Nature Comm.* **1** 149
[17] Huebener R, Mari A and J Eisert 2013 A Wick's theorem for matrix product states *Phys. Rev. Lett.* **110** 040401
[18] Steffens A, Friesdorf M, Langen T, Rauer B, Schweigler T, Hübener R, Schmiedmayer J, Riofrío C A and Eisert J 2015 Towards experimental quantum field tomography with ultracold atoms *Nature Comm.* **6** 7663
[19] Schwemmer C, Toth G, Niggebaum A, Moroder T, Gross D, Gühne O and Weinfurter H 2014 Experimental comparison of efficient tomography schemes for a six-qubit state *Phys. Rev. Lett.* **113** 040503

[20] Blume-Kohout R 2012 Robust error bars for quantum tomography arXiv:1202.5270
[21] Christandl M and Renner R 2012 Reliable quantum state tomography *Phys. Rev. Lett.* **109** 120403
[22] Faist P and Renner R 2016 Practical and reliable error bars in quantum tomography *Phys. Rev. Lett.* **117** 010404
[23] Carpentier A, Eisert J, Gross D and Nickl R 2015 Uncertainty quantification for matrix compressed sensing and quantum tomography problems ArXiv e-prints arXiv:1504.03234
[24] Rodionov A V, Veitia A, Barends R, Kelly J, Sank D, Wenner J, Martinis J M, Kosut R L and Korotkov A N 2014 Compressed sensing quantum process tomography for superconducting quantum gates *Phys. Rev.* B **90** 144504
[25] Kliesch M, Kueng R, Eisert J and Gross D 2016 Improving compressed sensing with the diamond norm *IEEE Trans. Inf. Th.* **62** 7445–63
[26] Shabani A, Kosut R L, Mohseni M, Rabitz H, Broome M A, Almeida M P, Fedrizzi A and White A G 2011 Efficient measurement of quantum dynamics via compressive sensing *Phys. Rev. Lett.* **106** 100401
[27] Riofrío C A, Gross D, Flammia S T, Monz T, Nigg D, Blatt R and Eisert J 2016 Experimental quantum compressed sensing for a seven-qubit system arXiv:1608.02263
[28] Carmeli C, Heinosaari T, Schultz J and Toigo A 2014 Tasks and premises in quantum state determination *J. Phys. A: Mathematical and Theoretical* **47** 075302
[29] Kalev A, Kosut R L and Deutsch I H 2015 Quantum tomography protocols with positivity are compressed sensing protocols *NPJ Quantum Information* **1** 15018
[30] Donoho D L 2006 Compressed sensing *IEEE Trans. Inf. Theor.* **52** 1289–306
[31] Candès E J and Recht B 2012 Exact matrix completion via convex optimization *Commun. ACM* **55** 111–9
[32] Candès E J and Tao T 2010 The power of convex relaxation: near-optimal matrix completion *IEEE Trans. Inf. Theor.* **56** 2053–80
[33] Gross D 2011 Recovering low-rank matrices from few coefficients in any basis *IEEE Trans. Inf. Theor.* **57** 1548–66
[34] Vandenberghe L and Boyd S 1996 Semidefinite programming *SIAM Rev.* **38** 49–95
[35] Fazel M 2002 Matrix rank minimization with applications *PhD Thesis* Stanford University
[36] Vandenberghe L and Boyd S 2004 *Convex Optimization* (New York: Cambridge University Press)
[37] Tropp J A 2015 Convex recovery of a structured signal from independent random linear measurements *Sampling Theory, a Renaissance* (Verlag: Birkhaeuser)
[38] Kueng R, Rauhut H and Terstiege U 2014 Low rank matrix recovery from rank one measurements arXiv:1410.6913
[39] Heinosaari T, Mazzarella L and Wolf M M 2013 Quantum tomography under prior information *Commun. Math. Phys.* **318** 355–74
[40] Kueng R and Gross D 2015 Qubit stabilizer states are complex projective 3-designs arXiv:1510.02767
[41] Liu Y-K 2011 Universal low-rank matrix recovery from Pauli measurements *Adv. Neural Inf. Process. Syst.* **24** 1638–46
[42] Rockafellar R T 1997 Convex analysis *Princeton Landmarks in Mathematics and Physics* (Princeton, NJ: Princeton University Press)
[43] Toh K C, Todd M J and Tütüncü R H 1999 SDPT3–a matlab software package for semidefinite programming *Optimization Methods and Software* **11** 545–81
[44] Grant M and S Boyd 2014 CVX: matlab software for disciplined convex programming, version 2.1, cvxr.com/cvx
[45] Cai J-F, Candès E J and Shen Z 2010 A singular value thresholding algorithm for matrix completion *SIAM J. Opt.* **20** 1956–82
[46] Bell B A, Tame M S, Clark A S, Nock R W, Wadsworth W J and Rarity J G 2013 Experimental characterization of universal one-way quantum computing *New J. Phys.* **15** 053030
[47] Browne D E and Rudolph T 2005 Resource-efficient linear optical quantum computation *Phys. Rev. Lett.* **95** 010501
[48] Gross D, Kieling K and Eisert J 2006 Potential and limits to cluster state quantum computing using probabilistic gates *Phys. Rev.* A **74** 042343
[49] Pittman T B, Jacobs B C and Franson J D 2001 Probabilistic quantum logic operations using polarizing beam splitters *Phys. Rev.* A **64** 062311
[50] Pan J-W, Simon C, Brukner C and Zeilinger A 2001 *Nature* **410** 1067–70
[51] Pan J-W, Gasparoni S, Ursin R, Weihs G and Zeilinger A 2003 Experimental entanglement purification of arbitrary unknown states *Nature* **423** 417–22
[52] Bodiya T P and Duan L-M 2006 Scalable generation of graph-state entanglement through realistic linear optics *Phys. Rev. Lett.* **97** 143601
[53] Bell B, McMillan A, McCutcheon W and Rarity J 2015 Effects of self- and cross-phase modulation on photon purity for four-wave-mixing photon pair sources *Phys. Rev.* A **92** 053849
[54] Burnham K P and Anderson D R 2003 *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (New York: Springer)
[55] Akaike H 1974 A new look at the statistical model identification *IEEE Trans. Aut. Cont.* **19** 716–23
[56] Schwarz G E 1978 Estimating the dimension of a model *Ann. Stat.* **6** 461–64
[57] Efron B and Tibshirani R J 1994 *An Introduction to the Bootstrap* (New York: Chapman and Hall/CRC Press)
[58] Suess D, Rudnicki Ł and Gross D 2016 Error regions in quantum state tomography: computational complexity caused by geometry of quantum states arXiv:1608.00374
[59] Hradil Z 1997 Quantum-state estimation *Phys. Rev.* A **55** R1561–4

3 QUANTUM FIELD TOMOGRAPHY

The *compressed sensing* techniques that where discussed in the last section allow for the robust recovery of low-rank (i.e., less mixed) quantum states, decreasing the amount of measurements to be taken and hence the experimental effort roughly by a square root factor in the Hilbert space dimension. This approach is feasible for intermediate-sized quantum systems, however, it reaches its inevitable limits for even larger systems, let alone continuous systems such as quantum fields.

In contrast, *tensor network states* allow for incorporating physical a priori information into the class of ansatz states for quantum tomography, making not only an exponential data reduction possible in various instances, but even allow for practical parametrizations of continuous quantum systems. Such a parametrization is represented by the *continuous matrix product states*, which will be discussed in section 3.2, after an introduction of tensor network states in section 3.1. Building on this, we have developed a protocol for the *tomography of quantum fields*, which is presented in the subsequent publication [2][12]. This protocol can be used to determine the states in experimental systems, which will be demonstrated in publication [3] in section 3.3. See also publication [5] in appendix A.1 for an application of *continuous matrix product states* in quantum transport experiments.

3.1 Tensor network states

Even with non-relativistic treatment, information in quantum systems propagates with finite speed, the *Lieb-Robinson velocity* v_{LR}. The respective, system-dependent upper bounds for the velocity are called *Lieb-Robinson bounds*.[98,99] This results in a certain locality and *clustering of correlations* for ground states of *gapped*[13] quantum lattice systems: Namely, the correlations of two local observables O_A and O_B at different subsystems A and B decay exponentially instead of algebraically with their distance $\mathrm{dist}(\mathrm{A},\mathrm{B})$:[100]

$$|\langle O_A O_B\rangle - \langle O_A\rangle\langle O_B\rangle| \leq C\,\mathrm{e}^{-\mathrm{dist}(A,B)\,\Delta E/(2v_{\mathrm{LR}})}||O_A||\,||O_B||. \tag{16}$$

With $\langle O\rangle$ we denote the expectation value of the observable O. This behavior is closely related to the fact that for many quantum systems, quantum correlations between a subsystem A and the complement subsystem B will grow not with the volume of A, but with the size of its lower-dimensional boundary ∂A: The *entanglement entropy* $S(\varrho_A) := -\mathrm{tr}(\varrho_A \log_2 \varrho_A)$[101],[14], which measures the

[12] Adrian Steffens, Carlos A. Riofrío, Robert Hübener, and Jens Eisert, "Quantum field tomography", *New Journal of Physics* 16:123010, 2014 (DOI:10.1088/1367-2630/16/12/123010). Published under a Creative Commons Attribution 3.0 License (creativecommons.org/licenses/by/3.0), © 2014 IOP Publishing.

[13] A quantum system is called *gapped* if there is a nonzero distance ΔE between its ground state energy and the energy of the first excited state.

[14] With ϱ_A, we denote the density matrix of the reduced subsystem A, after tracing out the complement system B, i.e., taking the partial trace of ϱ with respect to B: $\varrho_A := \mathrm{tr}_B(\varrho)$.

degree of *entanglement*[15] between A and B, quantifying their quantum correlations, scales linearly with the size $|\partial A|$ of ∂A

$$S(\varrho_A) = O(|\partial A|), \tag{17}$$

satisfying an *area law*.[102] Using Lieb-Robinson bounds, area laws have, e.g., been shown for one-dimensional gapped systems with a unique ground state.[103]

Importantly, the states that exhibit only comparably few and rather local correlations out of the many that can be modeled by the entire Hilbert space represent a large part of quantum states of interest appearing in nature, and they are confined to a tiny subset of the Hilbert space: The so-called *physical corner of Hilbert space*[47,104]. In the most extreme case, taking, e.g., a system $\mathscr{H} = \bigotimes_{j=1}^{n} \mathbb{C}^2$ of n spins, each in $\mathbb{C}^2$, with no correlations between the local spin systems, the state of the system is a *product state* $|\psi\rangle = \bigotimes_{j=1}^{n} |\varphi_j\rangle$ of the local spin states $|\varphi_j\rangle$. Only $O(n)$ parameters suffice to describe all such states—compared to the exponentially many of $\mathscr{H}$. It is relatively easy to experimentally prepare such uncorrelated states. Starting from these states, it was shown that it is hard using only local Hamiltonians to generate states that exhibit as much correlations as the *mathematically typical* state in $\mathscr{H}$ has: If one understands the preparation of arbitrary states as a sequence of local interactions within a *quantum circuit* with a product state as input, then such a circuit would also need to consist of exponentially many elements to produce most typical states.[104] Hence, in practice, large parts of the Hilbert space remain unreachable and—from a signal processing perspective—constitute an unused data overhead.

Conversely, many states that appear in nature can be parametrized using quantum circuits with $O(\operatorname{poly} N)$ elements, i.e., the physical corner can be covered by polynomially many parameters, and the suitable set of states to describe it in a compressed way is the set of *tensor network states*: A general pure state $|\psi\rangle$ in a Hilbert space $\mathscr{H} = \bigotimes_{j=1}^{n} \mathbb{C}^d$ can be written as

$$|\psi\rangle = \sum_{j_1,\dots,j_n=1}^{d} c^{j_1,\dots,j_n} |j_1\rangle \otimes \cdots \otimes |j_n\rangle, \tag{18}$$

where the coefficient array $c := (c^{j_1,\dots,j_n})$ can be understood as a tensor of order n and local dimension d. Decomposing c into a set of m lower-order tensors $c^{[1]}, \dots, c^{[m]}$,

$$c^{j_1,\dots,j_n} = \sum_{\{a_{k_l}^{[l]}\}} c^{[1]j_1,\dots,j_{l_1}}_{a_1^{[1]},\dots,a_{k_1}^{[1]}} \; c^{[2]j_{l_1+1},\dots,j_{l_2}}_{a_1^{[2]},\dots,a_{k_2}^{[2]}} \; \dots \; c^{[m]j_{l_{m-1}+1},\dots,j_{l_m}}_{a_1^{[m]},\dots,a_{k_m}^{[m]}} \tag{19}$$

[15] Fixing bases $\{|\varphi\rangle_A\}$ and $\{|\varphi\rangle_B\}$ of A and B, respectively, a general pure state in $\mathscr{H}_A \otimes \mathscr{H}_B$ can be written as $|\psi\rangle = \sum_{j,k} c_{jk} |\varphi_j\rangle_A \otimes |\varphi_k\rangle_B$. It is called *separable* if there exist two vectors a and b such that $c_{jk} = a_j b_k$ for all j and k, resulting in a *product state* $|\psi\rangle = \sum_{j,k} a_j b_k |\varphi_j\rangle_A \otimes |\varphi_k\rangle_B = |\chi\rangle_A \otimes |\chi\rangle_B$ with $|\chi\rangle_A = \sum_j a_j |\varphi_j\rangle_A$ and $|\chi\rangle_B = \sum_j b_j |\varphi_j\rangle_B$. In this case, there are no quantum correlations between A and B. If this decomposition into a product of subsystems does not exist, the state is called *entangled* and measurements on one subsystem affect measurements on the other. That notion of *bipartite* entanglement—for a system consisting of two subsystems—can be extended to *multipartite* entanglement as well as to mixed states.

results in a *tensor network*. The *physical indices* $j_1, \ldots, j_n$ get distributed among the m tensors, which are connected to each other via contracted shared indices $\{a_{k_l}^{[l]}\}$. The dimension of each contracted index connecting two tensors is called its *bond dimension*. A quantum state with a tensor network decomposition of the coefficient array c as in Eq. (19) is called a *tensor network state*[105]. For illustration, see Fig. 5.

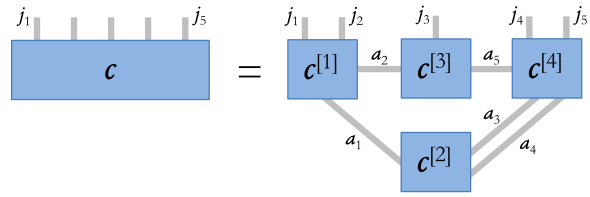

Figure 5: *Exemplary decomposition of a 5th-order tensor into a tensor network* $\sum_{a_1,\ldots,a_5} c_{a_1}^{[1]j_1,j_2}\ c_{a_1,a_3,a_4}^{[2]}\ c_{a_2,a_5}^{[3]j_3}\ c_{a_3,a_4,a_5}^{[4]j_4,j_5}$. *Here, we make use of a common graphical representation for tensor networks. Each box corresponds to a tensor, each line to an index. A line connecting two boxes corresponds to contracting the common index of the two respective tensors. The dimension of each internal index a_l is called its bond dimension. Unconnected, open lines (such as the ones labeled with $j_1, \ldots, j_5$) correspond to uncontracted indices.*

A particularly successful class of tensor network states for analyzing one-dimensional systems constitute the *matrix product states* (MPS)[106–108], which provide exact ground states of 1D quantum Heisenberg spin model extensions, the *AKLT model*[109] and the *Majumdar–Ghosh model*[110], classifications of quantum phases[111] and characterizations of quantum phase transitions[112]. Important algorithms for finding ground states and simulating time evolution such as the *density matrix renormalization group* (DMRG)[113] and *time-evolving block decimation* (TEBD)[114] are closely related to MPS. For reviews, see for example Refs. [115, 116]. Applications beyond quantum physics—e.g., for machine learning or the solution of partial differential equations—are being developed as well[117,118], under the heading of *tensor trains*[119].

Other types of tensor networks include *projected entangled pair states*[120] (PEPS), which generalize MPS to higher spatial dimensions, *tree tensor network states*[121], and *multiscale entanglement renormalization ansatz* (MERA) states[122], which are particularly practical for describing ground states of gapless Hamiltonians. Mixed states can be represented by *matrix product operators*[123] (MPO).

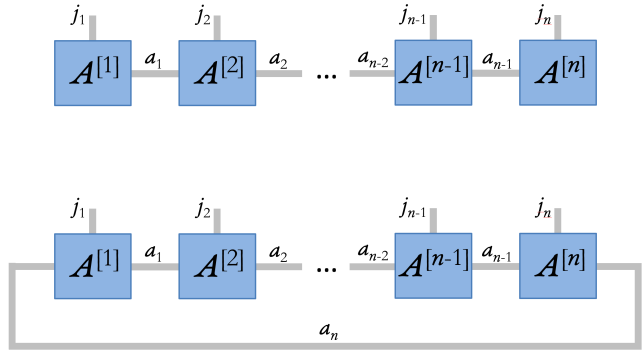

Figure 6: *Graphical representation of an* MPS *with open boundary conditions (above) and periodic boundary conditions (below). Above, the tensors* $\{A^{[1]j_1}\}$ *and* $\{A^{[n]j_n}\}$ *have only one internal index* a_1 *and* a_{n-1}*, respectively. Hence, they respectively correspond to row and column vectors.*

For MPS, the tensor c is factorized into products of matrices $A^{[l]j_l}$,

$$\begin{aligned} c^{j_1,\ldots,j_n} &= \sum_{a_1,\ldots,a_{n-1}} A^{[1]j_1}_{a_1} A^{[2]j_2}_{a_1,a_2} A^{[3]j_3}_{a_2,a_3} \ \ldots \ A^{[n-1]j_{n-1}}_{a_{n-2},a_{n-1}} A^{[n]j_n}_{a_{n-1}} \\ &= A^{[1]j_1} A^{[2]j_2} A^{[3]j_3} \ldots A^{[n-1]j_{n-1}} A^{[n]j_n}, \end{aligned} \tag{20}$$

resulting in a *matrix product state*

$$|\psi\rangle = \sum_{j_1,\ldots,j_n=1}^{d} A^{[1]j_1} A^{[2]j_2} \ldots A^{[n]j_n} |j_1,\ldots,j_n\rangle, \tag{21}$$

see Fig. 6 for illustration.[16] Each family of matrices $A^{[l]j_l}$ can, e.g., be associated with the l-th site of a one-dimensional spin chain. Matrix product states can be generated from a general state by sequentially regrouping the indices $j_1,\ldots,j_n$ of c, obtaining matrices $c^{j_1,(j_2,\ldots,j_n)}$, $\tilde{c}^{(j_1,j_2),(j_3\ldots,j_n)}$, etc., successively performing singular value decompositions on them and combining the resulting matrices, ultimately yielding the desired matrices $\{A^{[l]j_l}\}$.[115] Using this prescription, the magnitude of the number of parameters in c is reflected in an exponential increase of the bond dimension with increasing l up until $l = \lceil n/2 \rceil$. Specifically (for even n) the dimensions of the matrices

[16] In Eq. (21), we are given an MPS with *open* boundary conditions and the arrays $\{A^{[1]j_1}\}$ correspond to row vectors and $A^{[n]j_n}$ to column vectors. Contracting the first and the last site results in an MPS with *periodic* boundary conditions,

$$|\tilde{\psi}\rangle = \sum_{j_1,\ldots,j_n} \mathrm{tr}\left(A^{[1]j_1} A^{[2]j_2} \ldots A^{[n]j_n}\right) |j_1,\ldots,j_n\rangle. \tag{22}$$

$A^{[1]j_1},\ldots,A^{[n]j_n}$ are $1\times d, d\times d^2,\ldots,d^{n/2-1}\times d^{n/2}, d^{n/2}\times d^{n/2-1},\ldots,d^2\times d, d\times 1$. The crucial point is that for many physical states of interest, the spectra arising in the sequential singular value decompositions decay exponentially—which can be attributed to area laws[108]—and truncations of the singular spectra yield good approximations of the original state while massively reducing the number of required parameters: If for each site the bond dimension is fixed to b, we obtain $(n-2)d$ matrices in $\mathbb{C}^{b\times b}$ and two vectors in $\mathbb{C}^b$ (for the boundaries) with the total number of parameters scaling *linearly* instead of exponentially with the number of sites n. Tuning the bond dimension of an MPS allows for balancing out the computational operability with how well it describes the given physical system.

There exist quantum tomography protocols that allow for efficiently—i.e., using polynomially many parameters and still capturing the essential physics—attributing an MPS to a discrete complex quantum system.[48,124] This is already much more efficient than the more general compressed sensing regime, as discussed in the previous section. For continuous systems such as quantum fields, however, new concepts have to be introduced. This is the topic of the publications [2] and [3], which will be discussed in the following.

3.2 Continuous matrix product states for quantum tomography

The continuous analogues of MPS are the *continuous matrix product states*[125,126] (CMPS). These allow for an efficient characterization of one-dimensional quantum fields and were shown to accurately describe continuous models as, e.g., the *Lieb-Liniger model*[125,127], its fermionic equivalent, the *Gaudin-Yang model*[128,129], or the relativistic *Gross-Neveu model*[130]. There exist algorithms for efficiently computing the parameter sets of a CMPS approximation of the *Lieb-Liniger* ground state[131,132], but the involved methods are not restricted to this model. Continuous analogues of PEPS can be formulated as well[133], however, as in the discrete case, expectation values, which would be important for tomographic protocols, are in general computationally intractable.

A CMPS on an interval $[0,L]$ with periodic boundary conditions is defined as

$$|\Psi_{Q,R}\rangle = \mathrm{tr}_{\mathrm{aux}}\left(\mathscr{P}\,\mathrm{e}^{\int_0^L \mathrm{d}x\left(Q(x)\otimes\hat{\mathbb{1}}+R(x)\otimes\hat{\psi}^\dagger(x)\right)}\right)|\Omega\rangle\,. \tag{23}$$

Its characteristic parameters are contained within the matrix families $\{Q(x)\in\mathbb{C}^{b\times b}|x\in[0,L]\}$ and $\{R(x)\in\mathbb{C}^{b\times b}|x\in[0,L]\}$ with bond dimension b. These are coupled with the unit operator $\hat{\mathbb{1}}$ and field operators[17] $\hat{\psi}(x)$, respectively, that act on the Fock vacuum state $|\Omega\rangle$ within a path-ordered exponential $\mathscr{P}\,\mathrm{e}$, after which the $(b\times b)$-dimensional matrix space—the "auxiliary space"—is traced out. The matrices $R(x)$ can be interpreted as scattering matrices that result in a particle at position x, while the matrices $Q(x)$ can be related to free propagation.

[17] The field operators $\hat{\psi}(x)$ obey the canonical (anti-)commutation relations $[\hat{\psi}(x),\hat{\psi}^\dagger(y)]_\pm = \delta(x-y)$ for fermionic/bosonic fields. The definition in Eq. (23) can be extended to multiple fields by using sums of different field operators together with corresponding R matrices, $\sum_j R_j(x)\otimes\hat{\psi}_j(x)$, in the exponent.

Continuous matrix product states can be constructed as the continuum limit of MPS of the form

$$|\Psi_\varepsilon\rangle = \sum_{j_1,\ldots,j_n} \mathrm{tr}(A^{[1]j_1},\ldots,A^{[n]j_n})\,(\hat{a}_1^\dagger)^{j_1}\ldots(\hat{a}_n^\dagger)^{j_n}\,|\Omega_n\rangle \tag{24}$$

on discrete lattices $\mathscr{L}_\varepsilon \subset [0, L]$ with lattice parameter ε, $n = L/\varepsilon$ sites, site-dependent creation operators $\{\hat{a}_j^\dagger\}$ obeying $[\hat{a}_j, \hat{a}_k^\dagger]_\pm = \delta_{j,k}$, and matrices $\{A^{[k]j_k} \in \mathbb{C}^{b\times b}\}$. Setting

$$A^{[k]0} = \mathbb{1}_b + \varepsilon Q(k\,\varepsilon), \quad A^{[k]j_k} = \varepsilon^{j_k/2} R^{j_k}(k\,\varepsilon), \quad \hat{a}_k^\dagger = \varepsilon^{1/2}\hat{\psi}^\dagger(k\,\varepsilon) \tag{25}$$

and taking the limits $\varepsilon \to 0$, $n(\varepsilon) = L/\varepsilon \to \infty$ yields Eq. (23).[134] An equivalent construction of CMPS from continuous measurements is also possible.[126]

For translation invariant CMPS, the matrices $Q(x)$ and $R(x)$ become independent of the position x and the state is completely parametrized by $2\,b^2$ complex numbers. An important feature of translation invariant CMPS is that n-point correlation functions of the type

$$C(x_1,\ldots,x_n) = \langle\Psi_{Q,R}\,|\,\hat{\psi}^\dagger(x_1)\ldots\hat{\psi}^\dagger(x_m)\,\hat{\psi}(x_{m+1})\ldots\hat{\psi}(x_n)\,|\,\Psi_{Q,R}\rangle \tag{26}$$

can be reduced to 2- and 3-point correlation functions[135] and can be calculated in closed form as sums of exponentially damped sinusoids,

$$C(x_1,\ldots,x_n) = \sum_{j_1,\ldots,j_{n-1}=1}^{b^2} r_{j_1,\ldots,j_{n-1}}\,\mathrm{e}^{\lambda_{j_1}(x_2-x_1)}\ldots\mathrm{e}^{\lambda_{j_{n-1}}(x_n-x_{n-1})}, \tag{27}$$

with $r_{j_1,\ldots,j_{n-1}}, \lambda_{j_k} \in \mathbb{C}$, implying that only certain 2- and 3-point correlation functions—as opposed to arbitrarily large n for general field states[136]—are required to in principle determine the CMPS[135]. This makes these tensor network states particularly interesting for practical quantum tomography. As in many inverse problems, the quantities r and λ, however, are not straightforwardly and, due to various gauge degrees of freedom, not uniquely related to the parameter matrices Q and R. Since the λ values are also not linearly related to C, it is moreover not clear how to efficiently obtain r and λ from an experimentally sampled data array C. The tomographic protocol at the basis of the following publication [2] is therefore performed in several consecutive reconstruction steps from processing sampled correlation functions to computing the CMPS parameters in a feasible manner. To this end, state-of-the-art signal processing algorithms for precise reconstructability—*Prony*-like analysis[70,137] and the *matrix pencil method* [138,139]—were adapted and extended for handling input signals of arbitrary dimension. Implementing a non-linear least squares based optimization heuristic is in general not productive because of the non-convex nature of the problem and the disadvantageous scaling of the computational effort with the number of parameters, i.e., the bond dimension b. An important application of this protocol was to experimentally show that CMPS can be employed to adequately describe a continuous system that appears in nature. This will be

presented in publication [3] in the subsequent section 3.3. Apart from quantum state tomography, CMPS can also be used to characterize quantum transport experiments (see the publication [5] in appendix A.1).

FIRST-AUTHORED PUBLICATIONS

[1] Experimentally exploring compressed sensing quantum tomography
Adrian Steffens, Carlos A. Riofrío, Will McCutcheon, Ingo Roth, Bryn A. Bell, Alex McMillan, Mark S. Tame, John G. Rarity, and Jens Eisert, *Quantum Science and Technology* 2:025005, 2017.

[2] Quantum field tomography
Adrian Steffens, Carlos A. Riofrío, Robert Hübener, and Jens Eisert, *New Journal of Physics* 16:123010, 2014.

[3] Towards experimental quantum-field tomography with ultracold atoms
Adrian Steffens, Mathis Friesdorf, Tim Langen, Bernhard Rauer, Thomas Schweigler, Robert Hübener, Jörg Schmiedmayer, Carlos A. Riofrío, and Jens Eisert, *Nature Communications* 6:7663, 2015.

[4] An efficient quantum algorithm for spectral estimation
Adrian Steffens, Patrick Rebentrost, Iman Marvian, Jens Eisert, and Seth Lloyd *New Journal of Physics* 19:033005, 2017.

New Journal of Physics
The open access journal at the forefront of physics
Deutsche Physikalische Gesellschaft DPG | IOP Institute of Physics

Quantum field tomography

A Steffens, C A Riofrío, R Hübener and J Eisert
Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, D-14195 Berlin, Germany
E-mail: adrian.steffens@fu-berlin.de

Received 8 August 2014, revised 20 October 2014
Accepted for publication 31 October 2014
Published 2 December 2014
New Journal of Physics **16** (2014) 123010
doi:10.1088/1367-2630/16/12/123010

Abstract
We introduce the concept of quantum field tomography, the efficient and reliable reconstruction of unknown quantum fields based on data of correlation functions. At the basis of the analysis is the concept of continuous matrix product states (cMPS), a complete set of variational states grasping states in one-dimensional quantum field theory. We innovate a practical method, making use of and developing tools in estimation theory used in the context of compressed sensing such as Prony methods and matrix pencils, allowing us to faithfully reconstruct quantum field states based on low-order correlation functions. In the absence of a phase reference, we highlight how specific higher order correlation functions can still be predicted. We exemplify the functioning of the approach by reconstructing randomized cMPS from their correlation data and study the robustness of the reconstruction for different noise models. Furthermore, we apply the method to data generated by simulations based on cMPS and using the time-dependent variational principle. The presented approach is expected to open up a new window into experimentally studying continuous quantum systems, such as those encountered in experiments with ultra-cold atoms on top of atom chips. By virtue of the analogy with the input–output formalism in quantum optics, it also allows for studying open quantum systems.

Keywords: quantum state tomography, estimation methods, quantum fields, entanglement, tensor network methods, continuous matrix product states

New Journal of Physics **16** (2014) 123010
1367-2630/14/123010+32$33.00

1. Introduction

Quantum theory predicts probability distributions of outcomes in anticipated quantum measurements. The actual problem encountered in practice, however, is often not so much concerned with predicting certain outcomes of specific measurement procedures, but rather with reconstructing the unknown quantum state at hand that is compatible with precisely such measurement outcomes. This task of reconstructing states based on data—possible under certain conditions of completeness or other reasonable assumptions—is called *quantum state tomography*. For finite-dimensional quantum systems, this task is feasible and is routinely used in experiments. However, the number of parameters to be determined scales exponentially with the system size: full quantum state tomography is highly inefficient. This is even so much less of a problem than one might at first be tempted to think. It was one of the major insights in the field in recent years to recognize that economical or efficient quantum state tomography is distinctly possible for systems with many degrees of freedom. In fact, in most physically relevant questions, fully unconstrained quantum state tomography may be said to solve the 'wrong problem'. One is surely often not interested in arbitrary states, but only in those states that one is expected to encounter in practice, which are naturally more restricted.

In the context of *compressed sensing tomography* [9, 14] or *matrix product states* (MPS) tomography [1, 5, 34], identification of quantum systems with many degrees of freedom is indeed possible. The key step is to identify the right model in which to represent the states, e.g., approximately low-rank states or those with clustering correlation functions. In the context of MPS tomography, the notion of a model refers to a meaningful variational class of states that provably captures all states exhibiting low entanglement [8, 41]. In this sense, tomography is efficiently possible for any system size. In fact, by increasing the bond dimension, an arbitrary state can be well approximated. Quite similar to the mindset of compressed sensing, a 'sparsity of commonly encountered states' is heavily used for the benefit of tomography.

In quantum field theory, where one has to consider an infinite number of degrees of freedom, the situation is in principle aggravated. Analogously, a moment of thought reveals that to think about quantum field tomography in the sense of trying to 'fill an infinite table with numbers' is rather ill-guided. This is not the actual problem one aims at solving in any practical context—one again needs to identify the appropriate model and the right 'sparsity structure'.

In this work, we introduce the concept of *quantum field tomography*, tomography of continuous systems in quantum field theory, and provide a practical and feasible method for achieving this. We do so by drawing and further developing ideas from the study of *continuous matrix product states* (cMPS) [16, 35, 47], methods of how to assess higher order correlation functions in that context [24], as well as a machinery from statistical estimation theory, such as a *Prony analysis* [38] and *matrix pencil methods* (MPM) [21, 22], which are here brought to a new context. In fact, these methods of estimation have not been considered before in the context of quantum state reconstruction and are expected to be interesting in their own right. The basis of the analysis are low-order multi-point correlation functions directly accessible in many common current experiments.

This approach opens up a new window into grasping the physics of continuous quantum systems in equilibrium and non-equilibrium. Instead of having to make a physical model (e.g., define a Hamiltonian) and checking for the plausibility of it, one can—based on data of correlation functions—reconstruct the quantum field itself. Such an approach seems particularly appealing when studying one-dimensional continuous bosonic models such as *ultra-cold atoms*

on top of atom chips [13, 27, 28]. What is more, if only partial data is available, say, in the absence of a phase reference frame, higher-order correlation functions of the same type can be predicted as well. The starting point of the analysis is what is called 'Wick's theorem for MPSs' [24], which is here brought to a new level and transformed into a practical method of reconstructing unknown cMPS from correlation function data.

This work is structured as follows. In section 2, we will give a short overview of the concept of cMPS [16, 35, 47] as well as what can be called a 'Wick theorem' for this class of states [24], aiming as a preparation for the following technical sections. In section 3, we will describe in great technical detail how to reconstruct a field state from its low order correlation functions and give a complete MPS description of it. The limitations of this method are investigated in section 4. In section 5, we will demonstrate the method using simulated data from random cMPS and apply the method to the ground state of the Lieb–Liniger model, a prototypical integrable model in quantum field theory [3, 30]. The data used here have been generated using a cMPS-based simulation based on the time-dependent variational principle [6, 15, 17]. The impact of noise in real world-scenarios on the method is investigated here. In section 6, we summarize and conclude this work.

2. Background

In this work, we are concerned with one-dimensional quantum fields with fast decaying spatial correlations. Analogous to the case of many-body quantum systems, successfully described by the MPS formalism, there is a variational class of states specially suited to study such systems: the cMPS [35, 47].

2.1. cMPS

In this section, we briefly review the basics of the cMPS formalism. For a review and comprehensive discussion of the computation of correlation functions, see, e.g., [16].

2.1.1. Basic definitions. A translationally invariant cMPS with periodic boundary conditions and one species of bosonic particles is defined as

$$\left|\psi_{Q,R}\right\rangle = \mathrm{Tr}_{\mathrm{aux}}\left[\mathcal{P}\mathrm{e}^{\int_0^L \mathrm{d}x\left(Q\otimes\hat{1}+R\otimes\hat{\Psi}^\dagger(x)\right)}\right]|\Omega\rangle, \tag{1}$$

where the collection of field operators $\hat{\Psi}(x)$, $x \in [0, L]$, obey the bosonic commutation relations of the free field

$$\left[\hat{\Psi}(x), \hat{\Psi}^\dagger(y)\right] = \delta(x-y), \tag{2}$$

$|\Omega\rangle$ is the vacuum state vector, Q, $R \in \mathbb{C}^{d\times d}$ are matrices acting on an *auxiliary d-dimensional space* $\mathcal{A}$, the 'virtual space', and constitute the variational parameters of the class. L is the length of the closed physical system, $\mathcal{P}$ denotes the path ordering operator and $\mathrm{Tr}_{\mathrm{aux}}$ traces out the auxiliary space.

The parametrization in (1) by Q and R is not unique, i.e., there is and additional *gauge freedom*. Namely, when simultaneously conjugating Q and R with an invertible matrix G [16],

$$\tilde{Q} = G^{-1}QG, \tag{3}$$

$$\tilde{R} = G^{-1}RG, \tag{4}$$

then the two resulting state vectors still represent the same state, i.e., all expectation values are invariant under this transformation.

2.1.2. Related physical processes. A useful interpretation of the correlations in cMPS can be given in terms of a d-dimensional (auxiliary) quantum system $\mathcal{A} \cong \mathbb{C}^d$ interacting with a one-dimensional field $\mathcal{F}$ [35]. The Hamiltonian of the joint system is given by

$$\hat{H}(x) = K \otimes \hat{\mathbb{1}}_{\mathcal{F}} + R \otimes \hat{\Psi}^{\dagger}(x) + R^{\dagger} \otimes \hat{\Psi}(x), \tag{5}$$

where $\hat{\mathbb{1}}_{\mathcal{F}}$ is the identity on the field, $K \in \mathbb{C}^{d\times d}$ the Hamiltonian of the free evolution of the finite dimensional system, and $R \otimes \hat{\Psi}^{\dagger}(x)$ the coupling between the system and the field with $R \in \mathbb{C}^{d\times d}$. Note that H evolves in position, rather than time—in this picture, both are by construction equivalent. Starting with the state vector $|\varphi_i\rangle|\Omega\rangle$, where $|\varphi_i\rangle \in \mathcal{A}$ and the vacuum $|\Omega\rangle \in \mathcal{F}$, and evolving over $[0, L] \ni x$, we formally arrive at

$$\hat{U}(0, L)\,\Big|\varphi_i\Big\rangle|\Omega\rangle := \mathcal{P}\mathrm{e}^{-\mathrm{i}\int_0^L \mathrm{d}x\left(K\otimes\hat{\mathbb{1}}_{\mathcal{F}} - \frac{1}{2}R^{\dagger}R\otimes\hat{\mathbb{1}}_{\mathcal{F}} + \mathrm{i}R\otimes\hat{\Psi}^{\dagger}(x)\right)}\Big|\varphi_i\Big\rangle|\Omega\rangle, \tag{6}$$

using the Baker–Campbell–Hausdorff formula and the fact that $R^{\dagger} \otimes \hat{\Psi}(x)|\varphi_i\rangle|\Omega\rangle = 0$. By setting

$$Q = -\mathrm{i}K - \frac{1}{2}R^{\dagger}R, \tag{7}$$

projecting onto $\langle\varphi_i| \otimes \hat{\mathbb{1}}_{\mathcal{F}}$ to decouple $\mathcal{A}$ from $\mathcal{F}$, and summing over a complete orthonormal basis of all $|\varphi_i\rangle$, we again obtain equation (1). This shows the interpretation of the cMPS formalism in the sequential preparation picture of MPS [36].

In this picture, we interpret K to be the Hamiltonian of a virtual particle in the auxiliary space that mediates field interactions. Even more [35], the dynamical behaviour of the auxiliary system $\mathcal{A}$ can be modelled by computing the derivative of

$$\rho_{\mathcal{A}}(x) = \mathrm{Tr}_{\mathcal{F}}\Big[\hat{U}(x, L)\Big(\rho_{\mathcal{A}}(0) \otimes |\Omega\rangle\langle\Omega|\Big)\hat{U}^{\dagger}(x, L)\Big], \tag{8}$$

where $\mathrm{Tr}_{\mathcal{F}}$ means tracing out the physical system $\mathcal{F}$. This yields the ordinary differential equation

$$\frac{\mathrm{d}}{\mathrm{d}x}\rho_{\mathcal{A}}(x) = -\mathrm{i}[K, \rho(x)] + R^{\dagger}\rho(x)R - \frac{1}{2}\Big[R^{\dagger}R, \rho(x)\Big]_{+}, \tag{9}$$

which is a master equation in Lindblad form, governing the Markovian evolution of $\rho_{\mathcal{A}}$, where R plays the role of dissipative quantum jump (Lindblad) operators. Although arbitrary Q and R lead to a valid cMPS, not all pairs give rise to an effective Hamiltonian K via equation (7). For this, it is required that

$$Q + Q^{\dagger} + R^{\dagger}R = 0. \tag{10}$$

However, arbitrary Q and R can in general be transformed into a specific gauge where they fulfil this equation.

2.2. Correlation functions in cMPS

The mathematical relations between the n-point functions are the starting point for our tomography algorithms, hence we give a brief summary at this point. A quantum field state can be completely characterized by all the possible normal expectation values constructed from $\hat{\Psi}(.)$ and $\hat{\Psi}^{\dagger}(.)$ and their commutation relations. In this work, we will focus on *density-like* correlation functions, i.e., for each position $x_k \in [0, L]$, $k = 1,...,n$, both operators $\hat{\Psi}^{\dagger}(x_k)$ and $\hat{\Psi}(x_k)$ exist within the expectation values. Because of translational invariance, we can set $x_1 = 0$ without loss of generality. The expectation value $\langle\psi_{Q,R}|\hat{\Psi}^{\dagger}(x_1)...\,\hat{\Psi}^{\dagger}(x_n)\hat{\Psi}(x_n)...\,\hat{\Psi}^{\dagger}(x_1)|\psi_{Q,R}\rangle$ can be computed as

$$C^{(n)}(\tau_1,...,\tau_{n-1}) := \mathrm{Tr}\Big[\,\mathrm{e}^{T\tau_n}\big(\overline{R}\otimes R\big)...\,\mathrm{e}^{T\tau_2}\big(\overline{R}\otimes R\big)\mathrm{e}^{T\tau_1}\big(\overline{R}\otimes R\big)\Big], \tag{11}$$

(see, e.g., [16]), with the transfer matrix

$$T := \overline{Q}\otimes\mathbb{1}_d + \mathbb{1}_d\otimes Q + \overline{R}\otimes R, \tag{12}$$

and the positive distances $\tau_j = x_{j+1} - x_j$ for $j = 1,...,n-1$ and $\tau_n = L - x_n$; the overline denotes complex conjugation. Correlation functions of cMPS are given by expressions involving only the auxiliary space. Static properties of a quantum field with one spatial dimension are hence related to non-equilibrium properties of a zero-dimensional system. In this sense, they have been referred to as being 'holographic quantum states' [35].

For a normalized cMPS, the eigenvalues of T are all complex with negative or zero real parts, due to the analogy to quantum channels [48]. This leads to finite expectation values in the thermodynamic limit $L\to\infty$. Furthermore, assuming that T is diagonalizable, which is in particular the case if its spectrum is non-degenerate, the n-point function (11) can be further simplified to a sum of exponentially damped oscillatory terms

$$\lim_{L\to\infty} C^{(n)}(\tau_1,...,\tau_{n-1}) = \sum_{k_1,...,k_{n-1}=1}^{d^2} \rho_{k_1,k_2,...,k_{n-1}}\,\mathrm{e}^{\lambda_{k_1}\tau_1}...\,\mathrm{e}^{\lambda_{k_{n-1}}\tau_{n-1}}, \tag{13}$$

where

$$\rho_{k_1,k_2,...,k_{n-1}} = M_{1,k_{n-1}}M_{k_{n-1},k_{n-2}}...\,M_{k_1,1}. \tag{14}$$

The matrix $M\in\mathbb{C}^{d^2\times d^2}$ is defined as $M = X^{-1}\big(\overline{R}\otimes R\big)X$, where X is a change-of-basis matrix such that $X^{-1}TX$ is diagonal and compatible with the ordering of the eigenvalues $\{\lambda_k\}$. In the following, we will work exclusively in the thermodynamic limit and, for simplicity, use $C^{(n)}$ also to denote n-point correlation functions in this limit.

A first step to reconstruct a cMPS would be to identify $\{\rho_{k_1,k_2,...,k_{n-1}}\}$ and $\{\lambda_k\}$. That this is in principle possible can be seen by considering the Laplace transform of $C^{(n)}$

$$\mathcal{L}^{(n)}(\mathbf{s}) = \int_0^{\infty} \mathrm{d}^{n-1}\boldsymbol{\tau}\,\mathrm{e}^{-\mathbf{s}\cdot\boldsymbol{\tau}}C^{(n)}(\boldsymbol{\tau}), \qquad s_1,...,s_{n-1}\in\mathbb{C}, \tag{15}$$

which has the simple form

$$\mathcal{L}^{(n)}(\mathbf{s}) = \sum_{k_1,\ldots,k_{n-1}=1}^{d^2} \frac{\rho_{k_1,k_2,\ldots,k_{n-1}}}{\left(\lambda_{k_1} - s_1\right)\cdots\left(\lambda_{k_{n-1}} - s_{n-1}\right)}. \tag{16}$$

Each of the $d^{2(n-1)}$ combinations of T eigenvalues appears as a pole of $\mathcal{L}^{(n)}$ in $\mathbb{C}^{n-1}$ together with the corresponding residue in the numerator. If all the eigenvalues are different, i.e., the spectrum of T non-degenerate, and all residues non-zero, then all residues are distinguishable as well. Since the Laplace transform itself proved to be infeasible for practical reconstruction algorithms, we will present alternative ways in the following. Independently of this, we want to keep calling the eigenvalues $\{\lambda_k\}$ the poles and $\{\rho_{k_1,k_2,\ldots,k_{n-1}}\}$ the residues of the n-point function. In the following, we require the spectrum of T to be non-degenerate.

The structure of the correlation functions with the residues as products of entries of one matrix, equation (14), allows for expressing higher order correlation functions by lower order correlation functions, very much reminding of the Wick's theorem in quantum field theory [24]. In this sense, we will recover M from the residues. We will describe this in detail below.

2.3. Additional symmetries

In the remainder of this work, we will make use of some symmetries that the cMPS fulfil. Here, we briefly state them. By construction, for each non-real entry of $\overline{R} \otimes R$ and T there exists another entry containing its complex conjugate. More precisely, one can show that

$$\Lambda_d \overline{\overline{R} \otimes R} \Lambda_d = \overline{R} \otimes R \tag{17}$$

and $\Lambda_d \overline{T} \Lambda_d = T$, with

$$\Lambda_d := \sum_{j,k=1}^{d} E_{j,k} \otimes E_{k,j} \tag{18}$$

and $E_{j,k} = e_j\, e_k^T$, the dyadic product of the canonical column vectors e_j, [12, section 2.5]. Hence, if λ is an eigenvalue of T with eigenvector v then $\Lambda_d \overline{T} \Lambda_d v = \lambda v$, and since $(\Lambda_d)^2 = \mathbb{1}_{d^2}$, we obtain $T(\Lambda_d \overline{v}) = \overline{\lambda}\,(\Lambda_d \overline{v})$, such that the spectrum of T is closed under complex conjugation. This fact also follows from the channel property of cMPS as discussed in [48].

For the reconstruction algorithms we will discuss below, it is instrumental to fix an unambiguous ordering of the eigenvalues of the transfer matrix T, which makes its diagonal matrix D and furthermore the matrix M unambiguous, too. If we order the eigenvalues in D such that the $\kappa \in \{1,\ldots,d^2\}$ real eigenvalues constitute a block and the remaining $d^2 - \kappa$ are arranged in complex conjugate pairs (e.g., ordering by descending real part), then D obeys the symmetry relation $\Xi_{d,\kappa} \overline{D} \Xi_{d,\kappa} = D$ with the permutation matrix

$$\Xi_{d,\kappa} := \mathbb{1}_\kappa \oplus \left(\bigoplus_{j=1}^{(d^2-\kappa)/2} \sigma_x \right), \tag{19}$$

where σ_x is the x-Pauli matrix. In addition, since X consists of the eigenvectors v of T as column vectors, $\Lambda_d \overline{v}$ is the eigenvector of $\overline{\lambda}$, when v corresponds to λ. Moreover, since $\Xi_{d,\kappa}$ interchanges the columns back, we have that $\Lambda_d \overline{X} \Xi_{d,\kappa} = X$. Using this fact and the definition $M = X^{-1}\overline{R} \otimes RX$, we obtain the symmetry relation $\Xi_{d,\kappa} \overline{M} \Xi_{d,\kappa} = M$ for the matrix M. This

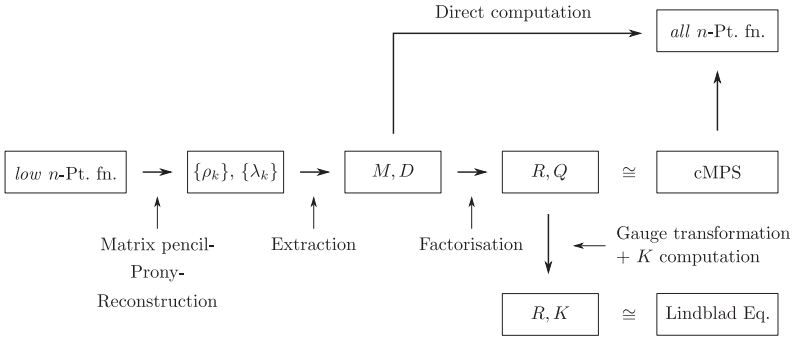

Figure 1. The particular reconstruction steps starting with the input data, an n-point correlation function of a cMPS $|\Psi_{Q,R}\rangle$, and ending with the variational parameter matrices Q and R, that fully characterize the state. Alternatively, the state can likewise be described by K and R. With this knowledge, one can compute other n'-point correlation functions and compare with the input data to obtain evidence for a successful reconstruction.

relation connects each entry of M with its complex conjugate and, via equation (14), each residue with its complex conjugate. As with the poles, the set of residues is closed under complex conjugation for density-like correlation functions. These symmetries can also be used for a systematic least squares approach to reconstruct the poles and residues, see section 3.2.

3. State reconstruction

Having established the structure of the correlation functions in cMPS, i.e., the structure of the data of our reconstruction problem, it remains to develop an appropriate protocol to extract the information encoded in the data. Given an n-point density-like correlation function of order 3 or higher corresponding to a cMPS $|\Psi_{Q,R}\rangle$, we will show that, in most cases, it is in principle possible to reconstruct the parameter matrices Q and R up to an arbitrary gauge and phase, and to reproduce all n-point functions.

We are dealing with a so-called inverse problem, a large class of problems that make 'use of the actual results of some measurements of the observable parameters to infer the actual values of the model parameters' [44]. Many inverse problems are ill-conditioned—a small change in the measurements can lead to a huge change in the model parameters. In this chapter we will examine the required steps for cMPS reconstruction, see figure 1, and the respective main factors that influence their performance regarding perturbed input data. Each step will be discussed in a separate section. We will see that in particular the first and the last step can be notably ill-conditioned.

3.1. Reconstruction steps

The reconstruction of a generic, translationally invariant cMPS in the thermodynamic limit comprises the following steps, which are represented in figure 1:

(1) The first step in processing the input data is to extract the poles $\{\lambda_k\}$ and the residues $\{\rho_{k_1,k_2,...,k_{n-1}}\}$ from a density-like n-point correlation function, $n \geqslant 3$,

$$C^{(n)}(\tau_1,...,\tau_{n-1}) = \sum_{k_1,...,k_{n-1}=1}^{d^2} \rho_{k_1,k_2,...,k_{n-1}} e^{\lambda_{k1}\tau_1} ... e^{\lambda_{k_{n-1}}\tau_{n-1}}, \tag{20}$$

which is measured and contains additional noise and experimental imperfections.

(2) In the second step, the matrix M is determined from the residues

$$\rho_{k_1,k_2,...,k_{n-1}} = M_{1,k_{n-1}} M_{k_{n-1},k_{n-2}} ... M_{k_1,1}, \tag{21}$$

and the matrix D is determined from the poles. This can be achieved using certain invariances in the correlation functions that led to the formulation of Wick's theorem for MPSs.

(3) In the final step, the cMPS parametrization matrices Q and R can be extracted from the matrices M and D by imposing a specific gauge. Additionally, and after another gauge transformation, the Hamiltonian K of the auxiliary system can be computed from the matrices Q and R.

In order to only generate and predict higher order density-like n-point functions, it is in general sufficient to use the matrices D and M from the second step without any further reconstruction steps. This is in general much more robust against noise than the full reconstruction. Furthermore, we can leave out some of the poles (together with the corresponding entries in M) that barely contribute to the n-point functions. We will follow this approach in accompanying work when analysing experimental data [43].

3.2. Reconstructing the poles and residues

When analysing spectra of sampled linear combinations of sinusoidal functions, methods based on integral transforms like the discrete Fourier transform seem like a natural choice. In our case, however, we deal with exponentially damped sinusoids with potentially similar frequencies, which results in heavy broadening and overlapping of the corresponding spectral peaks. In this case, the damping factors would have to be determined from the corresponding peaks' width, and, in view of experimental data, we cannot assume too many sampled data points. Hence, the spectral resolution would be rather low. Only for certain cases the peaks in the frequency spectrum are sufficiently separated to directly determine the poles in a feasible way using integral transforms.

Another class of methods for data fitting that may come to mind is based on nonlinear (e.g., least squares) minimization approaches. Clearly, the number of parameters critically determines the computational effort and the successful applicability of the algorithm. The results, however, can be improved by restricting ourselves to a likely parameter region as a result of a preceding Fourier transform. Taking into account the Λ_d and $\Xi_{d,\kappa}$ symmetries and assuming normalized n-point functions, the number of real parameters can be reduced to $nd^2 - 2$. Only for unambiguous global minima (which is usually not the case for high damping factors in combination with noise) and for very small bond dimension, we obtained satisfactory results in acceptable time. Least squares approaches for correlation functions with larger n are at best feasible when using Q and R as parameters, otherwise the number of parameters would become too large. In spite of these drawbacks, a least squares algorithm could be used as an additional refinement step with initial values from other procedures, like

the ones discussed below; nevertheless the number of parameters is still limiting. On the other hand, if we can only assume a small number of parameters and expect a considerable amount of noise, the least squares method can be a robust alternative. For example, for bond dimension $d = 2$, such nonlinear least squares approach can be feasibly and successfully used.

Alternative minimization methods, e.g., simulated annealing, did not lead to considerable improvements. However, the scaling of the computational effort with the number of parameters can be significantly mitigated using iterative quadratic maximum likelihood methods, but the application to correlation functions with $n > 2$ is not straightforward [19, section 1.2.3].

Realizing the challenges of solving a nonlinear estimation problem, it seems logical to exploit the structure of our particular model of the data to see if there are ways to more efficiently solve the estimation problem. It turns out that for data structures that consist of sums of damped oscillatory terms, it is possible to separate the estimation of poles and residues of the function in two different linear estimation processes. In the following sections, we describe two major approaches one can take to achieve such estimation.

3.2.1. Prony analysis. This technique is used in digital signal processing and its roots go back to a method that was originally established by Prony in 1795 in the context of fluids [38]. The main idea is to first recover the poles independently by determining the roots of a polynomial computed from the signal (the correlation function) and then to insert the poles into a system of linear equations for the coefficients, which is in principle solvable with the usual linear algebra procedures. Prony's method is a special case of *linear prediction* [19] and has many further applications, e.g., as the starting point for nearest-neighbour detection of atoms in optical lattices [26, 29]. The original method, however, is very sensitive to noise, so that for working on experimental data we need to use several modifications, which we will describe below. For further summaries and an introduction of the method, see for instance [18, 31, 37].

Prony's method is usually applied to $\mathbb{C}$-valued functions, corresponding to 2-point functions, and for our purposes has to be extended to work with higher order n-point functions, which can be done in a straightforward way. Therefore, in our description, we will start with the one-dimensional case with signal function

$$C^{(2)}(\tau) := \sum_{k=1}^{d^2} \rho_k \mathrm{e}^{\lambda_k \tau}. \tag{22}$$

The function is sampled at a finite number of points and is available only for $N + 1$ points $\{\tau_j\}$, which is $C^{(2)}(\tau_j) := C_j$, $j = 0, ..., N$. We, thus, obtain a system of linear equations

$$\rho_1 \mathrm{e}^{\lambda_1 \tau_0} + \cdots + \rho_{d^2} \mathrm{e}^{\lambda_{d^2} \tau_0} = C_0, \tag{23}$$

$$\rho_1 \mathrm{e}^{\lambda_1 \tau_1} + \cdots + \rho_{d^2} \mathrm{e}^{\lambda_{d^2} \tau_1} = C_1, \tag{24}$$

$$\vdots$$

$$\rho_1 \mathrm{e}^{\lambda_1 \tau_N} + \cdots + \rho_{d^2} \mathrm{e}^{\lambda_{d^2} \tau_N} = C_N. \tag{25}$$

Once we have identified all poles $\{\lambda_k\}$, we can easily solve this system and are finished with the reconstruction. As we will see, one requirement for Prony's method is to sample the signal at equidistant points $\tau_j = j \cdot \Delta\tau$, $j \in \{0,...,N\}$, and with $e^{\lambda_k \Delta\tau} =: \mu_k$ we arrive at

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \mu_1 & \mu_2 & \cdots & \mu_{d^2} \\ \vdots & \vdots & & \vdots \\ \mu_1^N & \mu_2^N & \cdots & \mu_{d^2}^N \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_{d^2} \end{pmatrix} = \begin{pmatrix} C_0 \\ C_1 \\ \vdots \\ C_N \end{pmatrix}, \tag{26}$$

where the poles are encoded in the (in general, non-square) Vandermonde matrix

$$\mathcal{V} := \left(\mu_k^{j-1}\right)_{\substack{j=1,...,N+1 \\ k=1,...,d^2}}. \tag{27}$$

We must take care not to choose the sampling interval $\Delta\tau$ too large, since, considering the Nyquist–Shannon sampling theorem [42], the sampling rate should in general be at least twice the highest frequency ω_{sup} of the signal spectrum $2\pi/(\Delta\tau) < 2\omega_{\text{sup}}$.

Vandermonde matrices will often be ill-conditioned—e.g., according to Gautschi [10], a lower bound for the norm of the inverse matrix of $\mathcal{V}$ (for $N = d^2$ and $\mathcal{V}$ invertible) is

$$\left\| \mathcal{V}^{-1} \right\|_\infty > \max_{1 \leqslant l \leqslant n} \prod_{\substack{m=1 \\ m \neq l}}^{d^2} \frac{\max\left(1, \left|\mu_m\right|\right)}{\left|\mu_l - \mu_m\right|}, \tag{28}$$

which will get very large if two poles get close to each other. This fact hints at the intrinsic limitations of this reconstruction method.

To determine the poles, we can regard the set $\{\mu_1,...,\mu_p\}$ as the roots of a polynomial $\mathcal{P}_{d^2}$ with real coefficients and degree d^2 in the variable z,

$$\begin{aligned} \mathcal{P}_{d^2}(z) &= \sum_{l=0}^{d^2} a_l z^l, \\ \mathcal{P}_{d^2}(z = \mu_k) &= 0 \end{aligned} \tag{29}$$

for each $k = 0,...,N$. Note that there are d^2 values of μ_k but $d^2 + 1$ of a_l. Such a polynomial naturally exists—it is just the product of the linear factors $(z - \mu_k)$,

$$\mathcal{P}_{d^2}(z) = \prod_{k=1}^{d^2} (z - \mu_k). \tag{30}$$

Our goal is to relate the set of coefficients $\{a_l\}$ to the set of function values $\{C_j\}$. Once we have all a_l, we can compute the roots of the corresponding polynomial (29) and obtain the poles $\lambda_k = \ln(\mu_k)/\Delta\tau$, $k = 1,...,d^2$. To this end, we multiply the first line of equation (26) by a_0, the second by a_1 and so on, and perform the sum,

$$\sum_{l=0}^{d^2} a_l C_l = \sum_{l=0}^{d^2} a_l \sum_{k=0}^{d^2} \rho_k \mu_k^l = \sum_{k=0}^{d^2} \rho_k \sum_{l=0}^{d^2} a_l \mu_k^l. \tag{31}$$

Now, by choice of the a_l, each μ_k is a root of $\mathcal{P}_{d^2}(z)$ for all $k = 1,...,d^2$ so that each sum over l in equation (31) vanishes. Accordingly, we see that

$$\sum_{l=0}^{d^2} a_l C_l = 0. \tag{32}$$

Since $\prod_{k=1}^{d^2}\left(z-\mu_k\right) = 1\cdot z^{d^2}+\cdots$, the coefficient a_{d^2} belonging to the highest power is equal to one. Hence, equation (32) becomes the recurrence relation

$$\sum_{l=0}^{d^2-1} a_l C_l = -C_n. \tag{33}$$

In order to compute the d^2 coefficients $\{a_0,...,a_{d^2-1}\}$, we need at least d^2 equations. More linear independent equations are easy to obtain because the argument in equation (31) is still valid if we shift C_l to C_{l+m} for any $m\in\mathbb{N}$ with $d^2+m\leqslant N$:

$$\begin{aligned}\sum_{l=0}^{d^2} a_l C_{l+m} &= \sum_{l=0}^{d^2} a_l \sum_{k=0}^{d^2} \rho_k {\mu_k}^{l+m} \\ &= \sum_{k=0}^{d^2} \rho_k {\mu_k}^{m}\left(\sum_{l=0}^{d^2} a_l {\mu_k}^{l}\right) = 0.\end{aligned} \tag{34}$$

For d^2 equations the largest index that appears is $2d^2-1$ and our equation system looks like

$$\begin{pmatrix} C_0 & C_1 & C_2 & \dots & C_{d^2-1} \\ C_1 & C_2 & C_3 & & \vdots \\ C_2 & C_3 & \ddots & & \vdots \\ \vdots & & & \ddots & C_{2d^2-3} \\ C_{d^2-1} & \dots & \dots & C_{2d^2-3} & C_{2d^2-2} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{d^2-1} \end{pmatrix} = -\begin{pmatrix} C_{d^2} \\ C_{d^2+1} \\ \vdots \\ C_{2d^2-1} \end{pmatrix}. \tag{35}$$

Therefore, for d^2 poles we need at least $2d^2$ sampling points $\{C_0,...,C_{2d^2-1}\}$. The square matrix on the left-hand side of equation (35) can be written as $(C_{j+k})_{j,k=0,...,d^2-1}$ and has the form of a Hankel matrix. If it is non-singular, the solution vector $(a_0,...,a_{d^2-1})^T$ is unique and can, together with $a_{d^2}=1$, directly be replaced in (29), which in turn will yield the d^2 poles in a unique way. Hence, when reconstructing a function with d^2 poles and residues, we need *precisely* $2d^2$ sampling points to *exactly* solve the Hankel and the Vandermonde system, provided that both matrices are not singular. This means that for small bond dimensions and without noise the necessary resolution of the signal for a complete reconstruction is very low.

There are many established criteria for the invertibility [25, section 18] and inversion algorithms [4, 45] of Hankel or Toeplitz matrices (equation (35) can also be rearranged as a Toeplitz system). They are known to be potentially ill-conditioned, which reflects the inverse nature of the problem, e.g., the spectral condition number of a real positive-definite $N\times N$ Hankel matrix is bounded from below by $3\cdot 2^{N-6}$ [46]. In practice, recovering the poles is more stable when oversampling the signal and using a higher pole estimate, i.e., working with a larger (not necessarily square) Hankel matrix and a larger solution vector in equation (35), and solving the equation system in a least squares sense. This boils down to applying the Moore–Penrose pseudoinverse to the right-hand side of equation (35) to obtain the coefficients of the polynomial, inserting the computed poles into equation (26) and discarding the $N+1-p$ surplus poles with the smallest associated residues.

Note that instead of solving equation (35), we can also determine the kernel of $(C_{j+k})_{j,k=0,...,d^2-1}$, whose dimension is larger or equal to one due to equation (32). Only in the latter case, which corresponds to the matrix in equation (35) being non-singular, we get a unique (up to multiplication by a constant) solution vector $(a_0,...,a_{d^2})^T$. The constant does not pose a problem because any multiple of $(a_0,...,a_{d^2})^T$ yields the same roots of the associated polynomial: $\sum_{l=0}^{d^2} \alpha\, a_l z^l = 0$ is equivalent to $\sum_{l=0}^{d^2} a_l z^l = 0$. This method has proven to be more robust towards noise in some cases [33] and can be generalized in an elegant way to higher order correlation functions [39].

Unfortunately, in many cases, Prony's method is highly susceptible to noise in the signal. However, it presents a beautiful framework that shows that, in principle, it is possible to reconstruct the poles and residues of a signal. Without noise, both poles and residues can be determined *exactly*. In the next section, we describe a better algorithm for solving this type of inverse problems, which is more stable for larger bond dimension and finer sample rates.

3.2.2. MPM. The original MPM was developed by Hua and Sarkar [21, 22] and can be directly applied to our problem. As with the Prony algorithm, the poles are determined first and independently from the residues. Although the MPM is related to Prony [40], it is considerably less sensitive to noise [19, section 1.2] and can deal with higher sampling rates in a more stable fashion. Once the poles are identified, the residues are found via a linear equation system in the same way as in Prony's method. Here, we will just describe how to determine the poles. For simplicity, we will begin with the case of reconstructing a 2-point function and generalize to higher order correlation functions in the following section.

A *matrix pencil* $\mathfrak{M}$ of degree $n \in \mathbb{N}$ is a polynomial over $\mathbb{C}$ with matrix valued coefficients $M_j \in \mathbb{C}^{d\times d}$, $\mathfrak{M}(\gamma) = \sum_{j=0}^{n} M_j \gamma^j$. As with the Prony algorithm, we start by forming the Hankel matrix

$$C^{[1]} := \begin{pmatrix} C_0 & C_1 & ... & C_{P-1} \\ C_1 & C_2 & ... & C_P \\ \vdots & \vdots & & \vdots \\ C_{N-P-1} & C_{N-P} & ... & C_{N-2} \end{pmatrix} \in \mathbb{C}^{(N-P)\times P}, \tag{36}$$

from the experimental data points $\{C_0,...,C_{N-2}\}$

$$C_j = \sum_{k=1}^{d^2} \rho_k e^{\lambda_k \Delta\tau \cdot j} = \sum_{k=1}^{d^2} \rho_k \mu_k^j, \tag{37}$$

with integers N, P, such that $N - P$, $P > d^2$. Generally, the larger the number of samples N, the better the estimation of poles becomes. The optimal value for P regarding noise sensitivity typically lies between $N/3$ and $N/2$ [23]. In this method, we make use of the fact that $C^{[1]}$ can always be decomposed as

$$C^{[1]} = \mathcal{V}_1 \mathcal{R} \mathcal{V}_2 \tag{38}$$

with Vandermonde matrices

$$\mathcal{V}_1 = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \mu_1 & \mu_2 & & \mu_{d^2} \\ \vdots & \vdots & & \vdots \\ \mu_1^{N-P-1} & \mu_2^{N-P-1} & \dots & \mu_{d^2}^{N-P-1} \end{pmatrix} \in \mathbb{C}^{(N-P)\times d^2} \tag{39}$$

and

$$\mathcal{V}_2 = \begin{pmatrix} 1 & \mu_1 & \dots & \mu_1^{P-1} \\ 1 & \mu_2 & & \mu_2^{P-1} \\ \vdots & \vdots & & \vdots \\ 1 & \mu_{d^2} & \dots & \mu_{d^2}^{P-1} \end{pmatrix} \in \mathbb{C}^{d^2\times P}, \tag{40}$$

and the diagonal matrix $\mathcal{R} = \mathrm{diag}\left(\rho_1,...,\rho_{d^2}\right)$, as can easily be verified by using equation (38). In addition to the Hankel matrix $C^{[1]}$, we construct a second Hankel matrix

$$C^{[2]} = \begin{pmatrix} C_1 & C_2 & \dots & C_P \\ C_2 & C_3 & \dots & C_{P+1} \\ \vdots & \vdots & & \vdots \\ C_{N-P} & C_{N-P+1} & \dots & C_{N-1} \end{pmatrix} \in \mathbb{C}^{(N-P)\times P}, \tag{41}$$

which in turn can be decomposed as

$$C^{[2]} = \mathcal{V}_1 \mathcal{R} \mathcal{V}_0 \mathcal{V}_2 \tag{42}$$

with $\mathcal{V}_0 = \mathrm{diag}\left(\mu_1,...,\mu_{d^2}\right)$, and consider the linear matrix pencil

$$C^{[2]} - \gamma C^{[1]} = \mathcal{V}_1 \mathcal{R} \left(\mathcal{V}_0 - \gamma \mathbb{1}_{d^2}\right) \mathcal{V}_2 \tag{43}$$

with $\gamma \in \mathbb{C}$. Since all μ_j of $\mathcal{V}_1$ and $\mathcal{V}_2$ are distinct for a non-degenerate spectrum of T and $N - L, L > d^2$, the matrices $\mathcal{V}_1$ and $\mathcal{V}_2$ have rank d^2 and we can see that

$$\mathrm{rank}\left(C^{[1]}\right) = \mathrm{rank}\left(C^{[2]}\right) = \mathrm{rank}\left(\mathcal{V}_1 \mathcal{R} \mathcal{V}_0 \mathcal{V}_2\right) = \mathrm{rank}(\mathcal{R}) = d^2. \tag{44}$$

Generically, the matrix pencil $C^{[2]} - \gamma C^{[1]}$ will have the same rank, except for $\gamma = \gamma_j \in \left\{\mu_1,...,\mu_{d^2}\right\}$. In that case, the jth row of $\left(\mathcal{V}_0 - \gamma \mathbb{1}_{d^2}\right)$ is zero, hence

$$\mathrm{rank}\left(C^{[2]} - \gamma C^{[1]}\right) = d^2 - 1, \tag{45}$$

and there exists a non-trivial vector v with

$$\left(C^{[2]} - \gamma C^{[1]}\right) v = 0. \tag{46}$$

In this form, the complex number γ can be regarded as a solution of the *generalized eigenvalue problem* (GEVP) (46). This means that the d^2 non-zero generalized eigenvalues of equation (46) are exactly the exponentiated poles $e^{\lambda_1 \Delta t},..., e^{\lambda_{d^2} \Delta t}$. Equation (46) can be solved by a generalized Schur decomposition of the matrix pair $\{C^{[2]}, C^{[1]}\}$ or by solving the ordinary eigenvalue problem

$$\left(C^{[1]}\right)^{+} C^{[2]} v = \gamma v \tag{47}$$

with the pseudoinverse $\left(C^{[1]}\right)^{+}$ of $C^{[1]}$ [22]. After having determined the poles this way, they can be inserted into a linear equation system to obtain the according residues, as with Prony's method.

3.2.3. Technical improvements. Several improvements can be made to the original MPM approach including features from other reconstruction methods, which led to algorithms like Pro-ESPRIT and TLS ESPRIT [23], which we mention for the sake of completeness. Modifications based on structured low rank approximations [2, 32] did not lead to significantly better results. Here, we will focus on the so-called state space MPM, which shows the highest robustness towards noise of all direct MPM descendants [19, 23] and is the one we prefer to implement.

In this context, we continue with equation (46), but instead of solving it directly, we perform additional noise filtering steps via SVD rank truncations [20]. Performing separate SVD truncations like in the original approach has proven to be less robust than performing a *joint* SVD on $C^{[1]}$ and $C^{[2]} \in \mathbb{C}^{(N-P)\times P}$ by

$$\left(C^{[1]}, C^{[2]}\right) = U\Sigma V^{\dagger} =: U\Sigma \left(V^{[1]\dagger}, V^{[2]\dagger}\right) \tag{48}$$

with a unitary matrix $U \in \mathsf{U}(N-P)$, $\Sigma \in \mathbb{C}^{(N-P)\times 2P}$ containing the singular values of the concatenated matrices $(C^{[1]}, C^{[2]}) \in \mathbb{C}^{(N-P)\times 2P}$, and $(V^{[1]\dagger}, V^{[2]\dagger}) \in \mathsf{U}(2P)$. Note that $V^{[1]}$ and $V^{[2]} \in \mathbb{C}^{P\times 2P}$ are not unitary, in contrast to the matrix $(V^{[1]\dagger}, V^{[2]\dagger})$, and are not directly related to the unitary matrices from the separate SVDs. We insert equation (48) into equation (46), yielding

$$\left(C^{[2]} - \gamma C^{[1]}\right) v = U\Sigma \left(V^{[2]\dagger} - \gamma V^{[1]\dagger}\right) v, \tag{49}$$

and see that if γ is a generalized eigenvalue of the matrix pair $\{V^{[2]\dagger}, V^{[1]\dagger}\}$, then so it is of $\{C^{[2]}, C^{[1]}\}$. Hence, we can just work with $\{V^{[2]\dagger}, V^{[1]\dagger}\}$ (or $\{V^{[2]}, V^{[1]}\}$ since the set of poles of our n-point functions is to be closed under complex conjugation), and can completely forget about the singular values in Σ. We now filter the signal given in equation (48) by keeping the d^2 largest singular values and the corresponding singular vectors of $V^{[1]\dagger}$ and $V^{[2]\dagger}$:

$$U\Sigma \left(V^{[1]\dagger}, V^{[2]\dagger}\right) \overset{\text{trunc}}{\longmapsto} \left(V_T^{[1]\dagger}, V_T^{[2]\dagger}\right). \tag{50}$$

The GEVP we want to solve now is

$$\left(V_T^{[2]} - \gamma' V_T^{[1]}\right) v = 0, \tag{51}$$

with the filtered eigenvalues $\gamma' \in \mathbb{C}$. Since $V_T^{[1]}, V_T^{[2]} \in \mathbb{C}^{P\times d^2}$ and $P \gg d^2$, there is still surplus information we can use to SVD filter equation (51) one more time. For higher robustness, we repeat the truncation process, applying it to the concatenated matrix $(V_T^{[1]}, V_T^{[2]}) \in \mathbb{C}^{P\times 2d^2}$,

$$\left(V_T^{[1]}, V_T^{[2]}\right) = U'\Sigma' \left(V'^{[1]\dagger}, V'^{[2]\dagger}\right) \overset{\text{trunc}}{\longmapsto} U'_T \, \Sigma'_T \left(V_T'^{[1]\dagger}, V_T'^{[2]\dagger}\right) \tag{52}$$

with $U' \in \mathsf{U}(P)$, $\Sigma' \in \mathbb{C}^{P\times 2d^2}$, $V' \in \mathsf{U}(2d^2)$, $V'_T \in \mathbb{C}^{d^2\times 2d^2}$ and $V_T'^{[1]}, V_T'^{[2]} \in \mathbb{C}^{d^2\times d^2}$. Equation (51) then becomes

$$V_T^{[2]} - \gamma' V_T^{'[1]} = U' \Sigma \left(V^{'[2]\dagger} - \gamma' V^{'[1]\dagger} \right) \mapsto U'_T \, \Sigma'_T \left(V_T^{'[2]\dagger} - \gamma'' V_T^{'[1]\dagger} \right) \quad (53)$$

with the doubly SVD filtered eigenvalues $\gamma'' \in \mathbb{C}$. If there is no noise, then all the d^2 generalized eigenvalues of the matrix pencil $\{V_T^{'[2]}, V_T^{'[1]}\}$ are generalized eigenvalues of $\{V^{'[2]}, V^{'[1]}\}$, thus generalized eigenvalues of $\{C^{[2]}, C^{[1]}\}$ and nothing else than the exponentiated poles $e^{\lambda_1 \Delta\tau}, ..., e^{\lambda_{d^2} \Delta\tau}$. With noise, we can assume that the filtered set of eigenvalues $\{\gamma''\}$ provide a better estimate than the unfiltered $\{\gamma\}$ [20, 23]. Since $V_T^{'[1]\dagger}$ is invertible by construction, everything boils down to solving an ordinary eigenvalue problem:

$$\left(V_T^{'[1]} \right)^{-1} V_T^{'[2]} v = \gamma'' v. \quad (54)$$

This concludes the description of the state space MPM, which is our preferred technique for pole reconstruction.

3.2.4. Generalization to higher dimensions. So far, we have developed the reconstruction techniques for 2-point correlation functions. In this section, we show how to deal with higher order functions and generalize the previous discussion. Additionally, we show how one can improve the signal-to-noise ratio by exploiting redundant information in the higher order correlation functions.

If, for an n-point function, we uniformly sample each tensor index with N sampling points, we obtain a $(n-1)$-dimensional array $\left(C_{l_1,...,l_{n-1}} \right)_{l_1,...,l_{n-1}=0,...,N-1} \in \mathbb{C}^{N^{n-1}}$ with

$$C_{l_1,...,l_{n-1}} = \sum_{k_1,...,k_{n-1}=1}^{d^2} \rho^{(n)}_{k_1,...,k_{n-1}} e^{\lambda_{k1} l_1 \Delta\tau} ... \, e^{\lambda_{kn-1} l_{n-1} \Delta\tau}. \quad (55)$$

To extract the poles, we carry forward the approach of Zhu and Hua [49, chapter 17.11]. We fix one index l_j of $C_{l_1,...,l_{n-1}}$ and sum over the other indices

$$\hat{C}^{(j)}_{l_j} := \sum_{\substack{\{l_i\}=0, \\ i \neq j}}^{N-1} C_{l_1,...,l_{n-1}}. \quad (56)$$

The summing provides averaging and hence increases noise stability. This procedure is only possible because the poles and the sampling interval are the same for each index of the n-point function data array. Inserting the definition for $C_{l_1,...,l_{n-1}}$ and separating $e^{\lambda_{kj} l_j \Delta\tau}$ from the summation of k_j yields

$$\hat{C}^{(j)}_{l_j} = \sum_{k_j=1}^{d^2} \check{C}^{(j)}_{k_j} e^{\lambda_{kj} l_j \Delta\tau} \quad (57)$$

with

$$\check{C}^{(j)}_{k_j} = \sum_{\substack{\{k_i\}=1,...,d^2, \\ \{l_i\}=0,...,N-1 \\ i \neq j}} \rho_{k_1,...,k_{n-1}} e^{\lambda_{k1} l_1 \Delta\tau} ... e^{\lambda_{kj-1} l_{j-1} \Delta\tau} e^{\lambda_{kj+1} l_{j+1} \Delta\tau} ... e^{\lambda_{kn-1} l_{n-1} \Delta\tau}. \quad (58)$$

Equation (57) can be be regarded as the components of a 2-point function with the sought-after poles and $\{\check{C}^{(j)}_{k_j}\}$, which only depend on k_j, as its residues. The concrete values of these effective

residues do not matter, since in this step we are only interested in the poles. We can average further by summing the vectors $(\hat{C}^{(j)}_{l_j})_{l_j=0,\ldots,N-1}$, each corresponding to the tensor direction j, which leads to the N-component vector

$$\left(\hat{C}_l\right)_l := \left(\hat{C}_l^{(1)}\right)_l + \left(\hat{C}_l^{(2)}\right)_l + \cdots + \left(\hat{C}_l^{(n-1)}\right)_l. \tag{59}$$

The counting indices $\{l_j\}$ do not depend on j, hence we omitted the j for clearness.

The vector $(\hat{C}_l)$ still corresponds to a 2-point function with the correct poles and we can now apply the established matrix pencil, Prony or a least squares method to obtain the poles. Additionally, the averaging results in an effective reduction of the standard deviation of the (white) noise by a factor of $((n-1)N^{n-1})^{-1}$. Regarding the residues, we can reshape the array of the poles into a matrix and obtain the residues as the solution vector of the corresponding linear equation system in the least squares sense.

3.3. Extracting M

After having determined the poles and residues of the input correlation function—our first reconstruction step as discussed in section 3.1—the next step is to identify the matrix M. From M together with D, the variational parameter matrices R and Q can be determined.

First, we note that conjugating M with a diagonal matrix whose first entry is equal to one does not change the density-like correlation functions. This observation can be used to require that $M_{1,j} = 1$ for $j = 2,\ldots,d^2$, which is possible if the $M_{1,j}$ are non-zero. For $M_{1,1}$ to be equal to one, we need to normalize the n-point function by dividing by

$$\left\langle \Psi_{Q,R} \middle| \hat{\Psi}^\dagger \hat{\Psi} \middle| \Psi_{Q,R} \right\rangle^n = M^n_{1,1}. \tag{60}$$

In particular, we obtain $\rho^{(n)}_{k_1,1,\ldots,1} = 1\cdots\cdots 1 \cdot M_{k_1,1} = \rho^{(2)}_{k_1}$. For clearness, in this section we mark the dimensions of the residues with an additional index. We can compute $M_{i,j}$ for any $i, j = 1,\ldots,d^2$ and $n \geqslant 3$ via

$$\frac{\rho^{(n)}_{j,i,1,\ldots,1}}{\rho^{(n)}_{j,1,\ldots,1}} = \frac{\rho^{(3)}_{j,i}}{\rho^{(2)}_j} = \frac{M_{i,j}M_{j,1}}{M_{j,1}} = M_{i,j}. \tag{61}$$

From this equation we can see that we need n to be larger than three, since a 2-point function can at best provide the first column of M.

In practice, we may want to reduce noise by averaging over multiple independent prescriptions for $M_{i,j}$, namely

$$M_{i,j} = \frac{1}{d^{2(n-3)}} \sum_{k_1,\ldots,k_{n-3}=1}^{d^2} \frac{\rho^{(n)}_{k_1,\ldots,k_{n-3},j,i}}{\rho^{(n)}_{k_1,\ldots,k_{n-3},j,1}}. \tag{62}$$

By rearranging the residues, we can express higher order expectation values in terms of lower order:

$$\begin{aligned}\rho^{(n)}_{k_1,\ldots,k_{n-1}} &= M_{1,k_{n-1}} M_{k_{n-1},k_{n-2}} M_{k_{n-2},k_{n-3}} \ldots M_{k_1,1} \\ &= M_{1,k_{n-1}} M_{k_{n-1},k_{n-2}} \frac{M_{k_{n-2},1} M_{1,k_{n-2}}}{M_{1,k_{n-2}} M_{k_{n-2},1}} M_{k_{n-2},k_{n-3}} \\ &\quad \cdots \frac{M_{k_2,1} M_{1,k_2}}{M_{1,k_2} M_{k_2,1}} M_{k_2,k_1} M_{k_1,1} \\ &= \rho^{(3)}_{k_1,\,k_2} \prod_{r=2}^{n-2} \frac{\rho^{(3)}_{k_r,\,k_{r+1}}}{\rho^{(2)}_{k_r}}. \end{aligned} \tag{63}$$

This is the Wick's theorem for MPSs [24]. At this point, we can check the validity of the reconstructed M, since it necessarily must obey the symmetry $\Xi_{d,\kappa}\overline{M}\Xi_{d,\kappa} = M$ for accordingly ordered spectrum of T.

3.4. Extracting R

To obtain a complete cMPS description of the system at hand, it is necessary to reconstruct the variational parameter matrices R and Q. We have that, by definition,

$$M = X^{-1}\big(\overline{R} \otimes R\big)X \tag{64}$$

and $D = \mathrm{diag}(\lambda_j) = X^{-1}TX$ with the change-of-basis matrix X indeterminate. Because of the gauge invariance of Q and R, we can determine them only up to conjugation with an invertible matrix and therefore will not need to determine the concrete form of X at all. In this sense, there are no specific R and Q matrices to be reconstructed. Nevertheless, we continue using the terms R and Q, thinking, without loss of generality, of matrices that are in a specific, yet arbitrary, gauge.

Our strategy to recover the variational parameter matrices is to choose R diagonal, which can be done in almost all cases, and determine Q accordingly. Equivalently, one could likewise require Q to be diagonal and determine R accordingly, but here we use the former approach. We first diagonalize $M \mapsto Y^{-1}MY = M_{\mathrm{diag}}$ with the change-of-basis matrix Y. Since M, as well as its similar matrix $\overline{R} \otimes R$, has the spectrum $\{\overline{r}_i r_j\}$ with $i, j = 1,\ldots,d$, where $r_1,\ldots,r_d$ are the eigenvalues of R, the entries of M_{diag} can be reordered with a permutation matrix O such that the resulting matrix has the form of a Kronecker product of two diagonal matrices R_{rec}

$$O^{-1}M_{\mathrm{diag}}O = \overline{R}_{\mathrm{rec}} \otimes R_{\mathrm{rec}}. \tag{65}$$

Since R_{rec} by construction is similar to R, we can write it as $R_{\mathrm{rec}} = W^{-1}RW$, where W is the change-of-basis matrix that diagonalizes R. Diagonalizing and reordering M thus yields R in a certain gauge, namely $W^{-1}RW$, and we can identify R_{rec} with a reconstruction of the matrix R.

Note that XYO has a Kronecker product structure as well, which will be important for reconstructing Q. Rewriting equation (65), we have

$$(XYO)^{-1}\big(\overline{R} \otimes R\big)XYO = O^{-1}Y^{-1}MYO \tag{66}$$

which is equal to $\overline{R}_{\mathrm{rec}} \otimes R_{\mathrm{rec}}$, and, by definition of R_{rec} and using a Kronecker product identity, hence equal to

$$\left(\overline{W} \otimes W\right)^{-1}\left(\overline{R} \otimes R\right)\left(\overline{W} \otimes W\right). \tag{67}$$

There is a little subtlety in that, in general, numerical diagonalization algorithms will not provide Y such that XYO is a Kronecker product, but usually such that each eigenvector, a column of Y, is normalized, yielding a matrix Y_N. This matrix can also be written as $Y_N = YD_Y$ with a diagonal matrix D_Y, where $XYD_Y\mathrm{O}$ in general will not correspond to a Kronecker product. This does not affect R_{rec}, since diagonal matrices are invariant under conjugation with other diagonal matrices.

To determine O and extract R_{rec} from $\overline{R}_{\mathrm{rec}} \otimes R_{\mathrm{rec}}$, it is important to take into account that multiplying R with an arbitrary complex phase factor $\mathrm{e}^{\mathrm{i}\varphi}$ does not change $\overline{R} \otimes R$. In the same way, $\overline{Q} \otimes \mathbb{1}_d + \mathbb{1}_d \otimes Q$ is left invariant when adding $\mathrm{i}\chi \cdot \mathbb{1}_d$ with $\chi \in \mathbb{R}$ to Q. Hence, the transfer matrix remains unchanged as well. Clearly, out of density-like correlation functions, R and Q can only be reconstructed up to these factors since Q and R only appear in these Kronecker product terms.

By fixing $\mathrm{e}^{\mathrm{i}\varphi}$, one diagonal entry r_j of R_{rec} can be assumed to be real and M_{diag} can be rearranged to a Kronecker product by successively checking if for an entry $M_{\mathrm{diag},l,l}$ the fraction $\left|M_{\mathrm{diag},l,l}/r_j\right|^2$ yields another (real) entry of M_{diag} (or, in practice with noise, is sufficiently close to it), which must be the case for a Kronecker product matrix with spectrum $\{\overline{r}_i r_j\}$. After repeating this procedure for all entries of M_{diag}, all eigenvalues $\{r_j\}$ are determined, in a fixed order that determines the order of R_{rec} and O as well. Now, it remains to determine Q, which will be done in the next section.

3.5. Extracting Q

The second parameter matrix to be reconstructed, Q, will in general not be diagonal in the same gauge where R is diagonal. The goal is to find Q in the appropriate gauge. First, we take the matrix D, which contains the eigenvalues of T, subtract the reconstructed matrix M, and see that in principle all the information about Q is stored here:

$$\begin{aligned} D - M &= X^{-1}TX - X^{-1} \cdot \overline{R} \otimes R \cdot X \\ &= X^{-1}\left(\overline{Q} \otimes \mathbb{1}_d + \mathbb{1}_d \otimes Q\right)X. \end{aligned} \tag{68}$$

By conjugating this with the matrix YO, which is the same change-of-basis matrix that directly led from M to $\overline{R}_{\mathrm{rec}} \otimes R_{\mathrm{rec}}$, we obtain

$$\begin{aligned} &(XYO)^{-1}\left(\overline{Q} \otimes \mathbb{1}_d + \mathbb{1}_d \otimes Q\right)XYO \\ &= \left(\overline{W} \otimes W\right)^{-1}\left(\overline{Q} \otimes \mathbb{1}_d + \mathbb{1}_d \otimes Q\right)\overline{W} \otimes W \\ &= \overline{W^{-1}QW} \otimes \mathbb{1}_d + \mathbb{1}_d \otimes \left(W^{-1}QW\right). \end{aligned} \tag{69}$$

We obtain in this way $Q_{\mathrm{rec}} := W^{-1}QW$ in the gauge corresponding to the gauge of $R_{\mathrm{rec}} = W^{-1}RW$ and thus it represents a valid set of parameters that define the state. To extract Q_{rec} out of equation (69), we can, as in the case of R_{rec}, assume one diagonal entry $q_{j,j}$ of Q_{rec} to be real, which corresponds to subtracting $\mathrm{i}\,\mathfrak{Im}\,(q_{j,j}) \cdot \mathbb{1}_d$ from Q. In this way, we can read each $q_{j,j}$ from the corresponding diagonal entry $\overline{q}_{j,j} + q_{j,j} = 2q_{j,j}$ in equation (69) and subsequently

the remaining diagonal entries. Because of the structure of equation (69) as a Kronecker sum, the off-diagonal entries can be read off without further preparation.

The fact that Y is only determined up to multiplication with a diagonal matrix D_Y, as mentioned in the previous section, does not pose an obstacle for the reconstruction of Q_{rec}: its gauge needs to be fixed only up to conjugation with a diagonal matrix if R_{rec} is in a diagonal gauge. Furthermore, it does not matter that also the matrix M is only determined up to conjugation with a diagonal matrix D_M, which we used to require that $M_{1,j} = 1$ for $j \geqslant 2$. Using $D_M^{-1} M D_M$ instead of M in equation (68) and $X^{-1}TX$ being diagonal, we have

$$X^{-1}TX - \big(XD_M\big)^{-1}\big(\overline{R} \otimes R\big)XD_M = \big(XD_M\big)^{-1}\big(T - \overline{R} \otimes R\big)XD_M, \tag{70}$$

which is equal to $\tilde{X}^{-1}\big(\overline{Q} \otimes \mathbb{1}_d + \mathbb{1}_d \otimes Q\big)\tilde{X}$ with $\tilde{X} = XD_M$. The particular structure of X or $\tilde{X}$ is not needed in the algorithm.

On the other hand, if we normalize the n-point function and hence M by multiplying it by a constant, we have to be careful since $D - cM$, for some $c \in \mathbb{R}$, will in general not result in a matrix similar to $\overline{Q} \otimes \mathbb{1}_d + \mathbb{1}_d \otimes Q$. Accordingly, we have to *renormalize* $M \mapsto \hat{M}_{1,1} \cdot M$. The number $\hat{M}_{1,1}$ can be read off the residue $\hat{\rho}^{(n)}_{1,\ldots,1} = (\hat{M}_{1,1})^n$ of the n-point function before normalizing it.

Note that computing eigenvectors, which the matrix X consists of, can be a very unstable (in extreme cases even discontinuous) procedure, especially for higher bond dimensions, when eigenvalues can cluster [11, cor. 7.2.6]. Hence the procedure of determining Q is highly susceptible to noise. To improve noise stability, we can average Y by using the symmetry property $\Xi_{d,\kappa}\overline{Y}\Lambda_d = Y$, which follows from the symmetries of M and $\overline{R}_{\text{rec}} \otimes R_{\text{rec}}$, and use $\big(Y + \Xi_{d,\kappa}\overline{Y}\Lambda_d\big)/2$ instead.

This concludes the reconstruction of the variational parameter matrices Q and R, which is the last step in our reconstruction procedure, section 3.1. Additionally, it is now possible to construct the Hamiltonian of the auxiliary system K as in equation (7) et sqq. and relate the cMPS to a Lindblad master equation. The fact that we can reconstruct Q only up to an additive term $\mathrm{i}\chi \cdot \mathbb{1}$ results in K being indeterminate up to an additive term $\chi \cdot \mathbb{1}$. This is reasonable since only the differences in the spectrum of the Hamiltonian are physically relevant and these are not affected by a global shift by χ.

4. Applicability and limitations

The proposed tomography method relies on assumptions. It is hence important to know its limitations and how to check the applicability of the method to given data. The basic assumption is that the correlations in the data are—at least approximately—of the type found in cMPS spatially, or equivalently of the type found in finite dimensional quantum systems whose dynamics are given by a Lindblad equation temporally. It is hence natural to assume that our method is applicable to settings similar to the ground states of gapped local Hamiltonians and for fields which originate from an interaction with finite level systems—think, e.g., of a light beam emitted by an atom trap. In this section, we aim to give a description of ways to gain confidence and check the consistency of the estimates obtained by our reconstruction methods for quantum fields.

Since it is our goal to produce usable estimation tools for experimental applications, it is very important to have a clear understanding of how to determine whether or not a particular

reconstruction was successful or even if the cMPS ansatz is applicable to a particular situation. In this context, we can recognize two different scenarios that can occur: (1) the idealized case, where the data actually comes from a cMPS, and (2) a realistic case, in which the data comes from a physical system (not a cMPS, but possibly well approximated by one) and is in general noisy. In the following, we will discuss both in more detail.

In the ideal case, data will be produced by a generic cMPS of unknown bond dimension d. From the 2-point correlation function, following the reconstruction methods discussed in section 3.2.2, we can extract an estimation of d by computing the rank of the (sufficiently sized) ansatz Hankel matrix in equation (36). Even if noise is present in the signal, an estimation of the bond dimension can be obtained, because noise-induced singular values are small. Since some of the elements of matrix M can be zero, some of the residues ρ corresponding to poles λ can also be zero, thereby hiding those poles. Correlators with different n, on the other hand, can reveal these poles at some point, but not necessarily so. Having found all the poles there are, also implying access to the whole matrix M, is indicated by an agreement of the poles of all available n-point functions. One should keep in mind, though, that one will never be able to verify this, even in the idealized case, with a finite amount of data, as it possible to construct a state which agrees with a given cMPS on e.g., a finite number of n-point functions but differs elsewhere. However, a non-increase of the set of poles over a wide range of n-point functions is sufficient to build confidence in the correctness of the reconstruction. It is a satisfactory feature of our method that we can quantify the confidence of the reconstruction in this way.

In contrast, *a priori* information about the number of expected poles and a guarantee that the number and numerical values of residues and poles will be consistent for all n-point functions is not available in most real-world tomographic settings. In fact, when data comes from an experiment, we expect a description in terms of cMPS to be possible only in an approximate sense. A similar situation is known for discrete MPS in a lattice setting, where an exact description of a state can be found only if its Schmidt rank is finite. However, many states whose Schmidt numbers form a fast decaying sequence allow for an efficient description with discrete MPS. Even if the physical system is well approximated by a cMPS in this sense, in general we expect to have an infinite number of poles to recover. However, only a small number of them will be associated to residues that are big enough to contribute to the correlation functions. The number of relevant residues and poles can be identified by looking for singular values of Hankel matrix equation (36) greater than an appropriate threshold. The tomographer, hence, has to formulate a hypothesis about the relevance of the observed poles and try to gain confidence in his/her assumption. The desired situation to observe in practice is that the recovered poles do not change too much (i.e., they are within some threshold, e.g., previously determined by the noise level) independently of the correlation function used to extract them.

In summary, if the set of poles has to be extended time and again over a wide range of correlation functions, the assumption that the state can be described by a cMPS is clearly wrong. In particular, such a situation would tell us that the cMPS ansatz is not a good model for the particular system and data set. Along the lines of the discussion above, in practice, what we propose to check and gain confidence of the applicability of our methods is the following. Use lower order correlation functions to extract a cMPS description of the system, use the reconstructed cMPS to predict higher order functions and compare them to available measured ones. This way, we can check the consistency of the reconstruction procedure and the validity of the cMPS ansatz for the field state under investigation.

5. Applications

In this section, we show how the formalism developed so far can be applied to real world scenarios. We demonstrate the applicability in two basic settings. First, we generate correlation functions similar to data obtainable in current experimental settings. For this, we use simulated data to study the performance of the reconstruction method in situations in which noise is present. Second, we analyse the applicability of our techniques to the Lieb–Liniger model, which is a well-known and well-investigated model in one-dimensional non-relativistic field theory.

5.1. Simulations and error analysis

Before typical noise models can be taken into consideration, we ask what kind of problems we are most likely to encounter. As we have seen, given an arbitrary cMPS n-point function with non-degenerate spectrum, its poles and residues can be obtained by matrix pencil or Prony's methods, provided there is sufficient accuracy. We keep in mind that formally it is required that T has a non-degenerate spectrum, which is, however, the case for almost all randomized T. Also, it is possible that M contains elements of value zero, which is, likewise, not to be expected. On the other hand, there are other more practical obstacles related to concrete implementation features of the numerical algorithms discussed above.

5.1.1. Typical problems to be expected. The identification of the poles when determining the matrices M and D is the most critical part of our procedure. More concretely, we face the problem of resolving maxima of the Laplace transform of the correlations in the complex plane. We do not do this directly, but the challenges remain the same.

The problem is to discern poles that lie close to each other and to identify poles that have comparatively small residues. Moreover, we might face large damping factors, which results in a broadening of the peaks in the Fourier spectrum. The required accuracy for the correct identification of poles and residues hence critically depends on the position of the poles $\{\lambda_j\}$ in the complex plane and the ratio between damping factor $\mathfrak{Re}(\lambda_j)$ and frequency $\mathfrak{Im}(\lambda_j)$. Not surprisingly, all these issues are aggravated for higher bond dimensions; the n-point functions consist of a larger number of oscillatory components, typically in the vicinity of other poles. Moreover, the reconstruction of the residues will also be affected if the poles are close to each other. This happens because the corresponding linear Vandermonde system of equations becomes more ill-conditioned.

When reconstructing Q from the matrix M, we face another type of typical problem. Determining R does not lead to significant additional numerical problems since it mainly involves an ordinary diagonalization procedure, whereas for reconstructing Q, we need the eigenvectors of M, which are very susceptible to perturbations of the matrix.

In the following, we want to test the robustness of our method by analysing typical noise cases independently. First, as a preparatory step, we generate typical cMPS. Second, we examine how the reconstruction of the poles is affected by adding noise to the input correlation functions. Third, we survey the reconstructability of R and Q when the input for this reconstruction step, the matrix M, is perturbed. Fourth, we study the influence of the presence of additional fields.

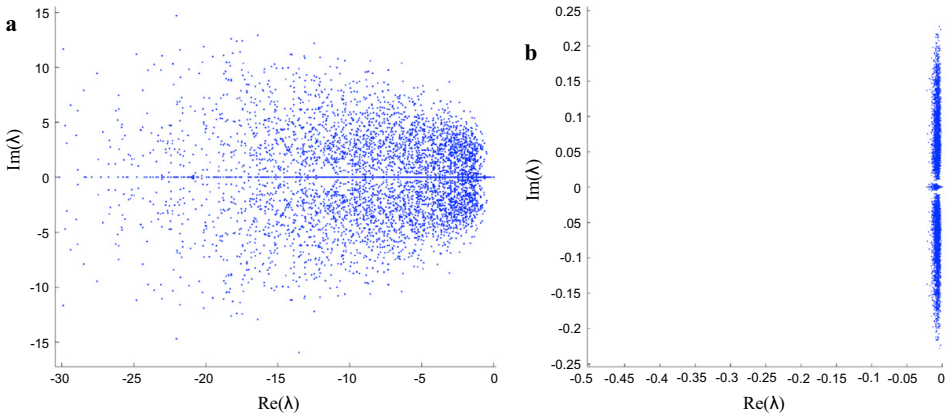

Figure 2. Distribution of the poles of the transfer matrices in the complex plane for 400 cMPS samples with bond dimension $d = 4$. The real and imaginary part of the entries of K and R are i.i.d. with $\mathcal{N}(0, 1)$ (a) and $\mathcal{N}(0, 0.01)$ (b). In (b), most damping factors corresponding to the real parts of the poles are much smaller than the respective imaginary parts, which correspond to the frequencies components of the correlation functions. This will lead to significantly better reconstructability properties of the cMPS.

5.1.2. Generating typical cMPS. In this section, we give a recipe to generate correlation functions with structural features on a desired length scale, based on a randomization-ansatz for the Q and R matrices. This is in principle a non-trivial task, as the length scales and damping of the fluctuations are directly derived from the spectrum of T, which depends nonlinearly on the entries of Q and R.

We make the ansatz of generating Q and R as complex Gaussian random matrices with mean μ and standard deviation σ—i.e., real and imaginary part of the entries are independently and identically normally distributed according to $\mathcal{N}(\mu, \sigma)$—and renormalize Q such that all eigenvalues of T have real part $\leqslant 0$. This results in a roughly uniform distribution of the eigenvalues of T within a disc left of the imaginary axis, which is not entirely unexpected when considering Girko's circular law [7] and the Kronecker product structure of T. The damping factors of the poles are of the same magnitude as their frequencies or larger, which is not the case if oscillations are actually to be observed and moreover aggravates the identification of such poles and increases the accuracy requirements.

In a more refined ansatz, we hence consider sampling K and R instead, from the same distribution, which leads to a drastically higher concentration of poles close to the imaginary axis, when scaling both matrices with a small number η, see figure 2, where we show a comparison of distributions of the poles in the complex plane between the naïve and the refined method of randomly sampled cMPS. This scaling of the matrices does not constitute a gauge of the cMPS but rather a transformation to another cMPS, see [47]. Matrix Q is mapped to $\frac{1}{2}\eta^2 R^\dagger R - \mathrm{i}\eta K$, see equation (7), such that for small η the eigenvalues of Q will typically feature much larger imaginary part than real part, since the spectrum of K is real and the $R^\dagger R$ term adds to Q in second order in η. This carries over to the construction of T, where $\overline{R} \otimes R$ also appears

in second order in η as opposed to $\overline{Q} \otimes \mathbb{1} + \mathbb{1} \otimes Q$, which are first order. Overall, for small η most damping factors become smaller than the frequencies by several orders of magnitude, a property expected to hold if oscillations are observed. Moreover, a distinct peak structure in the Fourier transform emerges, and the poles and residues of T are sufficiently separated and can be determined even with moderate amounts of noise present.

5.1.3. Effects of noisy correlation functions. Typical experimentally measured signals have inaccurate read-out of the signal. We model such noisy situations as Gaussian noise, and study the effect on the reconstruction procedure by adding noise to correlation functions originating from a cMPS.

In particular, we apply the MPM to the noisy amputated 2-point function

$$\hat{C}^{(2)}(\tau_k) + w(\tau_k) = \left\langle \hat{\Psi}^\dagger(\tau_k)\hat{\Psi}^\dagger(0)\hat{\Psi}(0)\hat{\Psi}(\tau_k) \right\rangle - \left\langle \hat{\Psi}^\dagger(0)\hat{\Psi}(0) \right\rangle^2 + w(\tau_k), \tag{71}$$

evaluated at 200 points τ_k, for cMPS with elements of R, K sampled from $\mathcal{N}(0, 0.01)$. The white noise function w is sampled from $\mathcal{N}(0, \text{mean}(|\hat{C}^{(2)}|)/\text{SNR})$, where SNR is the signal-to-noise ratio.

In figure 3, p is the percentage of pole sets with $\text{mean}_{j=2,\ldots,d^2}|(\lambda_j - \tilde{\lambda}_j)/\lambda_j| < 0.1$ as a function of the SNR, where $\{\lambda_j\}$ are the original poles, and $\{\tilde{\lambda}_j\}$ the pole estimates. Each point is computed for 5000 runs of our numerical experiment to gather enough statistics. What we observe is that for bond dimension $d = 2$, our reconstruction procedure is robust to reasonable amounts of noise. However, for bond dimension $d = 3$, we see that the robustness is much smaller, which hints to the practical limitations of our reconstruction procedure. The results can, for example, be improved by increasing the sampling rates, however this can be difficult to achieve in experiments.

Note that in both cases shown in figure 3 our procedure behaves as expected from a proper estimator as a function of the SNR: the lesser the noise, the better the reconstruction. In fact, for zero noise, we can in general expect 100% reconstructability, independent of the bond dimension. As already mentioned, for higher order correlation functions, $n > 2$, the reconstructability of the poles does not necessarily deteriorate—independent of the bond dimension d. In fact, since one can average over all projections that fix all but one τ, a significant part of the noise is effectively averaged out.

5.1.4. Reconstructability of Q and R when perturbing M. In this section, we look at the next step in the reconstruction process: recovering the cMPS parametrization matrices Q and R from an imperfectly recovered matrix M. We do so by simulating M and perturbing it directly, rather than using a reconstructed M matrix from noisy correlation functions. We do it this way to have control over the size of the perturbation and thus to separate these two different stages of the reconstructed problem and investigate their effect separately.

For this purpose, we prepare matrices R and Q with entries sampled from $\mathcal{N}(0, 1)$, then calculate T and M, and perturb M with an error matrix Δ. The perturbation has to be carefully designed in order to retain the symmetry $M = \Xi_{d,\kappa}\overline{M}\Xi_{d,\kappa}$ of the unperturbed matrix M. This is related to the fact that for any valid reconstruction of a density-like correlation function the residues together with the entries of the matrix M necessarily are either real or appear in pairs of complex conjugates, see section 2.3. Perturbing with the matrix

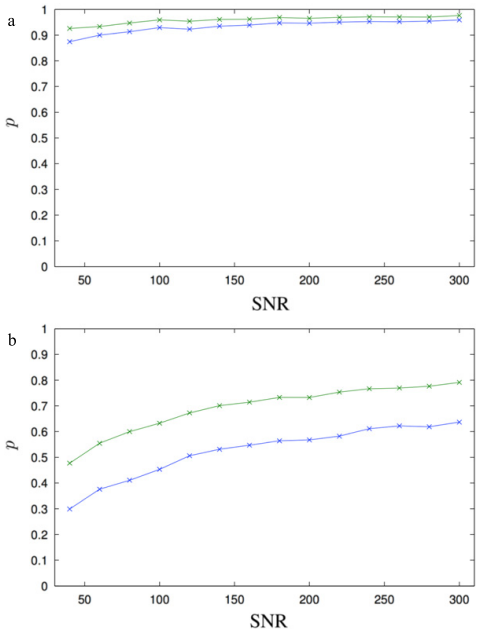

Figure 3. Application of the matrix pencil method to the signal vector with components $\hat{C}^{(2)}(\tau_k) + w(\tau_k))$ for $d = 2$ (above) and $d = 3$ (below). p is the percentage of pole sets with $\max_{j=2,\ldots,d^2}|(\lambda_j - \tilde{\lambda}_j)/\lambda_j| < 0.1$ (blue) and $(d^2-1)^{-1}\sum_{j=2}^{d^2}|(\lambda_j - \tilde{\lambda}_j)/\lambda_j| < 0.1$ (green) as a function of the signal-to-noise ratio, each point summarizing 5000 runs, where $\{\lambda_j\}$ are the original poles, and $\{\tilde{\lambda}_j\}$ the pole estimates.

$$\Delta := \frac{1}{2}\left(\Delta_0 + \Xi_{d,\kappa}\overline{\Delta}_0\Xi_{d,\kappa}\right) \tag{72}$$

with real and imaginary parts of the entries of Δ_0 sampled from $\mathcal{N}(0, 2^{-1/2}\,\mathrm{mean}(|M|))$ ensures the required symmetry since $\Delta = \Xi_{d,\kappa}\overline{\Delta}\,\Xi_{d,\kappa}$. Furthermore, since the first row of M is set to one due to normalization and this should not be changed for perturbed input, the first row of Δ is set to zero.

From the reconstructed matrices $\tilde{Q}$ and $\tilde{R}$ from $\tilde{M} = M + \epsilon\Delta$ with scaling parameter $\epsilon \in \mathbb{R}^+$ we build the transfer matrix $\tilde{T}$ and compare its spectrum with the spectrum of the original T. The ratio of samples with mean deviation $\sigma(\tilde{T})$ to $\sigma(T)$ not larger than 10% as a function of ϵ is depicted in figure 4 for bond dimensions $d = 2$ (blue) and $d = 3$ (green). As the error ϵ grows, the ratio of successfully reconstructed Q and R matrices drops for both bond dimensions. However, the $d = 2$ case is clearly more robust to perturbations. Additionally, we want to point out that any potential deviation of the spectra of T and $\tilde{T}$ is almost certainly due to the reconstruction of Q.

5.1.5. Effects of additional interactions. As discussed earlier in section 2.1.2, typical correlations under consideration can be seen as originating from processes where a field state is generated by an interaction with a finite dimensional system, and can be described by a

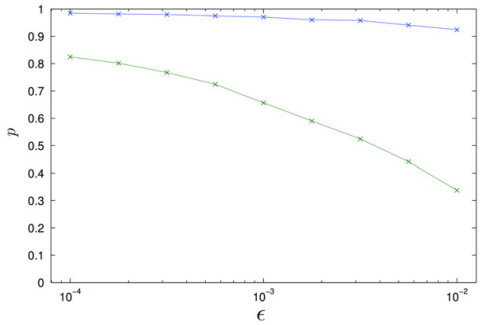

Figure 4. Reconstructability of Q depending on the perturbation of M: ratio p out of 5000 samples per point with $(d^2-1)^{-1}\sum_{j=2}^{d^2}|(\lambda_j-\tilde{\lambda}_j)/\lambda_j|<0.1$ as a function of ϵ with $\{\tilde{\lambda}_j\}=\sigma(\tilde{T})$ for $d=2$ (blue) and $d=3$ (green). $\tilde{Q}$ and $\tilde{R}$ depend on $\tilde{M}=M+\epsilon\Delta$. As $\epsilon\to 0$ we have that $p\to 1$.

Lindblad equation. In the ideal case, where the finite dimensional system interacts only with the field we measure, we obtain correlations which are perfectly described by a cMPS, or equivalently by a Lindblad equation with one Lindblad operator. In the case where the finite dimensional system interacts with other systems or fields, which we might not even know of, the Lindblad equation is altered and supplemented by more Lindblad operators, which correspond to the other systems or fields. In this case, the transfer matrix takes the form [35]

$$T=\mathrm{i}K\otimes\mathbb{1}-\mathbb{1}\otimes\mathrm{i}K+\sum_j\mathcal{R}_j, \tag{73}$$

where

$$\mathcal{R}_j=\frac{1}{2}\left(2\overline{R}_j\otimes R_j-\overline{R_j^\dagger R_j}\otimes\mathbb{1}-\mathbb{1}\otimes R_j^\dagger R_j\right) \tag{74}$$

and the additional fields are represented by the terms with $j\geqslant 2$. Each of the two latter summands in $\mathcal{R}_j$ are connected to Q via equation (7). The matrix M remains $\overline{R}_1\otimes R_1$, because it comes from measuring the field corresponding to it, but now in the diagonal basis of a different T than the one for a single field.

In order to analyse the sensitivity of reconstructing the variational parameter matrices, we consider one additional perturbation field. More additional fields within the same order of magnitude yield very similar outcomes. This results in $T=\mathrm{i}K\otimes\mathbb{1}-\mathbb{1}\otimes\mathrm{i}K+\mathcal{R}_1+\epsilon\mathcal{R}_2$. In this section, we study how well the spectrum of K can be matched depending on the scaling parameter $\epsilon\in\mathbb{R}^+$. Analogous to the last section, we prepare cMPS by randomly generating K, R_1, and R_2 with elements whose real and imaginary parts are sampled from $\mathcal{N}(0,1)$. We then generate M matrices and from this reconstruct $R_{1,\mathrm{rec}}$ and an effective Q_{rec}, assuming only a single field. From $R_{1,\mathrm{rec}}$ and Q_{rec} we compute K_{rec} and compare the differences of its eigenvalues, $\Delta\tilde{\kappa}_j=\tilde{\kappa}_{j+1}-\tilde{\kappa}_j$, with the differences of the eigenvalues κ_j of the actual K. Only the differences are reconstructable, see section 3.5. The reconstruction of K is said to be successful if

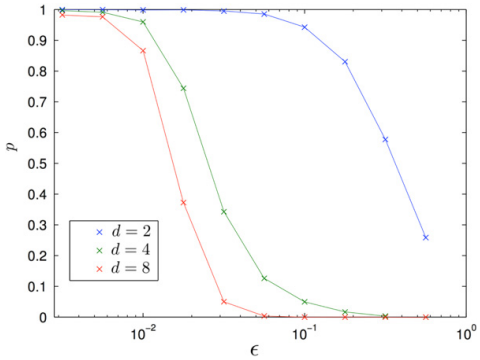

Figure 5. Reconstruction rate p depending on the size of an additional field and the bond dimension d from 5000 cMPS samples per point.

$$\max_{j=1,\ldots,d-1} \left| \frac{\Delta\tilde{\kappa}_j - \Delta\kappa_j}{\Delta\kappa_j} \right| < 10\%. \tag{75}$$

The reconstruction rate, depending on ϵ and the bond dimension, is shown in figure 5. For $\epsilon \to 0$ (single field case) all cMPS can be reconstructed. As the size of the additional field approaches the size of the main field, the reconstruction rate drops to zero. The smaller the bond dimension, the more perturbation by additional fields can be tolerated. We conclude that for sufficiently small additional fields, a successful reconstruction is in principle still feasible. Moreover, for $d = 2$, the most robust case, this is true even if the additional fields are merely one order of magnitude smaller than the main field.

5.2. The Lieb–Liniger model

In this section, we analyse the applicability of the results discussed above to the Lieb–Liniger model [30]. The model describes the dynamics of a one-dimensional system of bosons interacting via a delta-potential. In second quantization, the Hamiltonian describing such a model is given by

$$H = \int \mathrm{d}x \left(\frac{\mathrm{d}\hat{\Psi}^\dagger(x)}{\mathrm{d}x} \frac{\mathrm{d}\hat{\Psi}(x)}{\mathrm{d}x} + c\hat{\Psi}^\dagger(x)\hat{\Psi}^\dagger(x)\hat{\Psi}(x)\hat{\Psi}(x) \right), \tag{76}$$

where $x \in [0, L]$ is the position coordinate and c is the interaction strength.

For our application, we generate (Q,R) parametrizations of cMPS approximations for several bond dimensions of the Lieb–Liniger ground state for particular values of interaction strength c by using the algorithm and implementation of Hase [17]. This algorithm is an adaptation of the time-dependent variational principle for quantum lattices [15] to the continuous case (compare also [6]). It relates to an imaginary time evolution that exponentially damps all excited components of an initial state vector $|\Psi^{(d)}\rangle$ (a cMPS with bond dimension d) with increasing imaginary time and produces the ground state eigenvector of a Hamiltonian H, by applying $\mathrm{e}^{-\mathrm{i}Ht}$ with $t \in \mathrm{i}\mathbb{R}$ to $|\Psi^{(d)}\rangle$. The convergence of the energy of $\mathrm{e}^{-\mathrm{i}Ht}|\Psi^{(d)}\rangle$ indicates the approach to the cMPS ansatz ground state vector, which we denote by $|\Theta^{(d)}_{Q,R}\rangle$, together with

its characterizing matrices Q and R. Several interesting structural properties of the state in the cMPS representation are revealed, signifying a symmetry in the model: degeneracies and a block structure of the matrix M. These features emerge in the integrable Lieb–Liniger case, and do not appear in Gaussian-sampled cMPS as described above. These features, which will be discussed more in detail in the following, appear regardless of the bond dimension and interaction strength used. Moreover, they do not depend on the algorithm used to obtain the ground state.

5.2.1. Degeneracies in the eigenvalue structure of M. The topic of this section is to characterize the structure of the spectrum of M by understanding the degeneracy structure R in the exactly integrable case. In the case at hand, since all two-fold degenerate eigenvalues are equally spread into one of both blocks each, one is able to predict the spectrum of R from M even without reconstructing the second block. In our simulations, it is seen that the eigenvalues of Q and R appear in $\lfloor d/2 \rfloor$ pairs $\{q_j^{[1]}, q_j^{[2]}\}$ and $\{r_j^{[1]}, r_j^{[2]}\}$ with

$$q_j^{[1]} = \overline{q_j^{[2]}} + \mathrm{i}\chi, \quad r_j^{[1]} = \overline{r_j^{[2]}}\mathrm{e}^{\mathrm{i}\varphi}, \tag{77}$$

respectively, for each pair j, with $\chi, \phi \in \mathbb{R}$ independent of j. If d is odd, the two remaining unpaired eigenvalues take the form $q = \hat{q} + \mathrm{i}\chi$ and $r = \hat{r}\mathrm{e}^{\mathrm{i}\varphi}$, respectively, with $\hat{q}, \hat{r} \in \mathbb{R}$. We can simplify the structure by performing the transformations

$$Q \mapsto Q - \mathrm{i}\chi\mathbb{1}_d, \quad R \mapsto R\mathrm{e}^{-\mathrm{i}\varphi}, \tag{78}$$

which leave the transfer matrix T and all density-like n-point functions invariant. This ensures that the pairs now consist of complex conjugates and the spectra of Q and R are closed under complex conjugation, which we want to require for the further argument.

Since the spectrum of M by construction is the same as that of $\overline{R} \otimes R$ (up to a normalization constant and each $\lambda \in \sigma\left(\overline{R} \otimes R\right)$ can be written as $\bar{r}_j \cdot r_k$ with certain $j, k = 1, \ldots, d$, the appearance of complex conjugate pairs in the spectrum of R implies twofold degeneracies for the according eigenvalues in the spectrum of M as products of R eigenvalues, especially

$$\overline{r_j^{[1]}} r_k^{[1]} = r_j^{[2]}\,\overline{r_k^{[2]}} = \overline{r_k^{[2]}} r_j^{[2]}. \tag{79}$$

Not all eigenvalues are degenerate: $\overline{r_j^{[1]}} r_j^{[2]}$ and $\overline{r_j^{[2]}} r_j^{[1]}$ are complex conjugates, but since $j = k$, there are no other combinations that yield the same values. Assuming that R does not contain any other degeneracies, M will comprise d non-degenerate eigenvalues and $d^2 - d$ eigenvalues that are twofold degenerate each.

5.2.2. Block structure. Another structural observation we can make for the matrix M of the ground state of the Lieb–Liniger Hamiltonian is the fact that it can be transformed to a block diagonal matrix. We do this by simply grouping vanishing and non-vanishing elements in M and interchanging its rows and columns correspondingly, which amounts to a basis permutation. This way, we define the matrix $M^{\square} := M_1 \oplus M_2$, where M_1 and M_2 are block matrices and relate to the non-vanishing and vanishing residues of the cMPS. The block structure of $M^{\square}$ and the fact that e^T is diagonal imply a block structure of their products, which carries over to the correlation functions, lets M_2 decouple completely, and hence disappear from the reconstruction.

We can see why all the residues corresponding to M_2 vanish for every n-point function in the following way. Let us assume we reordered M and formed $M^{\square}$ by performing the basis permutations described above, and we consider a pole λ_l of the cMPS. For an arbitrary n-point function, each residue which contains the index l at least once can be written as

$$\rho_{k_1,\ldots,k_{j-1},l,k_{j+1},\ldots,k_{n-1}} = M^{\square}_{1,k_1}\ldots M^{\square}_{k_{j-1},l} M^{\square}_{l,k_{j+1}}\ldots M^{\square}_{k_{n-1},1} \tag{80}$$

with $j = 2,\ldots,n-2$.

We take $l \in \{\zeta+1,\ldots,d^2\}$, where ζ is the dimension of M_1, i.e., λ_l corresponds to a pole associated with M_2. In this situation, we note that two things can happen. Either $k_{j+1} \leqslant \zeta$, and $M^{\square}_{l,k_{j+1}} = 0$ since the entry is located in the lower left block of $M^{\square}$, which contains just zeros, and thus the entry vanishes. Or $k_{j+1} > \zeta$, and there exists an entry $M^{\square}_{k_m,k_{m+1}}$ with $m > j$, $k_m > \zeta$, and $k_{m+1} \leqslant \zeta$ such that $M^{\square}_{k_m,k_{m+1}} = 0$. This has to eventually happen since the last entry in the residue expression is of the form $M^{\square}_{k_{n-1},1}$ and $1 \leqslant \zeta$. Clearly, the residue vanishes again, and so does for the boundary indices $k_1 = l$ or $k_{n-1} = l$.

5.2.3. Reconstruction. Because of the block structure of M, we conclude that there is no direct way of obtaining all poles of cMPS approximations of the Lieb–Liniger ground state from an n-point density-like correlation function. In this case, the p-number [24], which is defined as the minimum order for a p-point function of a cMPS to reveal all poles, is infinite. There is a useful connection between the degeneracies in the spectrum of M and its block structure for the Lieb–Liniger model. It turns out that all the non-degenerate eigenvalues are related to $M^{\square}$ entries in the first block, while the degenerate pairs are distributed such that always one eigenvalue is associated with the first block and the other with the second. This way, since only the first block contributes to any density-like correlation function, all degeneracies are *effectively* lifted, and hence full reconstruction is possible. Since all eigenvalues of M that appear in the vanishing second block also appear in the visible first block one can in principle determine the spectrum of R even without full knowledge of M. The same holds for the spectrum of Q since also $D - M$ has the same spectral properties. For reconstructing *both* R and Q in the corresponding gauge, however, our procedure requires full knowledge of M. But again, note that for full reconstruction of the density-like correlation functions, this full knowledge is not required here.

This structure disappears if integrability is broken, and hence in a neighbourhood around the (cMPS approximation of the) Lieb–Liniger ground state. Imaginary time evolution gives us a notion of distance to the limit of the approximation process, as we can, e.g., observe convergence of matrix entries along imaginary time paths. The block structure and degeneracy become more clearly defined the closer one gets to the limit point. Ultimately, at the limit point of the imaginary time evolution, the degeneracies and block structure of M will prevent our methods to recover a full cMPS description in terms of matrices Q and R of the system. On the other hand, for each state along such a path, we can in principle apply our reconstruction method. The closer we get, the better all characteristic parameters can be reconstructed although the more ill-conditioned the problem becomes. A reconstruction of the n-point functions of arbitrary order is still possible, as it is based on the observable blocks of the matrices D and M alone and determining these quantities is in principle possible. Since the second block does not contribute to *any* n-point function, the applicability of 'Wick's theorem' for (continuous) MPSs

is maintained even in this case and we still can successfully predict higher order from lower order correlation functions.

6. Summary and outlook

In this work, we have introduced the concept of quantum field tomography. In spite of the inherent difficulties of attempting to reconstruct a continuous system, i.e., a system with infinite degrees of freedom, we have shown that this task can be done when only a relevant class of naturally occurring states is considered. This is physically well motivated since one expects naturally appearing states not to be of the most general form but restricted to a smaller class of states. This is clearly the case in physical applications in which, for example, MPSs have been shown to be a very successful model to describe correlations and dynamics. Here, we concentrated on developing tomographic tools for one-dimensional continuous many-body systems or quantum fields.

For this purpose, we employed the continuous generalization of the MPS variational class of states: the continuous MPS formalism. Based on this formalism and the predicted structure of the relevant data, i.e., the correlation functions, we developed a procedure to extract a best fit cMPS using state of the art statistical estimation tools. In this way, we are able to deliver a working and readily applicable tool to study this type of system. The procedure we offer can indeed be seen as the natural way to think of efficient quantum field tomography. This does not mean, however, that for tasks of direct estimation of fidelities and properties of states, alternative methods may not be advisable. The machinery here aims at reconstructing the states as such.

Formally, we have used the cMPS framework to describe the structure of correlation functions that can in principle be measured in experiments. Having identified this basic structure, we defined the tools needed to extract the pertinent information from the data. For this purpose, we employed the MPM as a viable way to determine the variational parameters of the cMPS from a correlation function. We showed that one can successfully extract a cMPS description of a system in principle for arbitrary bond dimensions. However, for noisy signals, one is in general limited to lower bond dimension approximations. Generally, this approach is applicable to states with low entanglement, similarly to matrix-product states approximating states that satisfy an area law for suitable Renyi entanglement entropies. In the discrete case, the connection of having ‘low entanglement’ and being approximable with a MPS of low bond dimension has been fully rigorously fleshed out already [8, 41]. In the continuous case, this connection is surely equally plausible, but is awaiting a similar fully rigorous treatment.

Moreover, we have given an in-depth study of the applicability of the reconstruction tools and their robustness for different noise models. Extensive numerical simulations were employed which provide at least empirical confidence of the performance of the reconstruction tools. We found that for the cases studied in this work, our methods are reasonably robust to noise when searching for low bond dimension cMPS estimates.

It is important to note that the methods developed in this work are likewise readily applicable to the translationally invariant discrete MPS case. Since in reality one deals with discrete (sampled) data even if the system is continuous in nature, all the methods developed here carry to the discrete case of MPSs, reflecting a finite lattice spacing, with minimal

modifications. Furthermore, there is evidence that the approach taken here reveals insight into the structure of the underlying model as such and can detect signatures of integrability.

The novel methods proposed in this work open a new avenue to explore continuous systems of many particles in both equilibrium and non-equilibrium. It constitutes a step towards assessing strongly correlated models with a topographic mindset, without having to make a model of the system in the first place: instead, one asks what the state is that is most compatible with the data found. This is a most healthy mindset specifically in the context of emergent quantum technologies, where one aims at assessing the state of a quantum system without making overly strong assumptions in the first place. In quantum information science, quantum state tomography is already a pillar on which the field rests, a technique routinely applied in most experiments. The present work opens up perspectives to think of quantum field tomography of strongly correlated quantum systems, as they feature in dynamical quantum simulators. Specifically in this context, the tools presented here can be used for partial benchmarking of analog quantum simulators. To fully explore the potential of such an approach to study many-body systems out of equilibrium constitutes a truly exciting perspective.

Acknowledgements

This work was supported by the the BMBF, the EU (RAQUEL, COST, SIQS, AQuS), and the ERC (TAQ). We thank T J Osborne and M Friesdorf for discussions and M von Hase for providing the code using the time-dependent variational principle.

References

[1] Baumgratz T, Gross D, Cramer M and Plenio M B 2013 Scalable reconstruction of density matrices *Phys. Rev. Lett.* **111** 020401
[2] Cadzow J A and Wilkes D M 1991 Enhanced rational signal modeling *Signal Process.* **25** 171
[3] Calabrese P and Caux J-S 2007 Dynamics of the attractive 1D Bose gas: analytical treatment from integrability *J. Stat. Mech.* P08032
[4] Chandrasekaran S and Sayed A H 1998 A fast stable solver for nonsymmetric Toeplitz and quasi-Toeplitz systems of linear equations *J. Soc. Ind. Appl. Math.* **19** 107
[5] Cramer M, Plenio M B, Flammia S T, Somma R, Gross D, Bartlett S D, Landon-Cardinal O, Poulin D and Liu Y-K 2010 Efficient quantum state tomography *Nat. Commun.* **1** 149
[6] Draxler D, Haegeman J, Osborne T J, Stojevic V, Vanderstraeten L and Verstraete F 2013 Particles, holes, and solitons: a matrix product state approach *Phys. Rev. Lett.* **111** 020402
[7] Edelman A and Rao N R 2005 Random matrix theory *Acta Numer.* **14** 233
[8] Eisert J, Cramer M and Plenio M B 2010 Area laws for the entanglement entropy *Rev. Mod. Phys.* **82** 277
[9] Flammia S T, Gross D, Liu Y-K and Eisert J 2012 Quantum tomography via compressed sensing: error bounds, sample complexity, and efficient estimators *New J. Phys.* **14** 095022
[10] Gautschi W 1978 On inverses of Vandermonde and confluent Vandermonde matrices: III. *Numer. Math.* **29** 445
[11] Golub G and van Loan C 1996 *Matrix Computations. Johns Hopkins Studies in the Mathematical Sciences* (Baltimore, MD: Johns Hopkins University Press)
[12] Graham A 1981 *Kronecker Products and Matrix Calculus with Applications* (New York: Wiley)
[13] Gring M, Kuhnert M, Langen T, Kitagawa T, Rauer B, Schreitl M, Mazets I, Smith D A, Demler E and Schmiedmayer J 2012 Relaxation and prethermalization in an isolated quantum system *Science* **337** 1318

[14] Gross D, Liu Y-K, Flammia S T, Becker S and Eisert J 2010 Quantum state tomography via compressed sensing *Phys. Rev. Lett.* **105** 150401
[15] Haegeman J, Cirac J I, Osborne T J, Pižorn I, Verschelde H and Verstraete F 2011 Time-dependent variational principle for quantum lattices *Phys. Rev. Lett.* **107** 070601
[16] Haegeman J, Cirac J I, Osborne T J and Verstraete F 2013 Calculus of continuous matrix product states *Phys. Rev.* B **88** 085118
[17] Hase M V 2013 Continuum versions of tensor network states *MSc Thesis* Freie Universität Berlin
[18] Hildebrand F 1987 *Introduction to Numerical Analysis* (New York: Dover)
[19] Hua Y, Gershman A and Cheng Q 2004 *High-Resolution and Robust Signal Processing* (*Signal Processing and Communications Series*) (New York: Dekker)
[20] Hua Y, Hu F and Sarkar T 1989 Matrix pencil with pre-filtering for direction finding *Proc. 23rd Asilomar Conf. on Signals, Systems and Computers* vol 2 pp 768–71
[21] Hua Y and Sarkar T 1989 Generalised pencil-of-function method for extracting poles of an em system from its transient response *IEEE Trans. Antennas Propag.* **37** 229
[22] Hua Y and Sarkar T 1990 Matrix pencil method for estimating parameters of exponentially damped/ undamped sinusoids in noise *IEEE Trans. Acoust. Speech Signal Process.* **38** 814
[23] Hua Y and Sarkar T 1991 On SVD for estimating generalised eigenvalues of singular matrix pencil in noise *IEEE Trans. Signal Process.* **39** 892
[24] Hübener R, Mari A and Eisert J 2013 Wick's theorem for matrix product states *Phys. Rev. Lett.* **110** 040401
[25] Iohvidov I 1982 *Hankel and Toeplitz Matrices and Forms: Algebraic Theory* vol 6 (Cambridge, MA: Birkhauser Boston)
[26] Karski M, Förster L, Choi J M, Alt W, Widera A and Meschede D 2009 Nearest-neighbor detection of atoms in a 1D optical lattice by fluorescence imaging *Phys. Rev. Lett.* **102** 053001
[27] Kitagawa T, Imambekov A, Schmiedmayer J and Demler E 2011 The dynamics and prethermalization of one-dimensional quantum systems probed through the full distributions of quantum noise *New J. Phys.* **13** 073018
[28] Langen T, Geiger R, Kuhnert M, Rauer B and Schmiedmayer J 2013 Local emergence of thermal correlations in an isolated quantum many-body system *Nat. Phys.* **9** 640
[29] Lee J 1997 Riemannian manifolds: an introduction to curvature *Graduate Texts in Mathematics* (Berlin: Springer)
[30] Lieb E H and Liniger W 1963 Exact analysis of an interacting Bose gas: I. The general solution and the ground state *Phys. Rev.* **130** 1605
[31] Lobos T, Rezmer J and Schegner P 2003 Parameter estimation of distorted signals using Prony method *IEEE Bologna Power Tech Conf. Proc.* vol 4
[32] Lu B, Wei D, Evans B L and Bovik A C 1998 Improved matrix pencil methods *Proc. IEEE Asilomar Conf. Signals, Systems, and Computers* p 1433
[33] Mackay A and McCowen A 1987 An improved pencil-of-functions method and comparisons with traditional methods of pole extraction *IEEE Trans. Antennas Propag.* **35** 435
[34] Ohliger M, Nesme V and Eisert J 2013 Efficient and feasible state tomography of quantum many-body systems *New J. Phys.* **15** 015024
[35] Osborne T J, Eisert J and Verstraete F 2010 Holographic quantum states *Phys. Rev. Lett.* **105** 260401
[36] Perez-Garcia D, Verstraete F, Wolf M M and Cirac J I 2007 Matrix product state representations *Quantum Inf. Comput.* **7** 401–30
[37] Potts D and Tasche M 2010 Parameter estimation for exponential sums by approximate Prony method *Signal Process.* **90** 1631
[38] Prony G 1795 Essai experimental et analytique: sur les lois de la dilatabilit des fluide lastiques et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alkool, diffrentes tempratures *J. l'École Polytech.* **22** 24

[39] Sacchini J, Steedly W and Moses R 1993 Two-dimensional Prony modeling and parameter estimation *IEEE Trans. Signal Process.* **41** 3127
[40] Sarkar T and Pereira O 1995 Using the matrix pencil method to estimate the parameters of a sum of complex exponentials *IEEE Antennas Propag. Mag.* **37** 48
[41] Schuch N, Wolf M M, Verstraete F and Cirac J I 2008 Entropy scaling and simulability by matrix product states *Phys. Rev. Lett.* **100** 030504
[42] Shannon C E 1949 Communication in the presence of noise *Proc. IRE* **37** 10
[43] Steffens A, Friesdorf M, Langen T, Rauer B, Schweigler T, Hübener R, Schmiedmayer J, Riofrío C A and Eisert J 2014 Towards experimental quantum field tomography with ultra-cold atoms (arXiv:1406.3632)
[44] Tarantola A 2005 *Inverse Problem Theory and Methods for Model Parameter Estimation* (Philadelphia: SIAM)
[45] Trench W F 1964 An algorithm for the inversion of finite Toeplitz matrices *J. Soc. Ind. Appl. Math.* **12** 515
[46] Tyrtyshnikov E E 1994 How bad are Hankel matrices? *Numer. Math.* **67** 261
[47] Verstraete F and Cirac J I 2010 Continuous matrix product states for quantum fields *Phys. Rev. Lett.* **104** 190405
[48] Wolf M M and Perez-Garcia D 2010 The inverse eigenvalue problem for quantum channels (arXiv:1005.4545)
[49] Yan H 1987 *Signal Processing for Magnetic Resonance Imaging and Spectroscopy* (*Signal Processing and Communications Series*) (London: Taylor and Francis)

3.3 Tomography of ultracold Bose gases

As noted in the previous section, the class of continuous matrix product states can be used to effectively describe various different theoretical models. How they would perform in a real-world setting, however, has not been analyzed so far. In the following, we present a first undertaking to efficiently characterize an experimentally prepared quantum system using CMPS and provide predictions for quantities that could independently be accessed and certified from the experiment.

In the experiment at the basis of the following publication [3][18], an ultracold gas of ^{87}Rb atoms is localized using an atom-chip[140], a magnetic micro-trap on a chip permitting very compact setups. The gas is transversally split into two mutually coherent halves, which corresponds to performing a sudden *quench*—abruptly changing the system parameters towards a new Hamiltonian—and results in an approximately pure state out of equilibrium.

One-dimensional ultracold Bose gases play a prominent role in experimentally analyzing the equilibration behavior in quantum many-body systems[141–143]. How and if closed quantum systems reach an equilibrium state that is close to a thermal state proportional to $\mathrm{e}^{-\beta H}$ (for Hamiltonian H and inverse temperature β) is a nontrivial and particularly interesting question, considering the time evolution of quantum states being governed by unitary transformations versus the monotonous increase of entropy with time according to the second law of thermodynamics (see, e.g., Ref. [144] for an overview).

The effectively one-dimensional ultracold Bose gas at hand can be captured by the *Lieb-Liniger model*, whose low-energy states, in turn, are well-approximated by CMPS with limited bond dimension.[125,145] This suggests the adaptation of the quantum field tomography protocol in publication [2] for determining the state of the system. The input n-point correlation functions, taking the form

$$C(x_1,\ldots,x_n) = \mathrm{Re}\left\langle \mathrm{e}^{\mathrm{i}\left(\hat{\theta}_{x_1}-\hat{\theta}_{x_2}+\hat{\theta}_{x_3}-\ldots+\hat{\theta}_{x_{n-1}}-\hat{\theta}_{x_n}\right)}\right\rangle, \tag{28}$$

were obtained by measuring via matter-wave interferometry the relative local phase difference θ_x at longitudinal position x. Using the polar decomposition $\hat{\psi}^\dagger(x) = \hat{n}(x)^{1/2}\,\mathrm{e}^{\hat{\theta}_x}$, the correlation functions can be reduced to a form as in Eq. (27). Specifically, reading in and processing 2-point and 4-point correlation functions allowed for a partial quantum-field reconstruction of the state of the system.[19] Using the reconstructed parameters, predictions about higher-order statistical behavior could be made: 6-point correlation functions could be produced that were in very good agreement with the ones directly obtained from the experiment, thus building further confidence in the suitability of the tomographic protocol. The robustness of the procedure was analyzed in a *boot-*

[18] Adrian Steffens, Mathis Friesdorf, Tim Langen, Bernhard Rauer, Thomas Schweigler, Robert Hübener, Jörg Schmiedmayer, Carlos A. Riofrío, and Jens Eisert, "Towards experimental quantum-field tomography with ultracold atoms", *Nature Communications* 6:7663, 2015 (DOI:10.1038/ncomms8663). Published under a Creative Commons Attribution 4.0 International License (creativecommons.org/licenses/by/4.0), © 2015 Springer Nature Publishing AG.

[19] Namely, the matrices M for applying *Wick's theorem for matrix product states* [135] for computing higher-order statistics from lower-order statistics could be recovered.

strapping[146,147] manner by varying the input data and recording the statistics of the corresponding output parameters, yielding effective error bars.

The reconstruction was performed for different times after the initial state preparation and we observed that with time, the agreement between the predicted correlation functions and the experimentally determined ones deteriorated. Whether this change is due to increasing noise or to entanglement growth, which happens in *quenched* systems[102,148] and requires to accordingly increase the bond dimension of the modeling CMPS, has to be investigated in future studies. If additionally, with increasing control over the experimental error, the CMPS parameter matrices Q and R were obtained, it would furthermore be possible to efficiently numerically evolve the CMPS in time[132] and to compute correlation functions for future times. It would be fascinating to relate this to the experimental values and ideally retrace the equilibration process of ultracold Bose gases. Due to experimental imperfections, the system is not perfectly translation invariant. Extending the tomographic protocol to also incorporate position-dependent features would be highly desirable to increase the accuracy in describing the state and its predictive power.

Note that directly applying a compressed sensing based quantum tomography protocol, as discussed in section 2.2, to this setup would have been entirely infeasible due to the extraordinary size of the involved density and measurement matrices without requiring tensor network properties. The strength of compressed sensing tomography rather lies in its broad scope concerning quantum systems of intermediate size.

FIRST-AUTHORED PUBLICATIONS

[1] Experimentally exploring compressed sensing quantum tomography
Adrian Steffens, Carlos A. Riofrío, Will McCutcheon, Ingo Roth, Bryn A. Bell, Alex McMillan, Mark S. Tame, John G. Rarity, and Jens Eisert, *Quantum Science and Technology* 2:025005, 2017.

[2] Quantum field tomography
Adrian Steffens, Carlos A. Riofrío, Robert Hübener, and Jens Eisert, *New Journal of Physics* 16:123010, 2014.

[3] Towards experimental quantum-field tomography with ultracold atoms
Adrian Steffens, Mathis Friesdorf, Tim Langen, Bernhard Rauer, Thomas Schweigler, Robert Hübener, Jörg Schmiedmayer, Carlos A. Riofrío, and Jens Eisert, *Nature Communications* 6:7663, 2015.

[4] An efficient quantum algorithm for spectral estimation
Adrian Steffens, Patrick Rebentrost, Iman Marvian, Jens Eisert, and Seth Lloyd *New Journal of Physics* 19:033005, 2017.

nature COMMUNICATIONS

ARTICLE

Received 29 Oct 2014 | Accepted 22 May 2015 | Published 3 Jul 2015

DOI: 10.1038/ncomms8663 OPEN

Towards experimental quantum-field tomography with ultracold atoms

A. Steffens[1], M. Friesdorf[1], T. Langen[2], B. Rauer[2], T. Schweigler[2], R. Hübener[1], J. Schmiedmayer[2], C.A. Riofrío[1] & J. Eisert[1]

The experimental realization of large-scale many-body systems in atomic-optical architectures has seen immense progress in recent years, rendering full tomography tools for state identification inefficient, especially for continuous systems. To work with these emerging physical platforms, new technologies for state identification are required. Here we present first steps towards efficient experimental quantum-field tomography. Our procedure is based on the continuous analogues of matrix-product states, ubiquitous in condensed-matter theory. These states naturally incorporate the locality present in realistic physical settings and are thus prime candidates for describing the physics of locally interacting quantum fields. To experimentally demonstrate the power of our procedure, we quench a one-dimensional Bose gas by a transversal split and use our method for a partial quantum-field reconstruction of the far-from-equilibrium states of this system. We expect our technique to play an important role in future studies of continuous quantum many-body systems.

[1] Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, Berlin 14195, Germany. [2] Vienna Center for Quantum Science and Technology, Atominstitut, TU Wien, Stadionallee 2, Vienna 1020, Austria. Correspondence and requests for materials should be addressed to J.E. (email: jenseisert@gmail.com).

Complex quantum systems with many degrees of freedom can now be controlled with unprecedented precision, giving rise to applications in quantum metrology[1], quantum information[1,2] and quantum simulation[3,4]. This holds true specifically for architectures based on trapped ions[5] and ultracold atoms[3,6–8], where large system sizes can now routinely be realized, while still maintaining control down to the level of single constituents. In the light of this development, the mindset has shifted when it comes to the assessment and verification of preparations of quantum states. Traditionally, experiments are being used as a vessel to test the validity of theoretical models by comparing their predictions to specific experimental output. With quantum experiments of many degrees of freedom becoming significantly more accurate, an attitude of 'quantum engineering' and quantum simulation is taking over. Compared with the traditional mindset, one does not compare the experimental data to predictions from theoretical models, but rather uses the full capabilities of the experimental setup as an investigative tool for the physical situation at hand. Triggered by this development and driven by the goal to maximize the information extracted from the experiment, the standards in quantum system identification have substantially risen. Quantum-state tomography[9–11] fulfils this need for precise and model-independent quantum-state identification. It asks the question: given data, what is the unknown quantum state compatible with those data? Maybe unsurprisingly, the interest in the field of quantum system identification and quantum-state tomography has exploded in recent years[10–13].

For many degrees of freedom, unqualified quantum state tomography must be inefficient in the system size, as exponentially many numbers need to be specified. This problem has given way to the insight that practically only the states found in experiments need to be reconstructed, which form only a small subset of the full Hilbert space[14,15]. Accordingly, more efficient tomography tools[9] have been developed, ranging from quantum compressed sensing[10] (for states of approximately low rank), over permutation-invariant tomography, to matrix-product state tomography[11–13,16]. These approaches are based on using the right 'data set' having the appropriate 'sparsity structure' to capture quantum many-body systems. For discrete systems, matrix-product states efficiently capture the low-energy behaviour of locally interacting models and a large body of literature in the condensed-matter context backs up this intuition of the 'physical corner of Hilbert space'[14,15,17].

In this work, we consider continuous systems, in which the tomographic problem is aggravated due to the fact that, in principle, infinitely many degrees of freedom need to be reconstructed. On the basis of the notion of sparsity, we present a novel quantum-field tomography procedure relying on the class of continuous matrix-product states (cMPS)[18,19]. This approach will allow us to give evidence that the state encountered in the laboratory is well approximated by a representative of this class.

Results

Quantum-field tomography. We apply our procedure to non-equilibrium experiments of a continuous quantum gas of one species of bosonic particles whose correlation behaviour can be captured by translation invariant states of the form

$$|\Psi_{Q,R}\rangle = \mathrm{Tr}_{\mathrm{aux}}\left[\mathcal{P}e^{\int_0^L \mathrm{d}x\left(Q\otimes\hat{\mathbb{1}}+R\otimes\hat{\psi}^\dagger(x)\right)}\right]|\Omega\rangle. \quad (1)$$

Here $\hat{\psi}(x)$, $x\in[0,L]$ are the canonical bosonic field operators, $|\Omega\rangle$ is the vacuum state vector, Q, $R\in\mathbb{C}^{d\times d}$ are matrices acting on an auxiliary d-dimensional space and completely parametrize the state. L is the length of the closed physical system, $\mathcal{P}$ denotes the path ordering operator and $\mathrm{Tr}_{\mathrm{aux}}$ traces out the auxiliary space. The bond dimension d takes the same role as the bond dimension for matrix-product states: Low entanglement states are expected to be well approximated by cMPS of low bond dimension; in turn, for suitably large d, every quantum-field state can be approximated.

We employ our reconstruction procedure to perform quantum state tomography for a one-dimensional (1D) system of ultracold Bose gases, an architecture that provides one of the prime setups for exploring the physics of interacting quantum fields[6,20,21]. The experiment consists of a large 1D quasi-condensate that is trapped using an atom chip[22]. To bring the system out of equilibrium, a split transversal to the condensate direction is performed. The subsequent out-of-equilibrium dynamics after the quench leads to apparent equilibration, prethermalization and thermalization[6,23,24]. In the middle of the trap, the system can be well approximated by two parallel quantum fields that are homogeneous and translationally invariant.

The experiment proceeds by performing a joint time-of-flight measurement of the two quasi-condensates. Since the experimentally measured images are single-shot measurements, repeating the experiment many times with identical initial conditions allows to extract the phase difference $\hat{\theta}_x$ of the two quasi-condensates at different longitudinal position x and construct higher order correlation functions[6,25]. The phase correlation functions are defined as

$$C^{(n)}(x_1,\ldots,x_n) = \mathrm{Re}\left\langle e^{\mathrm{i}\left(\hat{\theta}_{x_1}-\hat{\theta}_{x_2}+\hat{\theta}_{x_3}-\ldots+\hat{\theta}_{x_{n-1}}-\hat{\theta}_{x_n}\right)}\right\rangle, \quad (2)$$

where $\hat{\theta}_x$ are the measured phase differences and the angular brackets denote the ensemble average (Methods section).

To capture these correlation function in terms of a cMPS, we use a description in terms of effective field operators for the phase difference

$$\hat{\psi}^\dagger(x) = \hat{n}(x)^{\frac{1}{2}}e^{\hat{\theta}_x} \quad (3)$$

where $\hat{n}$ are density operators. As no density information could be obtained from the experiment in its current form, the expectation value of these operators remains unknown and our work is a partial reconstruction of the state. However, the obtained cMPS contains its full phase correlation behaviour.

Using this description, we can write an n-point phase correlation functions as

$$C^{(n)}(x_1,\ldots,x_n) = \mathrm{Re}\left\langle \hat{n}(x_1)^{-\frac{1}{2}}\hat{\psi}^\dagger(x_1)\hat{\psi}(x_2)\hat{n}(x_2)^{-\frac{1}{2}}\ldots\right\rangle. \quad (4)$$

Since it is sufficient for performing the tomography procedure, we will use the correlation information of the normal ordered subset with $x_1\le x_2\le\cdots\le x_n$ of the even-order correlation functions. In the cMPS language, assuming translation invariance and the thermodynamic limit, this can be reformulated as

$$C^{(n)}(\tau_1,\ldots,\tau_{n-1}) = \sum_{\{k_j\}=1}^{d^2}\rho_{k_1,\ldots,k_{n-1}}e^{\lambda_{k_1}\tau_1}\ldots e^{\lambda_{k_{n-1}}\tau_{n-1}} \quad (5)$$

with $\tau_k = x_{k+1}-x_k$,

$$\rho_{k_1,\ldots,k_{n-1}} = M^{-1}_{1,k_{n-1}}M_{k_{n-1},k_{n-2}}\ldots M^{-1}_{k_2,k_1}M_{k_1,1}, \quad (6)$$

λ_k being the eigenvalues of the transfer matrix T, and M being $\overline{R}^{\frac{1}{2}}\otimes R^{-\frac{1}{2}}$ in the diagonal basis of T (Methods section)[16]. The reconstruction proceeds by first extracting the eigenvalues λ_k from the two-point correlation function and in a second step, determining a compatible M matrix[26] from the four-point correlators.

Data analysis. We find that a cMPS with $d=2$, corresponding to four reconstructed poles and a 4×4 matrix M, matches the data. This indicates that the correlation function has a simple structure as one would expect from such local physical interactions (specifically based on previously explored descriptions in terms of a Luttinger liquid theory[6]). More importantly, no previously known theoretical description of the physical situation at hand is needed since the cMPS ansatz can be applied to any locally interacting quantum field. To estimate the performance of the reconstruction of the four-point correlation function, we use the mean relative deviation (Methods section), and find a small error of 1.4%, which is of the same magnitude as the experimental errors[6].

Approximating a correlation function can be done in many ways and it is, *a priori*, not clear that one has truly gained knowledge about the state. The advantage of the cMPS ansatz is that the approximation performed is sufficient to fully reconstruct the phase correlation behaviour of the cMPS. We build trust in the reconstructed state by using it to predict higher order correlation functions, which in turn can be experimentally checked. This provides an excellent benchmark for our procedure and allows us to estimate the quality of our guess for the unknown experimental state. Specifically, we obtain an error of 3.2% for the six-point function (Fig. 1), estimated with bootstrapping techniques. This shows that the reconstruction of the full correlation behaviour of the state was successful, providing a proof-of-principle application for efficient state tomography of interacting many-body quantum fields.

We have performed our reconstruction of the six-point correlator for different hold times after the quench and observe that the fit quality drops substantially with increasing time with mean relative deviations of 3.2%, 10.7% and 34.1% for times $t=3$, 7 and 23 ms, respectively (Fig. 2). There are several possible explanations for this decrease in reconstruction quality. While quantum-field tomography necessarily has to rely on a finite-dimensional 'data set', it is clear that not all situations can be captured equally well by the approach proposed here. This method applies to states of low entanglement, a situation expected to be present for ground states or states in non-equilibrium following quenches for short times. It will surely be difficult to capture highly entangled or thermal states, which are expected to have a high description complexity, with these tools[26].

Discussion

The physics of sudden quenches in discrete settings is usually connected to a linear entanglement growth with time[15,23,27], while for each time satisfying an area law in space[15]. Note that while the continuous physical system at hand can be well captured with a free Tomonaga–Luttinger liquid model[28,29], the states of the system can still be strongly entangled, in the sense that entanglement entropies across any real-space cut of the system are, in principle, arbitrarily large. It is precisely this spatial entanglement that will surely influence the quality of tensor network descriptions of the state and that is a key factor for the quality of any cMPS reconstruction[26]. Since our cMPS reconstruction with $d=2$ is only well-suited for states with low entanglement, a similar entanglement buildup for the performed sudden quench of quantum fields would be a natural explanation. Indeed, such light cone dynamics for the correlations of these systems[6,30,31] have recently been made explicit experimentally. Such entanglement growth could conceptually be unveiled by investigating how the fit quality changes when the bond dimension is increased. Given the structure of the data set (analysis contained in the Methods section) and the increase of experimental errors with hold time, the exploration of this observation lies outside the scope of this work, but is surely an interesting topic for the near future.

Experimental imperfections or the remaining actual temperature could be other sources for the decrease in fit quality with hold time, as they lead to a mixed state, thus impeding our description in terms of pure states. Previous studies, however, successfully described the system in terms of a pure state Luttinger liquid, even for long evolution times[31]. Moreover, the experimental data was taken in the middle of the trap, where, initially, the assumption of translational invariance holds up to excellent accuracy. For long hold times after the quench, however, regions outside of the center of the trap will have an influence on the behaviour of the system in the middle[6], thus making the data less translational invariant (Methods section).

The work presented here is surely a first step in the direction of a larger programme, advocating a paradigm change in the evaluation of experimental data from atomic-optical architectures. Instead of comparing predictions of an assumed theoretical model with data, one puts the data into the focus of attention and attempts a reconstruction in the mindset of quantum tomography. This, in particular, seem an important development in the context of quantum simulators, which have

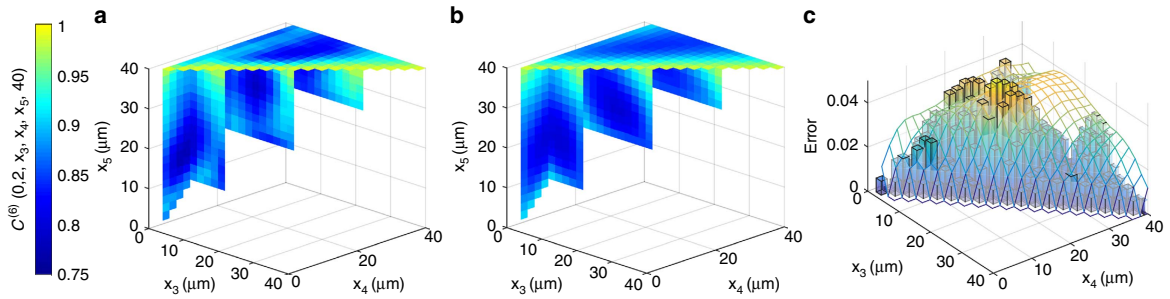

Figure 1 | Projections of the measured and predicted six-point correlation function. We show projections of the relevant sections of the (**a**) experimental and (**b**) predicted six-point function for a hold time after the quench of $t=3$ ms. This image shows the volumetric elements of certain projections of the high-dimensional six-point correlation function array and demonstrates a great overall agreement between experimental data and the predicted correlation data. In **c**, the absolute difference between the experimental and the predicted data points for the projection $C^{(4)}(0, 2, x_3, x_4)$ is shown as a bar plot, the statistical uncertainties of the data as a transparent mesh. More quantitatively, as a figure of merit for measuring the performance of the reconstruction, we use the mean relative deviation over all indices belonging to the relevant simplex of the data with $x_1\leq x_2\leq\cdots\leq x_6$ (Methods section) and find a mean error of 2.5% and a maximum relative deviation of 9.1%.

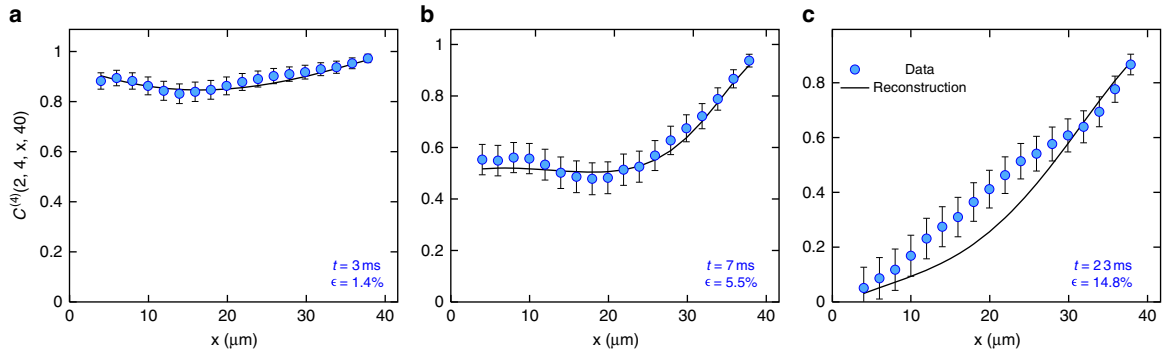

Figure 2 | Projections of the four-point correlation function. We show projections of the four-point correlators for a hold time of (**a**) 3 ms, (**b**) 7 ms and (**c**) 23 ms. The quality of the cMPS ansatz decreases substantially with the hold time, with a mean relative deviation ϵ of the full four-point correlator as indicated in the figures. This increase of the deviation with hold time could be seen as an indicator for the non-equilibrium processes in the system (see main text), but is presumably also related to the increase in s.e. in the experiment, as indicated by the error bars (Methods section).

the potential to address questions on interacting quantum systems that are inaccessible with classical means. While partial information of the results of a quantum simulator can easily be accessed, a full read-out necessarily corresponds to performing quantum tomography where feasible tools are still lacking. The present work offers a step forward and presents a novel tool to obtain and build trust in the complete results of a quantum simulation without having to include any information of the underlying Hamiltonian of the system.

Methods

Experiment. A single specimen of an ultracold gas of ^{87}Rb atoms is prepared using evaporative cooling on an atom chip. The final temperature and the chemical potential of the gas are both well below the first radially excited state of the trapping potential, implementing a 1D bosonic system that is well approximated by the Lieb–Liniger model. The systems contain several thousand atoms and spread over sizes as large as 100 μm. A sudden global quench is realised by transversally splitting the gas into two mutually coherent halves[32], leading to an out-of-equilibrium, approximately pure state. The setup in principle allows for different splitting procedures, in particular an experimental scheme to test the Unruh effect with a specially modelled split has recently been proposed[33]. Subsequently, this non-equilibrium system is let to evolve in the trap for a variable hold time. Its dynamical states are probed using matter wave interferometry in time-of-flight, which enables the direct measurement of the local relative phase θ_x. Since the experimentally measured images are single-shot measurements, repeating the experiment many times with identical initial conditions allows to measure not only the mean of the correlations, but also higher order correlation functions are accessible[6]. The corresponding correlation functions are constructed by averaging over $\sim$150 experimental realizations.

We are restricted to even-order correlation functions in the experiment. The reason for this is the fact that many experimental realizations are needed to construct the correlation functions. Each of these experimental realizations provides us with a measurement of the relative phase $\theta_x = \phi(x) + \varphi$. Here ϕ is the actual fluctuating phase that contains the interesting many-body physics and φ is a small global phase diffusion that is random in every experimental realization[32]. This global phase diffusion results from small shot-to-shot fluctuations in the electrical currents that create the trapping potential. These cause small random imbalances of the double well, leading to random and unknown values for φ. For the even-order correlation functions only differences between the θ at different positions need to be evaluated. Consequently, the global shifts φ cancel automatically. However, for odd-order correlation functions contributions $\sim e^{i\varphi}$ remain. Hence, the measured result does not only contain the pure dynamics, but is significantly perturbed by the unknown fluctuations of φ.

Reconstruction procedure. To make the correlation function in equation 2 directly accessible to our reconstruction procedure, we write it in terms of field operators $\hat{\psi}(x)$. For this purpose, we use the fact that $\hat{\theta}_x$ commutes for different positions and employ the polar decomposition to construct an effective field operator

$$\hat{\psi}^\dagger(x) = \hat{n}(x)^{1/2} e^{i\hat{\theta}_x}, \tag{7}$$

where $\hat{n}(x) = \hat{\psi}^\dagger(x)\hat{\psi}(x)$ is taken to be the density of one of the two condensates. The construction ensures that these effective field operators indeed fulfil the correct commutation relations. Equation 4 follows immediately.

In the cMPS formalism, the translationally invariant correlation functions in equation 4 can be directly calculated in terms of the cMPS variational parameter matrices R and Q in the thermodynamic limit as

$$C^{(n)}(x_1, \ldots, x_n) = \mathrm{Tr}\left(\lim_{L\to\infty} e^{T(L-x_n)}(\overline{R}^{\frac{1}{2}} \otimes R^{-\frac{1}{2}}) e^{T\tau_{n-1}} \cdot \ldots (\overline{R}^{-\frac{1}{2}} \otimes R^{\frac{1}{2}}) e^{T\tau_2} (\overline{R}^{\frac{1}{2}} \otimes R^{-\frac{1}{2}}) e^{T\tau_1} (\overline{R}^{-\frac{1}{2}} \otimes R^{\frac{1}{2}})\right)$$

with the transfer matrix

$$T := \overline{Q} \otimes \mathbb{1}_d + \mathbb{1}_d \otimes Q + \overline{R} \otimes R, \tag{9}$$

and positive distances $\tau_j = x_{j+1} - x_j$ for $j = 1,\ldots,n-1$. The overline denotes complex conjugation. This form of the correlator can be derived by the correspondences between field operators and variational matrices as described in refs 18,19.

By writing all the matrices in the basis where the transfer matrix T is diagonal and performing the limit $L \to \infty$, the correlation function takes the form

$$C^{(n)}(\tau_1, \ldots, \tau_{n-1}) = \sum_{\{k_j\}=1}^{d^2} \rho_{k_1,\ldots,k_{n-1}} e^{\lambda_{k_1}\tau_1} \cdot \ldots \cdot e^{\lambda_{k_{n-1}}\tau_{n-1}}. \tag{10}$$

The λ_k are the eigenvalues of the transfer matrix T, also known as poles and the pre-factors, usually refered to as residues, are

$$\rho_{k_1,\ldots,k_{n-1}} = M^{-1}_{1,k_{n-1}} M_{k_{n-1},k_{n-2}} \cdot \ldots \cdot M^{-1}_{k_2,k_1} M_{k_1,1}, \tag{11}$$

with

$$M = X^{-1}(\overline{R}^{\frac{1}{2}} \otimes R^{-\frac{1}{2}})X, \tag{12}$$

where X has been chosen such that $X^{-1}TX$ is diagonal[16,26]. For a fixed bond dimension, there are in general d^2 poles and $M \in \mathbb{C}^{d^2\times d^2}$. Note that this is different from the definition in ref. 26 where the matrix M stems from density-like correlation functions

$$\hat{O} = \prod_j \hat{n}_j. \tag{13}$$

There, according to the calculus of cMPS correlation functions, the field operator term for each position corresponds to the matrix $\overline{R} \otimes R$.

Note that equating two consecutive indices k_j, k_{j+1} in the n-point function in equation 10 leads to a $(n-2)$-point function, as expected from equation 2. Specifically, there are many equivalent projections of a four-point function that correspond to two-point functions. However, due to imperfections (that is, deviations from translational invariance), the experimental realizations of these projections are not identical. Averaging over the projections leads an expression of the same form of a two-point correlation function from a translationally invariant cMPS as follows,

$$C^{(2)}(\tau) = \sum_{k=1}^{d^2} \rho_k e^{\lambda_k \tau}. \tag{14}$$

The reconstruction starts by extracting the eigenvalues λ_k from the averaged two-point correlation function using a least-squares fit and under the assumption of translational invariance for the modelled system. The suitable bond dimension for

the data at hand can already be judged at this point, by analysing the structure of the two-point correlation function. To determine all entries of M, n-point functions with $n>2$ have to be taken into account, since for $n=2$, only the entries $M_{1,k}^{-1}$ and $M_{k,1}$ appear, see equation 15. Since multiplying M with a constant and conjugating it with a diagonal matrix whose first entry is equal to one leaves all properties considered in this work invariant, we can require that $M_{1,k}=1$ for each $k=1,\ldots,d^2$ (refs 16,26). The remaining independent entries of the M matrix are fixed by included four-point correlation data. For this, we use a Nelder–Mead simplex algorithm that varies the parameters of the M matrix, and calculates the corresponding residues according to

$$\rho_{k_1,k_2,k_3} = M^{-1}_{1,k_3} M_{k_3,k_2} M^{-1}_{k_2,k_1} M_{k_1,1}. \quad (15)$$

Each choice of an M matrix thus gives a prediction for the four-point correlators and the agreement with the experimental data is taken as the quality indicator for the algorithm. Working with a cMPS with bond dimension $d=2$ and relying on a set of 100 random initial numerical seeds proved to be sufficient for approximating the measurement data well. Taking into account the gauge and symmetry arguments[26], the employed cMPS, with bond dimension $d=2$ in terms of λ_k and M, has 15 independent parameters in total.

As discussed in the main text, we see a significant decrease of the fit quality with hold time. There are many issues entering here. One would naturally expect that entanglement entropies after the sudden quench grow over time leading to the need for a larger bond dimension. This is presumably the case, but in our analysis, this is mostly masked by two other effects. First, the statistical error in the experiment increases substantially with the hold time, making the data for longer times considerably less reliable (Fig. 2) and also questioning our fit in terms of a pure state. What is more, the translational invariance assumption is slowly violated as the hold time increases. This is not surprising, since the light-cone-like dynamics of the trapped system give good reason to believe that trap effects need time to enter the center part of the system. As a quantitative probe to estimate how translational invariant the data are, we consider the two-point correlation function at 21 different points and calculate the variance over those different positions for variable distances. The mean of those variances gives a good indicator on how much the two-point function varies depending on the position it is evaluated at. We find for the hold times $t=3$, 7 and 23 ms deviations from translational invariance of 0.3×10^{-2}, 5.4×10^{-2} and 8.3×10^{-2}, clearly indicating that for longer hold times, our assumption of translational invariance is considerably less accurate. Given these limitations of the data set and the fact that the two-point functions averaged over different positions does not possess a rich enough structure, we feel that using a bond dimension larger than $d=2$ would be overfitting. Let us point out that this is by no means a limitation of our method as such, as reconstructions with higher bond dimension could easily be performed using matrix-pencil methods as described in ref. 26.

Quantifying the statistical compatibility and error analysis. To quantify the error of our tomography procedure, we use the relative mean deviation with respect to the fitted (reconstructed) data,

$$\epsilon = \left(|S|^{-1} \sum_{\mathbf{x}\in S} \frac{|C(\mathbf{x}) - C_{\mathrm{rec}}(\mathbf{x})|^2}{|C_{\mathrm{rec}}(\mathbf{x})|^2} \right)^{1/2}, \quad (16)$$

where S is the set of all data points $\mathbf{x}=(x_1,\ldots,x_n)$ with $x_1\leq x_2\leq\cdots\leq x_n$, and $|S|$ denotes the number of elements in S. In addition, to estimate the robustness of our algorithm, we employ a bootstrapping method (see, for example, ref. 34). Namely, starting with the reconstructed four-point function from the experimental data, we add Gaussian noise with zero mean and s.d. given by the statistical uncertainties from the experiment. Subsequently, we perform our cMPS tomography procedure and reconstruct the six-point function. We repeated this procedure 100 times and computed the entry-wise relative standard deviation of the six-point functions. For the average over all entries, we obtain a deviation of 1.1% (with a maximum relative s.d. of 2.8%). This confirms that our reconstruction procedure is robust to the errors we expect in the experiment.

References

1. Wineland, D. J. & Leibfried, D. Quantum information processing and metrology with trapped ions. *Laser Phys. Lett.* **8,** 175–188 (2011).
2. Schindler, P. *et al.* Experimental repetitive quantum error correction. *Science* **332,** 1059–1061 (2011).
3. Bloch, I., Dalibard, J. & Nascimbene, S. Quantum simulations with ultracold quantum gases. *Nat. Phys.* **8,** 267–276 (2012).
4. Cirac, J. I. & Zoller, P. Goals and opportunities in quantum simulation. *Nat. Phys.* **8,** 264–266 (2012).
5. Blatt, R. & Roos, C. F. Quantum simulations with trapped ions. *Nat. Phys.* **8,** 277–284 (2012).
6. Langen, T., Geiger, R., Kuhnert, M., Rauer, B. & Schmiedmayer, J. Local emergence of thermal correlations in an isolated quantum many-body system. *Nat. Phys.* **9,** 640–643 (2013).
7. Hofferberth, S., Lesanovky, I., Fischer, B., Schumm, T. & Schmiedmayer, J. Non-equilibrium coherence dynamics in one-dimensional bose gases. *Nature* **449,** 324–327 (2007).
8. Trotzky, S. *et al.* Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional Bose gas. *Nat. Phys.* **8,** 325–330 (2012).
9. Paris, M. & Rehacek, J. *Quantum state estimation, Lecture Notes in Physics* (Springer, 2004).
10. Gross, D., Liu, Y.-K., Flammia, S. T., Becker, S. & Eisert, J. Quantum state tomography via compressed sensing. *Phys. Rev. Lett.* **105,** 150401 (2010).
11. Cramer, M. *et al.* Efficient quantum state tomography. *Nat. Commun.* **1,** 149–155 (2010).
12. Baumgratz, T., Gross, D., Cramer, M. & Plenio, M. B. Scalable reconstruction of density matrices. *Phys. Rev. Lett.* **111,** 020401 (2013).
13. Ohliger, M., Nesme, V. & Eisert, J. Efficient and feasible state tomography of quantum many-body systems. *N. J. Phys.* **15,** 015024 (2013).
14. Perez-Garcia, D., Verstraete, F., Wolf, M. M. & Cirac, J. I. Matrix product state representations. *Quant. Inf. Comp.* **5 & 6,** 401–430 (2006).
15. Eisert, J., Cramer, M. & Plenio, M. B. Area laws for the entanglement entropy. *Rev. Mod. Phys.* **82,** 277–306 (2010).
16. Hübener, R., Mari, A. & Eisert, J. Wick's theorem for matrix product states. *Phys. Rev. Lett.* **110,** 040401 (2013).
17. Verstraete, F., Cirac, J. I. & Murg, V. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Adv. Phys* **57,** 143–224 (2008).
18. Verstraete, F. & Cirac, J. I. Continuous matrix product states for quantum fields. *Phys. Rev. Lett.* **104,** 190405 (2010).
19. Osborne, T. J., Eisert, J. & Verstraete, F. Holographic quantum states. *Phys. Rev. Lett.* **105,** 260401 (2010).
20. Jacqmin, T., Armijo, J., Berrada, T., Kheruntsyan, K. V. & Bouchoule, I. Sub-poissonian fluctuations in a 1d bose gas: From the quantum quasi-condensate to the strongly interacting regime. *Phys. Rev. Lett.* **106,** 230405 (2011).
21. Gring, M. *et al.* Relaxation and prethermalization in an isolated quantum system. *Science* **337,** 1318–1322 (2012).
22. Reichel, J. & Vuletic, V. *Atom chips* (Wiley, 2011).
23. Cramer, M., Dawson, C. M., Eisert, J. & Osborne, T. J. Exact relaxation in a class of non-equilibrium quantum lattice systems. *Phys. Rev. Lett.* **100,** 030602 (2008).
24. Eisert, J., Friesdorf, M. & Gogolin, C. Quantum many-body systems out of equilibrium. *Nature Physics* **11,** 124 (2015).
25. Smith, D. A. *et al.* Prethermalization revealed by the relaxation dynamics of full distribution functions. *N. J. Phys.* **15,** 075011 (2013).
26. Steffens, A., Riofrío, C., Hübener, R. & Eisert, J. Quantum field tomography. *New J. Phys.* **16,** 123010 (2014).
27. Schuch, N., Wolf, M. M., Vollbrecht, K. G. H. & Cirac, J. I. On entropy growth and the hardness of simulating time evolution. *N. J. Phys.* **10,** 033032 (2008).
28. Kitagawa, T., Imambekov, A., Schmiedmayer, J. & Demler, E. The dynamics and prethermalization of one-dimensional quantum systems probed through the full distributions of quantum noise. *N. J. Phys.* **13,** 073018 (2011).
29. Langen, T. *et al.* Experimental observation of a generalized Gibbs ensemble. *Science* **348,** 207 (2015).
30. Cheneau, M. *et al.* Light-cone-like spreading of correlations in a quantum many-body system. *Nature* **481,** 484–487 (2012).
31. Geiger, R., Langen, T., Mazets, I. E. & Schmiedmayer, J. Local relaxation and light-cone-like propagation of correlations in a trapped one-dimensional Bose gas. *N. J. Phys.* **16,** 053034 (2014).
32. Schumm, T. *et al.* Matter-wave interferometry in a double well on an atom chip. *Nat. Phys.* **1,** 57–62 (2005).
33. Agarwal, K. *et al.* Chiral pre-thermalization in supersonically split condensates. *Phys. Rev. Lett.* **113,** 190401 (2014).
34. Shao, J. *Mathematical Statistics* (Springer, 2003).

Acknowledgements

We thank the EU (SIQS, AQuS, RAQUEL), the ERC (TAQ and QuantumRelax), the Studienstiftung des Deutschen Volkes, the FQXi and the BMBF (QuOReP) and the Austrian Science Fund (FWF) SFB-Foqus (P4010) for support. B.R., T.S., and T.L. acknowledge support by the FWF Doctoral Programme CoQuS (W1210). We thank the KITP in Santa Barbara for hospitality and acknowledge R. Geiger and A.H. Werner for useful discussions.

Author contributions

A.S., R.H., C.R. and J.E. developed the tomography procedure. M.F., A.S., and C.R. carried out the numerical analysis. T.L., B.R. and T.S. set up the experiment and carried out the measurements. J.S. conceived the experiment and supervised the measurements. M.F., J.E. and A.S. wrote the manuscript with substantial contributions from all authors.

Additional information

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Steffens, A. *et al.* Towards experimental quantum-field tomography with ultracold atoms. *Nat. Commun.* 6:7663 doi: 10.1038/ncomms8663 (2015).

4 QUANTUM ALGORITHMS

Classical computers have become orders of magnitude faster within the last decades, laying the foundation for the digital revolution. However, *Moore's law*, which roughly states that the performance of computers doubles about every two years and has accurately been describing the technological progress in computational power since the 1970s, is nearing its end[149]: The elements on integrated circuits have reached sizes of just few nanometers, the same scale as atomic diameters, and quantum effects such as unwanted quantum tunneling make the physical limitations of the current silicon-based technology apparent. Yet, even polynomially scaling problems become infeasible to compute for sufficiently large problem size.

To overcome a possible standstill, different approaches are being explored ranging from materials[150] to 3D integration of chips[151]. Beyond this, quantum effects could instead be harnessed with a fundamentally different architecture—by using *quantum computers*, promising large, in some instances even exponential speedups for algorithms. First discussed by Richard Feynman in the context of efficiently simulating quantum systems[20], a general framework has been developed[21,22,152,153], and possible physical implementations with, e.g., cold trapped ions[154] or photons[155] are being investigated[20].

Instead of working with bits that can have either the state "0" or the state "1", a quantum computer will work on quantum bits (*qubits*), which can be in the quantum state $|0\rangle$, the quantum state $|1\rangle$ or any superposition

$$|\varphi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \tag{29}$$

with $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$ for normalization. One can think of $|0\rangle$ and $|1\rangle$ as the ground state and the excited state of a two-level system. An n-qubit state consists of the tensor product of its constituting n single-qubit states. Quantum computations are performed by operations on a (multiple) qubit state $|\varphi\rangle$ that can be modeled by a unitary operator $U = \mathrm{e}^{-\mathrm{i}Ht}$ (preserving the norm of $|\varphi\rangle$) with a Hamiltonian H acting on $|\varphi\rangle$ for the time t. General unitary operations can be decomposed into single- and two-qubit operations, resulting in *quantum circuits* consisting of easier to realize *quantum gates*.[22] The outcomes of the computations are obtained by measuring the evolved quantum state.

The state of a qubit can be represented by a vector $(\alpha, \beta) \in \mathbb{C}^2$ consisting of its two amplitudes α and β, whereas a state $|\varphi_1\rangle \otimes |\varphi_2\rangle \otimes \cdots \otimes |\varphi_n\rangle$ with n qubits requires a vector in $\mathbb{C}^{(2^n)}$ for description, i.e., 2^n complex parameters (up to normalization)—as opposed to n numbers for classical bits. What is an obstacle for estimating a state via quantum tomography, as discussed in section 2.1, is a feature for quantum computations: The evolution of an n qubit state takes place in an exponentially larger Hilbert space, in principle allowing for the simultaneous manipulation of 2^n numbers, which is also called *quantum parallelism*. However, reading out the computed information in form

[20] A photonic system for potential quantum computations is also the experiment at the basis of publication [1] (see section 2.3).

of a general quantum state (again requiring quantum tomography) is hard and algorithms on a quantum computer need to be carefully designed to output efficiently obtainable results—not exponentially large state vectors—which poses an important condition for quantum algorithms. For example, the n-qubit *quantum Fourier transform*, providing the Fourier transform of a signal of size 2^n encoded in the amplitudes of a quantum state, requires only $\Theta(n^2)$ quantum gates, whereas the classical *fast Fourier transform* (FFT) needs $\Theta(n\,2^n)$ operations for processing a signal of the same size.[22] Quantum algorithms like the famous *Shor algorithm* for integer factorization[21], which build on the quantum Fourier transform, hence output an amplitude overlap or a qubit-wise measurable product state, consisting of uncorrelated qubits that can be measured independently and hence with effort linearly in the number of qubits.

But also accessing the input data needs to be sufficiently fast. The input could always be provided by the output of a previous quantum subroutine as a superposition in a quantum state. Other approaches involve *quantum random access memory* (QRAM)[158,159], which supplies input signals via quantum superpositions of memory cells and allows for an access time logarithmic in the data size. Quantum computers are susceptible to noise and decoherence, calling for *quantum error correction* protocols, such as *stabilizer codes* [160], which include appending ancillary qubits to the signal qubits. A succinct list of conditions on constructing quantum computers is presented in Ref. [161].

Polynomial quantum speedups are of interest as well: *Grover's algorithm*[162] is employed for searching an unsorted database with N entries just in $O(\sqrt{N})$ steps and can be generalized to *amplitude amplification* [163], iteratively evolving an input quantum state to the subspace associated with the desired solution of a problem. Building on this, an interesting set of quantum algorithms is presented in Ref. [164], providing a quadratic quantum speedup for semidefinite programs and connecting quantum computing with compressed sensing. With a quantum semidefinite solver, matrix completion problems as in Eq. (13) and in particular the compressive quantum state tomography problem Eq. (15) could be solved on a quantum computer. Using methods from *shadow tomography*[37], certain semidefinite problems can even be exponentially accelerated.[165]

A new class of quantum algorithms was sparked by a method[166] that allowed for the inversion of a linear equation system $Ax = y$ with sparse matrix $A \in \mathbb{R}^{N\times N}$ using only $O(\text{poly}\log N)$ operations—as opposed to $O(N)$ classically.[22] Applications and extensions included quantum versions of least squares data fitting[168,169], *support vector machines* [170] for determining a separating hyperplane between two classes of data points, an important tool for machine learning, and *principal component analysis* [171] for determining the leading eigenspaces of a matrix. The latter paper also expanded the scope of viable matrices in this class of quantum algorithms to dense (i.e., non-sparse), low-rank, positive semidefinite matrices.

[21] The computational effort for factorizing an n-bit integer, up to logarithmic factors, is $O(n^3)$ using Shor's quantum algorithm versus $e^{O(n^{1/3})}$ using the best known classical algorithm[156]. Since the classical computational complexity of integer factorization is yet unknown, polynomially scaling classical algorithms for this problem could theoretically be possible. Providing provable (superpolynomial) quantum speedups is a major topic of the field of *quantum supremacy* [157].

[22] This concerns sparse matrices; general matrices currently require $O(n^k)$, with $k \approx 2.37$, classical operations.[167]

A veritable "quantum algorithm zoo"[172] has emerged in recent years, including algorithms for, e.g., subset finding[173] or analyzing electrical networks[174]. See Ref. [175] for a quantum algorithm overview. Still, due to the intrinsic challenges in developing quantum algorithms, so far only few quantum analogues of classical signal processing algorithms exist.

The following publication [4][23] also builds on Ref. [166], complementing the quantum algorithm zoo with a procedure that provides a quantum speedup for the spectral estimation of a signal and scales polynomially in the logarithm of the signal size. At its core lies a *matrix pencil method* [139], similar to the one employed in publication [2] (see section 3.2), which had to be modified for an efficient quantum implementation. Special care was taken to meet the requirements for the feasible applicability of the arising unitary operations, the robust invertibility of the utilized matrices, and fast input/output, as emphasized in Ref. [176]. In the process, new methods were developed that could become useful components for future quantum algorithms. This work is also connected to the *quantum singular value decomposition* publication [6], which is presented in appendix A.2.

[23] Adrian Steffens, Patrick Rebentrost, Iman Marvian, Jens Eisert, and Seth Lloyd, "An efficient quantum algorithm for spectral estimation", *New Journal of Physics* 19:033005, 2017 (DOI:10.1088/1367-2630/aa5e48). Published under a Creative Commons Attribution 3.0 License (creativecommons.org/licenses/by/3.0), © 2017 IOP Publishing.

FIRST-AUTHORED PUBLICATIONS

[1] Experimentally exploring compressed sensing quantum tomography
Adrian Steffens, Carlos A. Riofrío, Will McCutcheon, Ingo Roth, Bryn A. Bell, Alex McMillan, Mark S. Tame, John G. Rarity, and Jens Eisert, *Quantum Science and Technology* 2:025005, 2017.

[2] Quantum field tomography
Adrian Steffens, Carlos A. Riofrío, Robert Hübener, and Jens Eisert, *New Journal of Physics* 16:123010, 2014.

[3] Towards experimental quantum-field tomography with ultracold atoms
Adrian Steffens, Mathis Friesdorf, Tim Langen, Bernhard Rauer, Thomas Schweigler, Robert Hübener, Jörg Schmiedmayer, Carlos A. Riofrío, and Jens Eisert, *Nature Communications* 6:7663, 2015.

[4] An efficient quantum algorithm for spectral estimation
Adrian Steffens, Patrick Rebentrost, Iman Marvian, Jens Eisert, and Seth Lloyd *New Journal of Physics* 19:033005, 2017.

IOP Publishing New J. Phys. 19 (2017) 033005 https://doi.org/10.1088/1367-2630/aa5e48

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
19 October 2016

REVISED
21 January 2017

ACCEPTED FOR PUBLICATION
3 February 2017

PUBLISHED
1 March 2017

PAPER

An efficient quantum algorithm for spectral estimation

Adrian Steffens[1,2,4], **Patrick Rebentrost**[2], **Iman Marvian**[2], **Jens Eisert**[1] **and Seth Lloyd**[2,3]

[1] Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany
[2] Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139, United States of America
[3] Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, United States of America
[4] Author to whom any correspondence should be addressed.

E-mail: adrian.steffens@fu-berlin.de

Keywords: quantum algorithm, spectral estimation, quantum phase estimation, matrix exponentiation, quantum signal processing

Abstract

We develop an efficient quantum implementation of an important signal processing algorithm for line spectral estimation: the *matrix pencil method*, which determines the frequencies and damping factors of signals consisting of finite sums of exponentially damped sinusoids. Our algorithm provides a quantum speedup in a natural regime where the sampling rate is much higher than the number of sinusoid components. Along the way, we develop techniques that are expected to be useful for other quantum algorithms as well—consecutive phase estimations to efficiently make products of asymmetric low rank matrices classically accessible and an alternative method to efficiently exponentiate non-Hermitian matrices. Our algorithm features an efficient quantum–classical division of labor: the time-critical steps are implemented in quantum superposition, while an interjacent step, requiring much fewer parameters, can operate classically. We show that frequencies and damping factors can be obtained in time logarithmic in the number of sampling points, exponentially faster than known classical algorithms.

1. Introduction

Algorithms for the spectral estimation of signals consisting of finite sums of exponentially damped sinusoids have a vast number of practical applications in signal processing. These range from imaging and microscopy [1], radar target identification [2], nuclear magnetic resonance spectroscopy [3], estimation of ultra wide-band channels [4], quantum field tomography [5, 6], power electronics [7], up to the simulation of atomic systems [8]. If the damped frequencies (poles) are known and merely the concomitant coefficients are to be identified, linear methods are readily applicable. In the practically relevant task in which the poles are to be estimated from the data as well, however, one encounters a nonlinear problem, and significantly more sophisticated methods have to be employed.

There are various so-called high resolution spectral estimation techniques that provide precisely such methods: they include matrix pencil methods (MPM) [9], *Prony's method* [10], *MUSIC* [11], *ESPRIT* [12], and *atomic norm denoising* [13]. These techniques are superior to discrete Fourier transform (DFT) in instances with damped signals and close frequencies or small observation time $T > 0$ [14–16] and are preferred over of the Fourier transform in those applications laid out in [1–5, 7, 8]: the DFT resolution in the frequency domain $\Delta\omega$ is proportional to $1/T$, which is especially critical for poles that are close to each other. If the poles are sufficiently damped and close, they cannot be resolved by DFT independently of T. Nonlinear least-squares fitting of the poles or considering higher-order derivatives of the Fourier transform is in general relatively imprecise, sensitive to noise, or unefficient. Nonlinear algorithms such as the MPM can still detect poles, where DFT fails, but are limited to signals composed of finitely many damped sinusoids.

With regard to *quantum algorithms* dedicated to tasks of spectral estimation—algorithms to be run on a quantum computer—the celebrated quantum Fourier transform (QFT) [17] provides an exponential speedup towards the fastest known classical implementations of DFT for processing discretized signals of N samples: classical fast Fourier transform algorithms, on the one hand, take $\Theta(N \log N)$ gates [18], whereas QFT takes

$\Theta(\log^2 N)$ gates to produce a quantum state encoding the Fourier coefficients in its amplitudes. The QFT constitutes a key primitive in various quantum algorithms. In particular, it paved the way for quantum speedups for problems such as prime factoring or order-finding [19]. Regarding spectral estimation, however, QFT inherits the above mentioned properties of its classical counterpart.

The aim of this work is to develop a quantum version of a powerful spectral estimation technique, the MPM, providing an analogous quantum speedup from $O(\text{poly } N)$ to $O(\text{poly } \log N)$ for data given in a suitable format. Hereto, we make use of the fact that establishing eigenvalues and eigenvectors of low-rank matrices—constituting major steps in this algorithm—can be achieved very fast on quantum computers [20]. Given signal data either via the amplitudes of a quantum state or stored in a quantum random access memory (QRAM) [21–23], phase estimation of these matrices can be performed directly. For exponentiating non-sparse operators for phase estimation, we employ quantum principal component analysis (QPCA) [20] and a recently developed oracle-based method [24]. In an additional step, we employ a quantum linear fitting algorithm [25, 26] to determine the summing coefficients and hence all parameters that determine the signal function. In this sense, we can understand our algorithm also as an instance of a *nonlinear* quantum fitting algorithm in contrast to linear curve fitting algorithms [25, 26]. Furthermore, our algorithm can also be employed as a sub-routine in a higher quantum algorithm that requires spectral estimation as an intermediate step. We expect the developed methods to provide valuable novel primitives to be used in other quantum algorithms as well.

2. The classical matrix pencil algorithm

We start by briefly recapitulating the original (classical) matrix pencil algorithm before in section 3, we turn to showing how to implement a quantum version of this algorithm in order to gain an exponential speedup. MPM [9] comprise a family of efficient signal processing algorithms for spectral estimation and denoising of equidistantly sampled complex-valued functions f of the type

$$f(t) = \sum_{k=1}^{p} c_k \, \mathrm{e}^{\lambda_k t} =: \sum_{k=1}^{p} c_k \, \mathrm{e}^{-\alpha_k t} \, \mathrm{e}^{\mathrm{i}\,\beta_k t}, \quad 0 \leqslant t \leqslant T, \tag{1}$$

with the poles $\lambda_k = -\alpha_k + \mathrm{i}\,\beta_k \in \mathbb{C}$, damping factors $\alpha_k \in \mathbb{R}_+$, frequencies $\beta_k \in \mathbb{R}$, and coefficients $c_k \in \mathbb{C}$ for $k = 1, \ldots, p$, where $p \in \mathbb{N}$ is the number of poles. The damping results in a broadening of the spectral lines towards Lorentzian curves. Real-valued functions as a special case can be analyzed as well: here, for each $k = 1, \ldots, p$ either $\lambda_k,\ c_k \in \mathbb{R}$—these terms are non-oscillatory—or there exist $\lambda_{k'},\ c_{k'}$ such that $\lambda_{k'} = \lambda_k^*$ and $c_{k'} = c_k^*$. Clearly, such signals, in which the number of poles p is small and finite, are ubiquitous, or in other instances provide an exceedingly well approximation of the underlying signal.

Algorithm 1. Matrix pencil algorithm.

Data: Discretized signal with components $f_j = \sum_{k=1}^{p} c_k \, \mathrm{e}^{\lambda_k \Delta t \cdot j}$, $j = 0, \ldots, N-1$, $c_k,\ \lambda_k \in \mathbb{C},\ \ \mathfrak{Re}(\lambda_k) \leqslant 0$.
Result: Frequencies $\{\lambda_k\}_{k=1}^{p}$ and coefficients $\{c_k\}_{k=1}^{p}$.
begin
Create the Hankel matrices $F^{(1)} := (f_{j+k-2})_{j,k=1}^{N/2}$ and $F^{(2)} := (f_{j+k-1})_{j,k=1}^{N/2}$ from the signal and compute their (truncated) singular decompositions $F^{(i)} = U^{(i)} S^{(i)} V^{(i)\,\dagger}, \quad i = 1, 2$.
Solve the generalized eigenvalue problem $U^{(1)\,\dagger} F^{(2)} V^{(1)} \, w_k = \mu_k \, S^{(1)} \, w_k$. The p frequencies $\{\lambda_k\}$ can directly be obtained from the p eigenvalues $\{\mu_k\}$.
Create the Vandermonde matrix W from the eigenvalues $\{\mu_k\}$ and invert the linear equation system $Wc = (f_j)$ to obtain the coefficients $\{c_k\}$.

The idea of MPM is to determine the poles $\{\lambda_k\}$ independently from the coefficients $\{c_k\}$ and compare the discretized signal with its translates. Assume that all c_k are nonzero and $\lambda_j \neq \lambda_k$ for $j \neq k$. First, sample the function f equidistantly,

$$f \mapsto (f_j)_{j=0}^{N-1}, \quad f_j = \sum_{k=1}^{p} c_k \, \mathrm{e}^{\lambda_k \Delta t \cdot j}, \tag{2}$$

with sampling interval $\Delta t > 0$. In general, the higher the number of samples N, the more robust the procedure becomes towards noise and the higher the frequencies that can be reconstructed (Nyquist–Shannon sampling theorem [27])—at the expense of computational effort. For clearness, assume that N is even. From the sampled signal, create the Hankel matrices $F^{(1)},\ F^{(2)} \in \mathbb{C}^{N/2 \times N/2}$, defined as

$$F^{(1)} \coloneqq (f_{j+k-2})_{j,k=1}^{N/2} = \begin{bmatrix} f_0 & f_1 & \cdots & f_{N/2-1} \\ f_1 & f_2 & \cdots & f_{N/2} \\ \vdots & \vdots & & \vdots \\ f_{N/2-1} & f_{N/2} & \cdots & f_{N-2} \end{bmatrix} \tag{3}$$

and

$$F^{(2)} \coloneqq (f_{j+k-1})_{j,k=1}^{N/2} = \begin{bmatrix} f_1 & f_2 & \cdots & f_{N/2} \\ f_2 & f_3 & \cdots & f_{N/2+1} \\ \vdots & \vdots & & \vdots \\ f_{N/2} & f_{N/2+1} & \cdots & f_{N-1} \end{bmatrix}. \tag{4}$$

Note that for complex signals, the matrices $F^{(1)}$ and $F^{(2)}$ are symmetric but in general not Hermitian. In other implementations, $F^{(1)}$ and $F^{(2)}$ do not even need to be square. To keep the notation clear, we proceed with square matrices as just defined. Set $\mu_k \coloneqq \mathrm{e}^{\lambda_k \Delta t}$ for $k = 1, \ldots, p$. It is easy to see that $F^{(1)}$ can be factorized as

$$F^{(1)} = M\, D_c\, M^T \tag{5}$$

with the Vandermonde matrix $M \in \mathbb{C}^{N/2 \times p}$,

$$M \coloneqq \underset{k=1,\ldots,p}{(\mu_k^j)_{j=0,\ldots,N/2-1}} = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ \mu_1 & \mu_2 & \cdots & \mu_p \\ \vdots & \vdots & & \vdots \\ \mu_1^{N/2-1} & \mu_2^{N/2-1} & \ldots & \mu_p^{N/2-1} \end{bmatrix} \tag{6}$$

and diagonal matrix $D_c \coloneqq \mathrm{diag}((c_k)_{k=1}^p) \in \mathbb{C}^{p\times p}$. The matrix $F^{(2)}$, on the other hand, can be decomposed as

$$F^{(2)} = M\, D_c\, D_\mu\, M^T \tag{7}$$

with $D_\mu \coloneqq \mathrm{diag}((\mu_k)_{k=1}^p) \in \mathbb{C}^{p\times p}$. Note that equations (5) and (7) are neither the eigenvalue nor the singular value decomposition of $F^{(1)}$ and $F^{(2)}$, respectively; the column vectors of M do not even have to be orthogonal. We can see from these equations that both $F^{(1)}$ and $F^{(2)}$ have rank p, which will in general also be the case for the linear *matrix pencil* [28]

$$F^{(2)} - \gamma\, F^{(1)} = M\, D_c\, (D_\mu - \gamma\, \mathbb{1})\, M^T, \tag{8}$$

unless $\gamma \in \mathbb{C}$ matches an element of the set $\{\mu_k\}_{k=1}^p$. Hence, all μ_k are solutions of the generalized eigenvalue problem (GEVP)

$$F^{(2)} v = \gamma\ F^{(1)} v, \tag{9}$$

with $v \in \mathbb{C}^{N/2}$. The matrix pair $(F^{(2)},\ F^{(1)})$ is in general *regular* and accordingly results in $N/2$ generalized eigenvalues [29]—not all of these correspond to a μ_k. There are different extensions that take care of this issue and increase algorithmic stability (see, e.g., [30]). To make the algorithm accessible to an efficient quantum implementation, we will consider a specific MPM variant, the *direct* MPM [9]: we make use of the singular value decompositions of $F^{(1)}$ and $F^{(2)}$, keeping only the nonzero singular values and the corresponding singular vectors,

$$F^{(i)} = U^{(i)} S^{(i)} V^{(i)\,\dagger},\ \ U^{(i)},\ V^{(i)} \in \mathbb{C}^{N/2\times p}, \tag{10}$$

with $S^{(i)} \in \mathbb{C}^{p\times p}$ for $i = 1,\ 2$. This singular value decomposition of a Hankel matrix of size order $N \times N$ is the time-critical step of the entire algorithm and it scales with $\Theta(N^2 \log N)$ using state-of-the-art classical algorithms [31, 32]. We multiply $U^{(1)\,\dagger}$ from the left and $V^{(1)}$ from the right to

$$F^{(2)} - \gamma\, F^{(1)} = F^{(2)} - \gamma\ U^{(1)} S^{(1)} V^{(1)\,\dagger} \tag{11}$$

and see that the resulting equivalent GEVP

$$U^{(1)\,\dagger} F^{(2)} V^{(1)}\, w = \gamma\, S^{(1)}\, w, \tag{12}$$

with $w \in \mathbb{C}^p$, yields exactly $\{\mu_k\}_{k=1}^p$ as eigenvalues and via $\lambda_k = \log(\mu_k)/\Delta t$ the corresponding poles. The eigenvalues can be retrieved in $\Theta(p^3)$ steps using the QZ algorithm [33]. Although in general it can be numerically favorable to solve the GEVP directly [29], $S^{(1)}$ is an invertible diagonal matrix and it is in practice sufficient to solve the equivalent ordinary eigenvalue problem

$$(S^{(1)})^{-1} U^{(1)\,\dagger} F^{(2)} V^{(1)}\, w = \gamma\, w. \tag{13}$$

The coefficients $\{c_k\}$ are linearly related to the signal and can be obtained by plugging $\{\mu_k\}_{k=1}^p$ into an overdetermined Vandermonde equation system,

$$W\ c = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \mu_1 & \mu_2 & \cdots & \mu_p \\ \vdots & \vdots & & \vdots \\ \mu_1^{N-1} & \mu_1^{N-1} & \cdots & \mu_p^{N-1} \end{bmatrix} \cdot \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix} = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_{N-1} \end{bmatrix}, \tag{14}$$

and computing the least squares solution

$$\hat{c} \coloneqq \arg\min_{\tilde{c} \in \mathbb{C}^p} \|W\ \tilde{c} - f\|_2 \tag{15}$$

in terms of the 2-norm, $\|\cdot\|_2$, e.g. via applying the Moore–Penrose pseudoinverse $W^+ \coloneqq (W^\dagger W)^{-1} W^\dagger$ to the signal vector f. Thus, all parameters that determine the signal are reconstructed.

3. Quantum implementation

In the following, we describe how to implement an efficient quantum analogue of the MPM.

Algorithm 2. Quantum matrix pencil algorithm.

Data: Discretized signal with components $f_j = \sum_{k=1}^{p} c_k\ \mathrm{e}^{\lambda_k \Delta t j}$, $j = 0, \ldots, N-1$, $c_k,\ \lambda_k \in \mathbb{C},\quad \mathfrak{Re}(\lambda_k) \leqslant 0$ either from QRAM or encoded in a quantum state.
Result: Frequencies $\{\lambda_k\}_{k=1}^{p}$ and coefficients $\{c_k\}_{k=1}^{p}$.
begin
Perform concatenated phase estimations via exponentiating Hermitian matrices $\tilde{F}^{(1)}$, $\tilde{F}^{(2)}$ that contain the matrices $F^{(1)}$, $F^{(2)}$, respectively, yielding the p biggest singular values and the overlaps $\{\langle u_j^{(1)} | u_k^{(2)} \rangle\}$ and $\{\langle v_j^{(1)} | v_k^{(2)} \rangle\}$ of the according left and right singular vectors.
Construct the according matrices and solve the eigenvalue problem classically to obtain the poles $\{\lambda_k\}$.
Build a fitting matrix from the poles and obtain the coefficients $\{c_k\}$ via quantum linear fitting.

For an efficient quantum algorithm, we assume that the number of poles p is constant and small relative to the number of samples N, which is a natural setting since in practice, we are often interested in damped line spectra with fewer constituents and higher sampling rates for robustness towards noise. The guiding idea is to condense all arrays of size $O(N)$ in equation (13) into arrays of size $O(p)$ by rewriting the first term in equation (12),

$$\begin{bmatrix} \text{——} \langle u_1^{(1)} | \text{——} \\ \vdots \\ \text{——} \langle u_p^{(1)} | \text{——} \end{bmatrix} \begin{bmatrix} | & \cdots & | \\ |u_1^{(2)}\rangle & & |u_p^{(2)}\rangle \\ | & \cdots & | \end{bmatrix} \begin{bmatrix} s_1^{(2)} & & 0 \\ & \ddots & \\ 0 & & s_p^{(2)} \end{bmatrix} \begin{bmatrix} \text{——} \langle v_1^{(2)} | \text{——} \\ \vdots \\ \text{——} \langle v_p^{(2)} | \text{——} \end{bmatrix} \begin{bmatrix} | & \cdots & | \\ |v_1^{(1)}\rangle & & |v_p^{(1)}\rangle \\ | & \cdots & | \end{bmatrix},$$

as

$$\begin{bmatrix} \langle u_1^{(1)} | u_1^{(2)} \rangle & \ldots & \langle u_1^{(1)} | u_p^{(2)} \rangle \\ \vdots & & \vdots \\ \langle u_p^{(1)} | u_1^{(2)} \rangle & \ldots & \langle u_p^{(1)} | u_p^{(2)} \rangle \end{bmatrix} \begin{bmatrix} s_1^{(2)} & & 0 \\ & \ddots & \\ 0 & & s_p^{(2)} \end{bmatrix} \begin{bmatrix} \langle v_1^{(2)} | v_1^{(1)} \rangle & \ldots & \langle v_1^{(2)} | v_p^{(1)} \rangle \\ \vdots & & \vdots \\ \langle v_p^{(2)} | v_1^{(1)} \rangle & \ldots & \langle v_p^{(2)} | v_p^{(1)} \rangle \end{bmatrix} =: \mathcal{U}\ S^{(2)}\ \mathcal{V}, \tag{16}$$

with $\mathcal{U},\ \mathcal{V} \in \mathbb{C}^{p \times p}$. The singular values $\{s_k^{(j)}\}$ will be obtained via quantum phase estimation [34, 35], the overlaps $\langle v_k^{(i)} | v_l^{(j)} \rangle$ via two concatenated quantum phase estimations. The eigenvalue problem equation (13),

$$(S^{(1)})^{-1}\ \mathcal{U}\ S^{(2)}\ \mathcal{V}\ w = \gamma\ w, \tag{17}$$

is now determined by $2p^2$ complex and $2p$ real numbers, and can easily be evaluated classically in $\Theta(p^3)$ operations, yielding the required poles $\lambda_k = \log(\mu_k)/\Delta t$ for $k = 1, \ldots, p$. Thus, as other efficient quantum algorithms [36, 37], the classical result is a low-dimensional read-out quantity. Otherwise, the read-out costs would neutralize any performance gain in the algorithm. After that, the poles are used as input for a quantum linear fitting algorithm yielding the coefficients $\{c_k\}$. In the following, we describe the individual steps of the quantum algorithm in detail. We start by discussing the quantum preparation of the Hankel matrices.

3.1. Accessing the data

In order to realize a quantum speedup, the signal has to be accessible in a fast and coherent way—otherwise, the read-in process alone would be too costly. The data input for the matrix pencil algorithm consists of a time series

$(f_j)_{j=0}^{N-1}$. We consider two crucially different approaches of data access/availability for the quantum algorithm, with the main focus of this work being on the first approach:

(i) The signal is stored in a quantum accessible form such as quantum RAM. In other words, we are provided with access to the operation

$$|j\rangle|0\rangle \mapsto |j\rangle|f_j\rangle \tag{18}$$

for $j = 0, \ldots, N-1$, with the signal values encoded in binary form in the second quantum register. In order to create the Hankel matrix $F^{(i)} = (f_{j+k+i-3})_{j,k=1}^{N/2} \in \mathbb{C}^{N/2\times N/2}$ and $i = 1, 2$, we can perform the following operation with straightforward index manipulations,

$$|j\rangle|k\rangle|i\rangle|0\rangle \longmapsto |j\rangle|k\rangle|i\rangle|f_{j+k+i-3}\rangle \tag{19}$$

for $j, k = 1, \ldots, N/2$. The ancilla prepared in $|i\rangle$, $i = 1, 2$, will be used in an entirely classical manner. This operation can be used to simulate Hankel matrices via the non-sparse matrix simulation methods of [24, 38].
One way to implement signal access in equation (18) is via QRAM [21, 22]. As discussed in [21, 22], the expected number of hardware elements that are activated in a QRAM call is $O(\text{poly log } N)$. For each memory call, the amount of required energy and created decoherence thus scales logarithmically with the memory size. Note that because of their peculiar structure, $(N \times N)$-Hankel matrices require only $O(N)$ elements to be stored. In comparison, a general s-sparse matrix requires storage of $O(Ns)$ elements.

(ii) As a second approach, we have been given multiple copies of particular quantum state vectors encoding the data in their amplitudes. This approach does not require quantum RAM and operates using the quantum principal component algorithm (QPCA). Importantly, our method then compares to the QFT in the sense that it operates on a given initial state that contains the data to be transformed.
The states that are processed by QPCA correspond to positive semidefinite matrices, which is in general not the case for the Hankel matrices $F^{(i)}$. Adding a sufficiently scaled unit matrix would enforce positivity, but the resulting matrix would not have the required low rank anymore. It turns out, however, that by employing a new type of extended matrix, we can use QPCA to compute singular value decompositions of indefinite matrices and make it applicable for our algorithm, as is fleshed out in appendix B. The given state vectors have to be of a particular form such as

$$|\chi^{(i)}\rangle = \frac{1}{\sqrt{C^{(i)}}} \sum_{j,k=1}^{N/2} |j\rangle|k\rangle (F^{(i)}_{j,k}|0\rangle + a^{(i)}(F^{(i)\dagger}F^{(i)})_{j,k}|1\rangle), \tag{20}$$

with $C^{(i)} = (\|F^{(i)}\|_2^2 + a^{(i)\,2}\|F^{(i)\dagger}F^{(i)}\|_2^2)$ and a known scaling constant $a^{(i)}$ such that $(a^{(i)})^{-1} = O(\max_{j,k}|(F^{(i)\dagger}F^{(i)})_{j,k}|)$, where $\|F^{(i)}\|_2$ is the Frobenius norm of $F^{(i)}$. This state includes in its amplitudes information about the Hankel matrix $F^{(i)}$ and $F^{(i)\dagger}F^{(i)}$. The particular form of $|\chi^{(i)}\rangle$ will become clear in the next section. The advantages of the matrix pencil algorithm over the usual Fourier transform come at a price in the quantum algorithm: we require availability of the state vectors $|\chi^{(i)}\rangle$ instead of the signal state vector $\sum_j f_j|j\rangle$.

In the next section, we show how the operation in equation (18) or, alternatively, multiple copies of $|\chi^{(i)}\rangle$ can be used to efficiently simulate a Hermitian matrix that encodes the eigenvalues and associated eigenvectors of the Hankel matrices.

3.2. Simulating the Hankel matrices

We would like to obtain the singular values and vectors of $F^{(1)}$ and $F^{(2)}$ with a quantum speedup via phase estimation, which for real signals correspond, up to signs, to their eigenvalues and vectors. Since the procedure is the same for $F^{(1)}$ and $F^{(2)}$, for clarity we will drop the index in this section and use F for both matrices. Phase estimation requires the repeated application of powers of a unitary operator generated by a Hermitian matrix to find the eigenvalues and eigenvectors of that matrix. Thus, we need to connect both Hankel matrices, generally non-Hermitian, to Hermitian matrices. Depending on the input source discussed in the previous section, this is done in different ways.

Generally, since F is not sparse, we cannot make use of the sparse simulation techniques described in [39]. Although both matrices have low rank $p \ll N$, they will in general not be positive definite, so that QPCA [20] cannot readily be used either. Note that although $F^\dagger F$ and $FF^\dagger$ are positive definite, provide the correct singular vectors of F, and can be efficiently exponentiated, the phase relations between left and right singular vectors, which are necessary for the matrix pencil algorithm, are not preserved. This insight can be taken as yet another

motivation to look for more general efficient methods to exponentiate matrices that exhibit a suitable structure, such as being low-rank, sparse or having a low tensor rank.

For the oracular setting (i), we construct a Hermitian matrix $\widetilde{F}$ and apply the unitary operator $\mathrm{e}^{-\mathrm{i}\,\widetilde{F}t}$ to an initial quantum state. Hereto, we employ the 'extended matrix'

$$\widetilde{F} \coloneqq \begin{bmatrix} 0 & F \\ F^\dagger & 0 \end{bmatrix} \in \mathbb{C}^{N\times N}, \tag{21}$$

which is Hermitian by construction. Its eigenvalues correspond to the singular values $\pm s_j$, $j = 1,\ldots,N/2$, of F and its eigenvectors are proportional to $(u_j,\ \pm v_j) \in \mathbb{C}^N$. Importantly, the phase relations between left and right singular vectors are preserved. Note that an operation analogous to equation (18) for the extended matrix can be easily constructed from equation (18). The method developed in [24] allows us to exponentiate non-sparse Hermitian matrices in this oracular setting. Following their discussion, equation (19) is mapped to the corresponding entry of a modified swap matrix $S_{\widetilde{F}}$, resulting in the matrix

$$S_{\widetilde{F}} \coloneqq \sum_{j,k=1}^{N} \widetilde{F}_{j,k}\ |k\rangle\langle j| \otimes |j\rangle\langle k| \ \in \mathbb{C}^{N^2\times N^2}. \tag{22}$$

In [24], it is shown that performing infinitesimal swap operations with $S_{\widetilde{F}}$ on an initial state $\rho \otimes \sigma$ with auxiliary state $\rho \coloneqq (1/N)_{j,k=1}^{N}$ is equivalent to just evolving σ in time with the Hamiltonian $\widetilde{F}$ for small $\Delta t > 0$, i.e.

$$\mathrm{tr}_1(\mathrm{e}^{-\mathrm{i}\,S_{\widetilde{F}}\Delta t}\ \rho\otimes\sigma\ \mathrm{e}^{\mathrm{i}\,S_{\widetilde{F}}\Delta t}) \approx \mathrm{e}^{-\mathrm{i}\,\widetilde{F}\ \Delta t/N}\ \sigma\ \mathrm{e}^{\mathrm{i}\,\widetilde{F}\ \Delta t/N}. \tag{23}$$

The modified swap matrix $S_{\widetilde{F}}$ is one-sparse within a quadratically larger space and can be efficiently exponentiated with the methods in [39–41] with a constant number of oracle calls and run time $\widetilde{O}(\log N)$, where we omit polylogarithmic factors in O by use of the symbol $\widetilde{O}$. Achieving an accuracy $\varepsilon > 0$ for the eigenvalues requires

$$O\left(\frac{\|\widetilde{F}\|_{\max}^2}{\varepsilon^2}\right) \tag{24}$$

steps in the algorithm [24], where $\|\widetilde{F}\|_{\max}$ denotes the maximal absolute element of $\widetilde{F}$. The phase estimation is performed as discussed in [42] to obtain the $1/\varepsilon^2$ scaling compared to the $1/\varepsilon^3$ scaling of the original work [20, 24]. Note that in our setting $|\widetilde{F}_{j,k}| = \Theta(1)$ and in particular $\|\widetilde{F}\|_{\max} = \Theta(1)$. The run time is the number of steps multiplied by the run time of the swap matrix simulation, i.e. $\widetilde{O}(\log N/\varepsilon^2)$. In appendix A, we discuss an alternative approach for efficient matrix exponentiation developed in [38], and check its applicability to our algorithm.

In setting (ii), where we are given multiple copies of state vectors, we proceed in a different way employing QPCA. The state vector $|\chi\rangle$ can be reduced to a particular quantum density matrix as

$$|\chi\rangle\langle\chi| \longmapsto \frac{1}{C}\begin{bmatrix} FF^\dagger & a\ F(F^\dagger F) \\ a\ (F^\dagger F)F^\dagger & a^2\ (F^\dagger F)(F^\dagger F) \end{bmatrix} =: G. \tag{25}$$

With quantities $C = (\|F\|_2^2 + a^2\|F^\dagger F\|_2^2)$ and $a^{-1} = O(\max_{j,k}|(F^\dagger F)_{j,k}|)$ as before. In the same way,

$$\frac{1}{C}\begin{bmatrix} a^2\ (F^\dagger F)(F^\dagger F) & a\ F(F^\dagger F) \\ a\ (F^\dagger F)F^\dagger & FF^\dagger \end{bmatrix} =: \widetilde{G} \tag{26}$$

can be prepared from a permuted state vector $|\widetilde{\chi}\rangle$. The matrix

$$Z \coloneqq (G + \widetilde{G})/2 \tag{27}$$

is positive semi-definite with unit trace by construction, just as required by the QPCA. Invoking the singular value decomposition of $F = USV^\dagger$, its eigenvalues in terms of the singular values of F are given by $s_j^2(as_j \pm 1)^2/(2C)$, its eigenvectors are $(u_j,\ \pm v_j) \in \mathbb{C}^N$. The matrix Z has twice the rank of F. The application of QPCA then allows resolving its eigenvalues to an accuracy $\varepsilon > 0$ using

$$O\left(\frac{1}{\varepsilon^2}\right) \tag{28}$$

copies of $|\chi\rangle$ and $|\widetilde{\chi}\rangle$ [20] for a total run time of again $\widetilde{O}(\log N/\varepsilon^2)$. In appendix B, we provide further details on this method.

Both the oracular and the QPCA setting can be employed in quantum phase estimation to obtain the singular values and associated singular vectors of the Hankel matrices in quantum form. Phase estimation allows the preparation of

$$\sum_{k=1}^{2p} \beta_k |s_k\rangle\,|u_k,\ v_k\rangle, \tag{29}$$

where $F = USV^\dagger$ is the singular value decomposition with right and left singular vectors u_k and v_k. The associated singular value s_k is encoded in a register. The β_k arise from the choice of the initial state. The next section describes concretely how consecutive phase estimation steps are used for the matrix pencil algorithm as a building block to obtain the signal poles and expansion coefficients.

3.3. Twofold phase estimation

In this section, we describe how to obtain the singular vector overlaps $\{\mathcal{U}_{j,k}\}$ and $\{\mathcal{V}_{j,k}\}$. Hereto, we perform two concatenated phase estimation procedures to obtain states that encode these overlaps in their amplitudes, which are essentially determined by tomography. It is important to pay attention to the correct phase relations between the overlaps. Phase estimation is applied to a specific initial state and an additional eigenvalue register. Initial states with large overlap with the eigenstates of $\widetilde{F}$, equation (21), or Z, equation (27), respectively, can be prepared efficiently. For example, $FF^\dagger/\mathrm{tr}(FF^\dagger)|0\rangle\langle 0|$ or $F^\dagger F/\mathrm{tr}(F^\dagger F)|1\rangle\langle 1|$ are suitable initial states and can be prepared from the oracle equation (18) [20]. For both initial states, the trace with an eigenvector $|u_k,\ v_k\rangle$ is $\sigma_k^2/(2\sum_j \sigma_j^2)$. Alternatively, if we have been given multiple copies of $|\chi\rangle$, we can simply take Z to be the initial state [20].

We append two registers for storing the singular values to the initial state, obtaining $|0\rangle\,|0\rangle\,|\psi_0\rangle$ with the notation $|0\rangle := |0,\ldots,0\rangle$, and perform the phase estimation algorithm with $\mathrm{e}^{-\mathrm{i}\,S_{\widetilde{F}^{(2)}}\,\Delta t}$ as a unitary operator to obtain a state proportional to

$$\sum_{k=1}^{2p} \langle u_k^{(2)},\ v_k^{(2)}|\psi_0\rangle\,|0\rangle\,|s_k^{(2)}\rangle\,|u_k^{(2)},\ v_k^{(2)}\rangle, \tag{30}$$

where for clarity we order the eigenspaces such that positive singular values are put first, i.e. $s_{k+p}^{(2)} = -s_k^{(2)}$, $u_{k+p}^{(2)} = u_k^{(2)}$, and $v_{k+p}^{(2)} = -v_k^{(2)}$ for $k = 1,\ldots,p$. To obtain the overlaps of the matrices $U^{(1)}$ and $U^{(2)}$, the v-part of the eigenvector of $\widetilde{F}^{(2)}$ is projected out, yielding

$$|\psi_1\rangle = \frac{1}{\nu_1}\sum_{k=1}^{2p} \langle u_k^{(2)},\ v_k^{(2)}|\psi_0\rangle\,|0\rangle\,|s_k^{(2)}\rangle\,|u_k^{(2)},\ 0\rangle =: \sum_{j=1}^{2p} g_k|0\rangle\,|s_k^{(2)}\rangle\,|u_k^{(2)},\ 0\rangle \tag{31}$$

with normalization factor $\nu_1 \in \mathbb{R}_+$ and $\sum_{k=1}^{2p}|g_k|^2 = 1$. Each singular value $s_k^{(2)} \in \mathbb{R}_+$ can be determined efficiently from this with accuracy ε_σ in a runtime of $\tilde{O}(\log N/\varepsilon_\sigma^3)$ (see section 3.2). We need to determine the amplitudes $\{g_k\}$, which have to be removed from the overlap values. For this, we essentially perform standard tomography of the quantum state equation (31). The singular register vectors $\{|s_k^{(2)}\rangle\}_{k=1}^{2p}$ are pairwise orthogonal, so that the amplitudes $\{g_k\}_{k=1}^{p}$ can be efficiently obtained—up to a global complex phase $\mathrm{e}^{\mathrm{i}\,\vartheta_1}$—via measurements e.g. of the form

$$|s_{k_1}^{(2)}\rangle\langle s_{k_1}^{(2)}|,\ \ |s_{k_2}^{(2)}\rangle\langle s_{k_2}^{(2)}|,\ \ (|s_{k_1}^{(2)}\rangle + |s_{k_2}^{(2)}\rangle)(\langle s_{k_1}^{(2)}| + \langle s_{k_2}^{(2)}|),\ \ (|s_{k_1}^{(2)}\rangle - \mathrm{i}|s_{k_2}^{(2)}\rangle)(\langle s_{k_1}^{(2)}| + \mathrm{i}\langle s_{k_2}^{(2)}|), \tag{32}$$

with probabilities

$$|g_{k_1}|^2,\ \ |g_{k_2}|^2,\ \ |g_{k_1}|^2 + |g_{k_2}|^2 + 2\,\mathfrak{Re}(g_{k_1}g_{k_2}^*),\ \ |g_{k_1}|^2 + |g_{k_2}|^2 + 2\,\mathfrak{Im}(g_{k_1}g_{k_2}^*), \tag{33}$$

respectively. Suppose g_{k_1} is known. Then g_{k_2} can easily be obtained from equation (33). Hence, by fixing one global phase $\mathrm{e}^{\mathrm{i}\,\vartheta_1}$ (e.g. corresponding to $g_1 \overset{!}{=} +|g_1|$), all values $\{g_k\}_{k=1}^{2p}$ are unambiguously determined. Requiring an accuracy

$$\varepsilon_g = \mathbb{V}(g)^{1/2}/\mathbb{E}(g) \tag{34}$$

of the probabilities in equation (33) for $k = 1,\ldots,p$, denoting expected value and variance with $\mathbb{E}$ and $\mathbb{V}$, respectively and with ξ_g the reciprocal of the smallest probability, we require $O(\xi_g/\varepsilon_g^2)$ measurement repetitions for each amplitude. We thus have established the values

$$g_k\mathrm{e}^{\mathrm{i}\,\vartheta_1} = \langle u_k^{(2)},\ v_k^{(2)}|\psi_0\rangle\,\frac{\mathrm{e}^{\mathrm{i}\,\vartheta_1}}{\nu_1},\ \ k = 1,\ldots,2p. \tag{35}$$

Next, the state vector $|\psi_1\rangle$ is used as input for a second phase estimation procedure with $\mathrm{e}^{-\mathrm{i}\,S_{\widetilde{F}^{(1)}}\,\Delta t}$ as unitary operator, yielding

$$|\psi_2\rangle = \frac{1}{\nu_2}\sum_{j,k=1}^{2p}\langle u_k^{(2)}, v_k^{(2)}|\psi_0\rangle\,\langle u_j^{(1)}, v_j^{(1)}|u_k^{(2)}, 0\rangle\,|s_j^{(1)}\rangle\,|s_k^{(2)}\rangle\,|u_j^{(1)}, v_j^{(1)}\rangle$$
$$=: \sum_{j,k=1}^{2p} h_{j,k}|s_j^{(1)}\rangle\,|s_k^{(2)}\rangle\,|u_j^{(1)}, v_j^{(1)}\rangle \tag{36}$$

with normalization factor $\nu_2 \in \mathbb{R}_+$ and $\sum_{j,k=1}^{2p}|h_{j,k}|^2 = 1$. The inner product $\langle u_j^{(1)}, v_j^{(1)}|u_k^{(2)}, 0\rangle$ reduces to $\langle u_j^{(1)}|u_k^{(2)}\rangle$ with vectors in $\mathbb{C}^N$. The same way as above, we determine the singular values $\{s_j^{(1)}\}$ and the values

$$h_{j,k}\mathrm{e}^{\mathrm{i}\,\vartheta_2} = \langle u_k^{(2)}, v_k^{(2)}|\psi_0\rangle\,\langle u_j^{(1)}|u_k^{(2)}\rangle\,\frac{\mathrm{e}^{\mathrm{i}\,\vartheta_2}}{\nu_2},\quad j,\,k = 1,\ldots,2p, \tag{37}$$

up to ε_h with global phase $\mathrm{e}^{\mathrm{i}\,\vartheta_2}$ with $O(\xi_h/\varepsilon_h^2)$ repetitions for each amplitude. Dividing the values in equation (37) by the ones in equation (35), we obtain

$$\mathcal{U}_{j,k}\,\nu_{\mathcal{U}}\,\mathrm{e}^{\mathrm{i}\,\vartheta_{\mathcal{U}}} = \langle u_j^{(1)}|u_k^{(2)}\rangle\,\nu_{\mathcal{U}}\,\mathrm{e}^{\mathrm{i}\,\vartheta_{\mathcal{U}}},\quad j,\,k = 1,\ldots,2p, \tag{38}$$

with $\vartheta_{\mathcal{U}} := \vartheta_2 - \vartheta_1$, $\nu_{\mathcal{U}} := \nu_1/\nu_2$ and accuracy $\sim\varepsilon_g + \varepsilon_h$. The established overlaps

$$\langle u_j^{(1)}|u_k^{(2)}\rangle,\quad \langle u_{j+p}^{(1)}|u_k^{(2)}\rangle,\quad \langle u_j^{(1)}|u_{k+p}^{(2)}\rangle,\quad \langle u_{j+p}^{(1)}|u_{k+p}^{(2)}\rangle \tag{39}$$

correspond to the same matrix entry of $\mathcal{U}$ for $j,\,k = 1,\ldots,p$ and can be averaged over. This way, the matrix $\mathcal{U}$ is determined up to a global phase and a normalization factor. Repeating the entire procedure, but with projecting out the u-part,

$$|u_k^{(2)}, v_k^{(2)}\rangle \mapsto |0, v_k^{(2)}\rangle,\qquad k = 1,\ldots,2p, \tag{40}$$

yields all overlaps $\{\langle v_j^{(1)}|v_k^{(2)}\rangle\}_{j,\,k=1}^{p}$, the entries of $\mathcal{V}$, up to a factor $\nu_{\mathcal{V}}\,\mathrm{e}^{\mathrm{i}\vartheta_{\mathcal{V}}}$. Note that

$$\langle v_j^{(1)}|v_k^{(2)}\rangle = -\langle v_{j+p}^{(1)}|v_k^{(2)}\rangle = -\langle v_j^{(1)}|v_{k+p}^{(2)}\rangle = \langle v_{j+p}^{(1)}|v_{k+p}^{(2)}\rangle \tag{41}$$

for $j,\,k = 1,\ldots,p$ because the v-parts of the $\widetilde{F}^{(i)}$ eigenvectors from $k = 1,\ldots,p$ and $k = p+1,\ldots,2p$ have opposite signs. For real-valued signals and Hermitian $F^{(i)}$, we can perform the procedure with $\mathrm{e}^{-\mathrm{i}\,S_{F^{(i)}}\,\Delta t}$ instead of $\mathrm{e}^{-\mathrm{i}\,S_{\widetilde{F}^{(i)}}\,\Delta t}$ and do not need to project the u- and v-parts.

In summary, we have determined the singular values forming matrix $S^{(i)}$ to accuracy ε_σ in time $\widetilde{O}(p/\varepsilon_\sigma^2)$. In addition, we have determined the overlaps of the right and left singular vectors of the two Hankel matrices $F^{(1)}$ and $F^{(2)}$. The required number of repetitions is

$$n_{\mathcal{U}} = O\left(\frac{p}{\varepsilon_g^2}\xi_g + \frac{p^2}{\varepsilon_h^2}\xi_h\right) \tag{42}$$

for obtaining the entries of $\mathcal{U}$ and analogously $n_{\mathcal{V}}$ for obtaining the entries of $\mathcal{V}$. With

$$n_\phi = \widetilde{O}\left(\frac{\log N}{\varepsilon^2}\right) \tag{43}$$

for the cost of the phase estimation, this leads to a total run time of

$$n := n_\phi\,(n_{\mathcal{U}} + n_{\mathcal{V}}) = \widetilde{O}\left(\frac{p^2\,\xi}{\varepsilon^4}\log N\right), \tag{44}$$

with $\xi := \max\{\xi_g, \xi_h\}$. The performance scales as $n = O(\mathrm{poly}\,\log N)$ for example in the following regime: first, the number of poles is small compared to N, which is a natural regime, as mentioned above; second, regarding ξ, if the overlaps are not too small, $\xi = O(\mathrm{poly}\,\log N)$; and third, an error $1/\varepsilon = O(\mathrm{poly}\,\log N)$ can be tolerated.

3.4. Solving the small classical problem

Having determined the values via phase estimation, the reconstructed eigenvalue equation (17) now reads

$$\hat{\mathcal{F}}\,w := \nu_{\mathcal{U}}\nu_{\mathcal{V}}\mathrm{e}^{\mathrm{i}(\vartheta_{\mathcal{U}}+\vartheta_{\mathcal{V}})}\,(S^{(1)})^{-1}\mathcal{U}\,S^{(2)}\,\mathcal{V}\,w = \gamma\,w. \tag{45}$$

All (scaled) matrix entries of equation (45) are available classically and we can solve the problem with a classical algorithm [33] running with time $O(p^3)$. The errors in the matrix entries are amplified within the entries of the matrix product entries $\hat{\mathcal{F}}_{j,k}$ by a factor of poly p at worst. Taking the inverse of $S^{(1)}$ amounts to inverting its diagonal entries, hence the relative errors of $(S^{(1)})_{j,j}^{-1}$ are unchanged. These are only small if the effective singular values of $F^{(1)}$ (the ones bigger than a suitable threshold θ_1) are sufficiently bigger than zero, resulting in a condition number of $S^{(1)}$ bounded by $\max_j(S_{j,j}^{(1)})/\theta_1$. $\mathcal{F}$ as well as the perturbed matrix $\hat{\mathcal{F}} = \mathcal{F} + \Delta\mathcal{F}$ will in general not be normal, but diagonalizable: $\mathcal{F} = X\,\mathrm{diag}(\lambda_j)\,X^{-1}$. According to the Bauer–Fike theorem [43], we can order the eigenvalues $\{\hat{\lambda}_j\}$ of $\hat{\mathcal{F}}$ such that

$$|\lambda_j - \hat{\lambda}_j| \leqslant \kappa(X)\|\mathcal{F} - \hat{\mathcal{F}}\|_2 \tag{46}$$

for $j = 1, \ldots, p$, where $\kappa(X) := \|X\|_2 \|X^{-1}\|_2$ is the condition number of X, which represents the amplification factor of the matrix perturbation towards the perturbation of the eigenvalues. The matrix perturbation contributes linearly, while the condition number of X, which is independent of the perturbation $\Delta\mathcal{F}$, is related to the condition of the underlying inverse spectral estimation problem. This could in principle be ill-conditioned (e.g. for the reconstruction of extremely small or highly damped spectral components relative to the other ones), but we are more concerned with problems that are also of interest in the classical world and hence sufficiently well-conditioned. Note that p, the number of poles, is small by assumption so that this classical step does not pose a computational bottleneck for the algorithm. For noisy signals, the rank of $F^{(i)}$ will in general be larger than p, $F^{(i)}$ could even be full rank—for not too large noise, however, the additional noise components will remain small such that the effective rank will still be at p. Since only the biggest components of $F^{(i)}$ are taken into account, this results in a rank-p approximation that is best in the Frobenius norm sense (Eckart–Young theorem [44]) and an effective noise filtering of the underlying signal.

The eigenvalues γ_k of equation (45) are determined up to $\mathrm{e}^{-\mathrm{i}(\varphi_\mathcal{U}+\varphi_\mathcal{V})-\log(\nu_\mathcal{U}\nu_\mathcal{V})}$, which corresponds to a uniform translation of all poles. We can take care of this ambiguity by introducing an additional reference pole $\lambda_{\mathrm{ref}} := 0$ (corresponding to the eigenvalue $\mu_{\mathrm{ref}} = 1$) that has to be incorporated into the original signal. This can easily be achieved by adding any constant to the original signal vector (its normalizability is not affected). Since for exponentially damped signals $\mathfrak{Re}(\lambda_k) \leqslant 0$ holds for each k, the eigenvalue γ_{ref} corresponding to the reference pole will still be identifiable as the one with the biggest absolute value $|\gamma_k|$. Simply dividing all γ_k by γ_{ref} (corresponding to the transformation $\lambda_k \Delta t \mapsto \lambda_k \Delta t + \mathrm{i}(\varphi_\mathcal{U} + \varphi_\mathcal{V}) + \log(\nu_\mathcal{U}\nu_\mathcal{V})$ for each k) then yields the correct values $\{\mu_k\}$ and poles.

3.5. Quantum linear fitting

We feed the poles back into the quantum world by using the quantum fitting algorithm described in [25, 26] to obtain the coefficients $\{c_k\}$ in $O(\log(N)p)$ steps and hence the entire parametrization of the input function. We consider real and imaginary parts of the signal f, the poles $\lambda_k \Delta t =: -\alpha_k + \mathrm{i}\,\beta_k$ and the coefficients $c_k = a_k + \mathrm{i}\,b_k$ separately, and equation (14) becomes

$$\widetilde{W}\,\tilde{c} = \tilde{f} \tag{47}$$

with

$$\widetilde{W} := \begin{bmatrix} \mathrm{e}^{-\alpha_1 \cdot 0}\cos(\beta_1 \cdot 0) & \ldots & \mathrm{e}^{-\alpha_p \cdot 0}\cos(\beta_p \cdot 0) & -\mathrm{e}^{-\alpha_1 \cdot 0}\sin(\beta_1 \cdot 0) & \ldots & -\mathrm{e}^{-\alpha_p \cdot 0}\sin(\beta_p \cdot 0) \\ \vdots & & \vdots & \vdots & & \vdots \\ \mathrm{e}^{-\alpha_1 \cdot \widetilde{N}}\cos(\beta_1 \cdot \widetilde{N}) & \ldots & \mathrm{e}^{-\alpha_p \cdot \widetilde{N}}\cos(\beta_p \cdot \widetilde{N}) & -\mathrm{e}^{-\alpha_1 \cdot \widetilde{N}}\sin(\beta_1 \cdot \widetilde{N}) & \ldots & -\mathrm{e}^{-\alpha_p \cdot \widetilde{N}}\sin(\beta_p \cdot \widetilde{N}) \\ \mathrm{e}^{-\alpha_1 \cdot 0}\sin(\beta_1 \cdot 0) & \ldots & \mathrm{e}^{-\alpha_p \cdot 0}\sin(\beta_p \cdot 0) & \mathrm{e}^{-\alpha_1 \cdot 0}\cos(\beta_1 \cdot 0) & \ldots & \mathrm{e}^{-\alpha_p \cdot 0}\cos(\beta_p \cdot 0) \\ \vdots & & \vdots & \vdots & & \vdots \\ \mathrm{e}^{-\alpha_1 \cdot \widetilde{N}}\sin(\beta_1 \cdot \widetilde{N}) & \ldots & \mathrm{e}^{-\alpha_p \cdot \widetilde{N}}\sin(\beta_p \cdot \widetilde{N}) & \mathrm{e}^{-\alpha_1 \cdot \widetilde{N}}\cos(\beta_1 \cdot \widetilde{N}) & \ldots & \mathrm{e}^{-\alpha_p \cdot \widetilde{N}}\cos(\beta_p \cdot \widetilde{N}) \end{bmatrix},$$

$$\widetilde{W} := (w_{j,k}) = \begin{bmatrix} (\mathfrak{Re}\,\mu_k^{j-1}) & (-\mathfrak{Im}\,\mu_k^{j-1}) \\ (\mathfrak{Im}\,\mu_k^{j-1}) & (\mathfrak{Re}\,\mu_k^{j-1}) \end{bmatrix} \in \mathbb{R}^{2N \times 2p}, \quad \tilde{c} := \begin{bmatrix} \mathfrak{Re}\,c_1 \\ \vdots \\ \mathfrak{Re}\,c_p \\ \mathfrak{Im}\,c_1 \\ \vdots \\ \mathfrak{Im}\,c_p \end{bmatrix} \in \mathbb{R}^{2p}, \quad \tilde{f} := \begin{bmatrix} \mathfrak{Re}\,f_0 \\ \vdots \\ \mathfrak{Re}\,f_{\widetilde{N}} \\ \mathfrak{Im}\,f_0 \\ \vdots \\ \mathfrak{Im}\,f_{\widetilde{N}} \end{bmatrix} \in \mathbb{R}^{2N},$$

and $\widetilde{N} := N - 1$. The vector 2-norm of the kth column of $\widetilde{W}$ can be established in closed form as

$$\frac{1 - \mathrm{e}^{-2\alpha_k N}}{1 - \mathrm{e}^{-2\alpha_k}}, \quad \text{if } \alpha_k > 0, \qquad \text{and} \quad N, \quad \text{if } \alpha_k = 0. \tag{48}$$

Hence, $\|\widetilde{W}\|_2$ can be computed in time $O(p)$. We will rescale the solution for c such that we can assume that $\|\widetilde{W}\|_2 = 1$. The norms of matrices $\|\widetilde{W}\|_2$ for real-valued signals can be calculated as well by combining the norms of the kth with the $(k + p)$th column. Since each row consists of $2p$ elements, the row norms can be computed in $O(p)$ as well.

Since $\alpha := (\alpha_k)$, $\beta := (\beta_k)$ are known, we can construct a quantum oracle, providing quantum access to the matrix entries $w_{j,k}(\alpha, \beta)$,

$$|\alpha\rangle\,|\beta\rangle\,|j\rangle\,|k\rangle\,|0\rangle \longmapsto |\alpha\rangle\,|\beta\rangle\,|j\rangle\,|k\rangle\,|w_{j,k}(\alpha, \beta)\rangle. \tag{49}$$

The matrix $\widetilde{W}$ can be prepared as a state vector

$$|w\rangle = \sum_{j=1}^{2N}\sum_{k=1}^{2p} w_{j,k}|j\rangle|k\rangle \tag{50}$$

following the procedure described in [26] with time $\widetilde{O}(\text{poly log}(N)\, p\, \xi_W \log(1/\zeta))$, where ζ is the accuracy of the preparation of $|w\rangle$ and

$$\xi_W := \max \|w_j\|_2 / \min \|w_j\|_2 . \tag{51}$$

Here, we set $\widetilde{O}(g(N)) := O(g(N)\text{poly log}(g(N)))$ for functions g. For the preparation of $|\tilde{f}\rangle$, we require time $\widetilde{O}(\text{poly log}(N)\, \xi_f \log(1/\zeta))$ with

$$\xi_f := \max|\tilde{f}_j| / \min|\tilde{f}_j|. \tag{52}$$

With $|w\rangle$ and $|f\rangle$ prepared, we then can proceed as described in [26, theorems 2 and 3] and obtain with probability bigger than 2/3 an estimate $\hat{c}$ in time $\widetilde{O}(\text{poly log}(N)\kappa_W p^{3/2}(\sqrt{2p}\,\xi_f/\varepsilon + \kappa_W^2 \xi_f/\Phi + \kappa_W^6 (2p)^5 \xi_W/\varepsilon^4\Phi)/\varepsilon\Phi)$, with 2-norm accuracy ε, $\kappa_W = \|\widetilde{W}\|_2 / \|\widetilde{W}^+\|_2$, and norm Φ of the projection of $\tilde{f}$ onto the column space of $\widetilde{W}$, the fit quality. Importantly, we can estimate the quality of the fit with time $\widetilde{O}(\text{poly log}(N)(\xi_f + \xi_W (2p)^3 \kappa_W^4/\varepsilon)/\varepsilon)$. Note that sampling $\hat{c}$ is efficient because it comprises $O(p)$ components. Altogether, we have determined the sought-after coefficients and hence all parameters that characterize the signal f in poly log N. This concludes the description of the quantum matrix pencil algorithm.

4. Summary and discussion

We have developed a quantum implementation of an important algorithm for spectral estimation, the MPM, taking a tool from signal processing to the quantum world and significantly improving upon the effort required. Given the arguable scarcity of quantum algorithms with this feature, progress in this respect seems highly desirable. The quantum MPM is a useful alternative to QFT in many practical applications such as imaging or simulation of atomic systems, in the same way that classical MPMs and related algorithms are useful alternatives to the classical Fourier transform. This is especially the case for signals with close damped poles and limited total sampling time. The presented algorithm can be applied to classical data to solve the classical problem at hand.

For a signal given by N equidistant samples, we have made use of the fact that the eigenvalue problem equation (17) consisted of large matrices of size $O(N)$ that could, however, be contracted into manageable matrices of size $O(p)$ via concatenated use quantum phase estimations in $O(\text{poly log}\, N)$. This justifies the use of a quantum version of the MPM as opposed to quantum versions of related algorithms like Prony's method, where the p quantities leading the corresponding poles are determined in a later step, during the fitting of the coefficients, and the critical step would already be $O(\text{poly}\, N)$.

The quantum phase estimation was shown to be implementable in two complementary ways: either by retrieving the input signal via quantum oracle calls such as quantum RAM, or by using multiple copies of a state with the signal encoded in its amplitudes for QPCA. The developed extended matrix construction for indefinite matrices significantly expands the set of matrices that can be exponentiated via QPCA. Since QPCA so far solely relied on positive semidefinite matrices, we expect this to be a useful new primitive also for other quantum algorithms.

The actual step to determine the poles from an eigenvalue problem of a $p \times p$ matrix can be performed classically since p is assumed to be small. Subsequently, feeding back the established poles into a quantum fitting algorithm allows the coefficients of the signal again to be determined efficiently in $\widetilde{O}(\text{poly log}\, N)$. This way, we have an effective division of labor between classical and quantum algorithms, to the extent that such a hybrid algorithm is possible efficiently. Classical intermediate steps are for example reminiscent of quantum error correction, where error syndromes are measured and the quantum state is processed according to the classical measurement results [45].

In order to create an efficient quantum algorithm, it is essential to adress certain caveats, which are succinctly listed in Aaronson [46] using the example of the groundbreaking work in [47]: both for the QRAM and the QPCA setting, the input data can be accessed quickly enough and the Hankel matrices can be exponentiated efficiently—due to being sparse in a quadratically larger space or by fulfilling the QPCA requirements, respectively. For this, it is necessary that the entries of the Hankel matrices and hence the input signal have a similar magnitude $\Theta(1)$. Furthermore, for twofold phase estimation, as for general phase estimation, we need to be able to prepare initial states that provide sufficiently large overlap with the states we use for further processing. In the QRAM setting as well as in the QPCA setting, one can employ initial states that are closely related to the input signal. Analogously, the overlaps in the matrices $\mathcal{U}$ and $\mathcal{V}$ need to be sufficiently large. Reading-out the

$O(N)$ components of the state vectors would foil the achieved quantum speedup; however, as in [36, 37], the number of necessary output quantities in our algorithm is condensed down to $O(p)$. Each output can be determined with time $\widetilde{O}(\text{poly}\log(N))$, provided that Vandermonde matrix $\widetilde{W}$ from the established frequencies is sufficiently well-conditioned, analogous to the requirements related to the condition number in the matrix inversion algorithm [47]. Naturally, we are interested in sufficiently well-behaved signals where a classical MPM algorithm could in principle reconstruct all of its components, excluding e.g. highly damped or relatively small terms, which manifest themselves again in the conditioning of the matrix inversion. In this respect, the quantum MPM inherits the properties related to the conditioning of its classical analogue.

The outlined procedure is generalizable to arbitrary signal dimensions d, i.e. signals of the type $f(t_1,\ldots,t_d) = \sum_{k_1,\ldots,k_d=1}^{p} c_{k_1,\ldots,k_d}\, \mathrm{e}^{\lambda_{k_1} t_1+\ldots+\lambda_{k_d} t_d}$, with $c \in \mathbb{C}^{p^d}$ by suitable tensor contractions of the array of signal samples $(f_{j_1,\ldots,j_d})_{(j_l)=0}^{N-1}$ [5] or fixing all time indices but one and applying the MPM on the remaining vector. This yields the sought-after poles since they are the same for the different time indices t_i. For time index-dependent poles, one can consider 'enhanced matrices'—embeddings of Hankel matrices that correspond to one-dimensional projections of the multidimensional signal within a larger block Hankel matrix—as in [48]. There are many potential applications for this, e.g. in radar imaging and geophysics [49].

Beyond the potential use of reducing the computation time of the MPM in its classical applications or classical postprocessing in quantum applications, it is also worthwhile to consider the possibilities in a pure quantum setting: these include the examination of quantum systems that feature a discrete set of damped oscillations such as the vibronic modes of molecules in a condensed-phase environment where the data—as opposed to what is usually done—would also have to be taken in a quantum coherent manner in order to replace quantum RAM or to build a state as in appendix B and subsequently be processed by the quantum MPM.

We expect the methods and primitives that we develop and introduce here to be highly useful also when devising other quantum algorithms. This includes the new ideas on the computation of overlaps by suitably concatenating quantum phase estimation procedures and on the efficient exponentiation of a novel type of structured matrices on a quantum computer. We hope that the present work stimulates such further research.

Acknowledgments

AS thanks the German National Academic Foundation (Studienstiftung des deutschen Volkes) and the Fritz Haber Institute of the Max Planck Society for support. SL and PR were supported by ARO and AFOSR. JE thanks the Templeton Foundation, the DFG (CRC 183, EI 519/7-1), the ERC (TAQ), and the EC (RAQUEL, AQuS) for support.

Appendix A. Alternative non-sparse quantum oracle method

Berry *et al* present a method to exponentiate matrices sublinear in the sparsity [38]. In this section, we summarize the performance and requirements of this method and the application to the low-rank Hankel matrices of the present work. The number of oracle queries for simulating a matrix such as the Hermitian $\widetilde{F}^{(i)}$ in equation (21) is given by

$$O(t^{3/2}\sqrt{s\Lambda_{\text{tot}}/\varepsilon}), \tag{A1}$$

where s is the sparsity and ε is the error. The quantity $\Lambda_{\text{tot}} > 0$ depends on the norms of the matrix as $\Lambda_{\text{tot}} = \Lambda\Lambda_1\Lambda_{\max}$ with the spectral norm $\Lambda = \|\widetilde{F}^{(i)}\|_\infty$, the maximum column sum norm $\Lambda_1 = \|\widetilde{F}^{(i)}\|_1$, and the maximum matrix element $\Lambda_{\max} = \|\widetilde{F}^{(i)}\|_{\max}$. The conditions for this to work are given by $\Lambda t \geqslant \sqrt{\varepsilon}$,

$$t \geqslant \frac{\Lambda}{\Lambda_{\max}\Lambda_1 s}, \tag{A2}$$

and $\Lambda \leqslant \Lambda_1$.

We confirm that under reasonable assumptions the low-rank non-sparse Hankel matrices under consideration in this work can be simulated with $O(\log N)$ queries. Assume that the signal is reasonably small with not too many zeros. This implies that the matrix $\widetilde{F}^{(i)}$ is non-sparse with $s = \Theta(N)$ and the individual elements scale as $\widetilde{F}^{(i)}_{jk} = \Theta(1)$. If we assume that the signal is generated by a few (in fact, p) components, then the matrix is low rank with rank $2p$. Since $\text{tr}((\widetilde{F}^{(i)})^2) = \sum_{j=1}^{2p}\lambda_j^2 \leqslant N^2\|\widetilde{F}^{(i)}\|_{\max}^2$, we have that the significant eigenvalues scale as $\lambda_j = \Theta(N)$, $j = 1, \ldots, 2p$. These assumptions have the following straightforward implications:

(i) The spectral norm (largest eigenvalue) is $\Lambda = \Theta(N)$,

(ii) the induced 1-norm (maximum column sum) is $\Lambda_1 = \Theta(N)$, and

(iii) the maximum element is $\Lambda_{\max} = \Theta(1)$.

Thus, $\Lambda_{\text{tot}} = \Theta(N^2)$ and the total number of queries is $O(t^{3/2}\sqrt{\Theta(N^3)/\varepsilon})$. We need time $t = \Theta(1/N)$ to resolve the eigenvalues $\lambda_j = \Theta(N)$ via phase estimation. Thus, at an error ε, we need $O(1/\sqrt{\varepsilon})$ queries, which is again efficient.

We show that we can satisfy the conditions as follows. Since we have $t = \Theta(1/N)$ already from phase estimation, we can assume that with constant effort $t \geqslant \sqrt{\varepsilon}/\Lambda = \Theta(\sqrt{\varepsilon}/N)$. Next, by using (i)–(iii) and $s = \Theta(N)$, we have

$$t \geqslant \frac{\Lambda}{\Lambda_{\max}\Lambda_1 s} = \Theta\left(\frac{1}{N}\right). \tag{A3}$$

The third criterion $\Lambda \leqslant \Lambda_1$ is satisfied by Gershgorin's theorem, since the eigenvalues are bounded by the maximum sum of the absolute elements in a row/column.

Appendix B. Matrix exponentiation via QPCA

In this appendix, we present an alternative way to efficiently exponentiate indefinite matrices, in order to give more substance to ideas of exponentiating structured matrices while at the same time preserving a phase relationship. Since exponentiating matrices $F \in \mathbb{C}^{N/2\times N/2}$ while a preserving phase relationship is key to the above algorithm and is expected to be important in other quantum algorithms, we briefly present an alternative method that accomplishes this task via QPCA. This method compares to the QFT in the sense that it operates on a given initial state that contains the data to be transformed in its amplitudes without querying QRAM. We assume that we have been presented with many copies of the state vector

$$|\chi\rangle = \frac{1}{\sqrt{C}}\sum_{j,k=1}^{N/2}|j\rangle|k\rangle(F_{j,k}|0\rangle + a(F^\dagger F)_{j,k}|1\rangle), \tag{B.1}$$

with $C := (\|F\|_2^2 + a^2\|F^\dagger F\|_2^2)$ and $a^{-1} := O(\max_{j,k}|(F^\dagger F)_{j,k}|)$. The matrix F takes the role of $F^{(1)}$ and $F^{(2)}$ of the main text, so again the classical index i is suppressed. Note that even though a is exponentially small, the individual amplitudes of this state are of similar size. Reducing the state in terms of the k index leads to

$$\mathrm{tr}_2(|\chi\rangle\langle\chi|) = \frac{1}{C}\left(\sum_{j,j'}|j\rangle\langle j'|\sum_{k=1}^{N/2}(F_{j,k}|0\rangle + a(F^\dagger F)_{j,k}|1\rangle)(F^*_{j',k}\langle 0| + a(F^\dagger F)^*_{j',k}\langle 1|)\right).$$

In matrix form, this reduced density matrix is written as

$$G := \frac{1}{C}\begin{bmatrix} FF^\dagger & a\,F(F^\dagger F) \\ a\,(F^\dagger F)F^\dagger & a^2\,(F^\dagger F)(F^\dagger F)\end{bmatrix}. \tag{B.2}$$

By the use of the singular value decomposition of $F = USV^\dagger$, this matrix—positive semi-definite by construction—can be written as

$$G = \frac{1}{C}\begin{bmatrix}U & 0\\ 0 & V\end{bmatrix}\begin{bmatrix}S^2 & a\,S^3\\ a\,S^3 & a^2S^4\end{bmatrix}\begin{bmatrix}U^\dagger & 0\\ 0 & V^\dagger\end{bmatrix}. \tag{B.3}$$

In precisely the same way, we are given multiple copies of the state

$$|\widetilde{\chi}\rangle = \frac{1}{\sqrt{C}}\sum_{j,k=1}^{N/2}|j\rangle|k\rangle(a(FF^\dagger)_{j,k}|0\rangle + F^\dagger_{j,k}|1\rangle). \tag{B.4}$$

Again reducing the state in terms of the k index leads to

$$\mathrm{tr}_2(|\widetilde{\chi}\rangle\langle\widetilde{\chi}|) = \frac{1}{C}\left(\sum_{j,j'}|j\rangle\langle j'|\sum_{k=1}^{N/2}(a(FF^\dagger)_{j,k}|0\rangle + F^\dagger_{j,k}|1\rangle)(a(FF^\dagger)^*_{j',k}\langle 0| + (F^\dagger_{j',k})^*\langle 1|)\right),$$

leading to the matrix

$$\widetilde{G} := \frac{1}{C}\begin{bmatrix}a^2\,(FF^\dagger)(FF^\dagger) & a\,(FF^\dagger)F\\ a\,F^\dagger(FF^\dagger) & F^\dagger F\end{bmatrix}, \tag{B.5}$$

which can be decomposed as

$$\widetilde{G} = \frac{1}{C}\begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix}\begin{bmatrix} a^2S^4 & a\,S^3 \\ a\,S^3 & S^2 \end{bmatrix}\begin{bmatrix} U^\dagger & 0 \\ 0 & V^\dagger \end{bmatrix}. \tag{B.6}$$

The matrix

$$Z := \frac{1}{2}(G + \widetilde{G}) \tag{B.7}$$

has still low rank, as it has just twice the rank of F. Its eigenvectors are $(u_j, \pm v_j) \in \mathbb{C}^N$ and its eigenvalues in terms of the singular values of F are given by $s_j^2(as_j \pm 1)^2/(2C)$ since

$$Z = \frac{1}{2C}\begin{bmatrix} FF^\dagger + a^2\,(FF^\dagger)(FF^\dagger) & 2a\,FF^\dagger F \\ 2a\,F^\dagger FF^\dagger & a^2\,F^\dagger F + (F^\dagger F)(F^\dagger F) \end{bmatrix} \tag{B.8}$$

and

$$\begin{aligned} &\frac{1}{2C}\begin{bmatrix} FF^\dagger + a^2\,(FF^\dagger)(FF^\dagger) & 2a\,FF^\dagger F \\ 2a\,F^\dagger FF^\dagger & a^2\,F^\dagger F + (F^\dagger F)(F^\dagger F) \end{bmatrix}\begin{bmatrix} u_j \\ \pm v_j \end{bmatrix} \\ &= \frac{1}{2C}\begin{bmatrix} (s_j^2 + a^2 s_j^4 \pm 2as_j^3)u_j \\ (2as_j^3 \pm s_j^2 \pm a^2 s_j^4)v_j \end{bmatrix} = \frac{1}{2C}s_j^2(as_j \pm 1)^2\begin{bmatrix} u_j \\ \pm v_j \end{bmatrix}. \end{aligned} \tag{B.9}$$

This renders standard QPCA [20] readily applicable and allows us to determine the singular spectra of matrices F, even if they are indefinite, by constructing the positive semidefinite matrix Z.

References

[1] Karski M, Förster L, Choi J M, Alt W, Widera A and Meschede D 2009 Nearest-neighbor detection of atoms in a 1d optical lattice by fluorescence imaging *Phys. Rev. Lett.* **102** 053001
[2] Naishadham K and Piou J E 2008 A robust state space model for the characterization of extended returns in radar target signatures *IEEE Trans. Antennas Propag.* **56** 1742–51
[3] Viti V, Petrucci C and Barone P 1997 Prony methods in NMR spectroscopy *Int. J. Imaging Syst. Technol.* **8** 565–71
[4] Maravic I, Kusuma J and Vetterli M 2003 Low-sampling rate UWB channel characterization and synchronization *J. Commun. Netw.* **5** 319–27
[5] Steffens A, Riofrío C A, Hübener R and Eisert J 2014 Quantum field tomography *New J. Phys.* **16** 123010
[6] Steffens A, Friesdorf M, Langen T, Rauer B, Schweigler T, Hübener R, Schmiedmayer J, Riofrío C A and Eisert J 2015 Towards experimental quantum field tomography with ultracold atoms *Nat. Commun.* **6** 7663
[7] Leonowicz Z, Lobos T and Rezmer J 2003 Advanced spectrum estimation methods for signal analysis in power electronics *IEEE Trans. Ind. Electron.* **50** 514–9
[8] Andrade X, Sanders J N and Aspuru-Guzik A 2012 Application of compressed sensing to the simulation of atomic systems *Proc. Natl Acad. Sci.* **109** 13928–33
[9] Hua Y and Sarkar T K 1990 Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise *IEEE Trans. Signal Process.* **38** 814–24
[10] de Prony B G R 1795 Essai éxperimental et analytique: sur les lois de la dilatabilitéde fluides élastique et sur celles de la force expansive de la vapeur de l'alkohol, a différentes températures *J. Ec. Poly.* **1** 24–76
[11] Schmidt R O 1986 Multiple emitter location and signal parameter estimation *IEEE Trans. Antennas Propag.* **34** 276–80
[12] Roy R and Kailath T 1989 Esprit-estimation of signal parameters via rotational invariance techniques *IEEE Trans. Signal. Process.* **37** 984–95
[13] Bhaskar B N, Tang G and Recht B 2013 Atomic norm denoising with applications to line spectral estimation *IEEE Trans. Signal. Process.* **61** 5987–99
[14] Park J I and Kim K T 2010 A comparative study on ISAR imaging algorithms for radar target identification *Prog. Electromagn. Res.* **108** 155–75
[15] del Río J E F and Sarkar T K 1996 Comparison between the matrix pencil method and the fourier transform technique for high-resolution spectral estimation *Digit. Signal Process.* **6** 108–25
[16] Baqai F A and Hua Y 1993 Matrix pencil methods for ISAR image reconstruction *1993 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, 1993, ICASSP-93* vol 5 (IEEE) pp 473–6
[17] Nielsen M A and Chuang I L 2010 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)
[18] Cooley J W and Tukey J W 1965 An algorithm for the machine calculation of complex fourier series *Math. Comput.* **19** 297–301
[19] Shor P W 1999 Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer *SIAM Rev.* **41** 303–32
[20] Lloyd S, Mohseni M and Rebentrost P 2014 Quantum principal component analysis *Nat. Phys.* **10** 631–3
[21] Giovannetti V, Lloyd S and Maccone L 2008 Quantum random access memory *Phys. Rev. Lett.* **100** 160501
[22] Giovannetti V, Lloyd S and Maccone L 2008 Architectures for a quantum random access memory *Phys. Rev.* A **78** 052310
[23] De Martini F, Giovannetti V, Lloyd S, Maccone L, Nagali E, Sansoni L and Sciarrino F 2009 Experimental quantum private queries with linear optics *Phys. Rev.* A **80** 010302
[24] Rebentrost P, Steffens A and Lloyd S 2016 Quantum singular value decomposition of non-sparse low-rank matrices arXiv:1607.05404
[25] Wiebe N, Braun D and Lloyd S 2012 Quantum algorithm for data fitting *Phys. Rev. Lett.* **109** 050505
[26] Wang G 2014 Quantum algorithms for curve fitting arXiv:1402.0660
[27] Shannon C E 1949 Communication in the presence of noise *Proc. IRE* **37** 10–21
[28] Golub G H and Van Loan C F 2012 *Matrix Computations* vol 3 (Baltimore, MD: John Hopkins University Press)

[29] Stewart G W and Sun J-G 1990 *Matrix Perturbation Theory* (*Computer Science and Scientific Computing*) (London: Academic)
[30] Hua Y and Sarkar T K 1991 On SVD for estimating generalized eigenvalues of singular matrix pencil in noise *IEEE Int. Symp. Circuits and Systems 1991* 39 (IEEE) pp 892–900
[31] Browne K, Qiao S and Wei Y 2009 A Lanczos bidiagonalization algorithm for hankel matrices *Linear Algebr. Appl.* **430** 1531–43
[32] Xu W and Qiao S 2008 A fast symmetric SVD algorithm for square hankel matrices *Linear Algebr. Appl.* **428** 550–63
[33] Moler C B and Stewart G W 1973 An algorithm for generalized matrix eigenvalue problems *SIAM J. Numer. Anal.* **10** 241–56
[34] Kitaev A Y 1995 Quantum measurements and the Abelian stabilizer problem arXiv:quant-ph/9511026
[35] Cleve R, Ekert A, Macchiavello C and Mosca M 1998 Quantum algorithms revisited *Proc. R. Soc.* A **454** 339–54
[36] Aaronson S 2009 Bqp and the polynomial hierarchy arXiv:0910.4698
[37] Rebentrost P, Mohseni M and Lloyd S 2014 Quantum support vector machine for big data classification *Phys. Rev. Lett.* **113** 130503
[38] Berry D W and Childs A M 2012 Black-box hamiltonian simulation and unitary implementation *Quantum Inf. Comput.* **12** 29–62
[39] Berry D W, Ahokas G, Cleve R and Sanders B C 2007 Efficient quantum algorithms for simulating sparse hamiltonians *Commun. Math. Phys.* **270** 359–71
[40] Childs A M, Cleve R, Deotto E, Farhi E, Gutmann S and Spielman D A 2003 Exponential algorithmic speedup by a quantum walk *Proc. 35th Ann. ACM Sym. Th. Comp.* (ACM) pp 59–68
[41] Aharonov D and Ta-Shma A 2003 Adiabatic quantum state generation and statistical zero knowledge *Proc. 35th Ann. ACM Sym. Th. Comp.* (ACM) pp 20–9
[42] Kimmel S, Lin C, Low G, Ozols M and Yoder T J 2016 Hamiltonian simulation with optimal sample complexity arXiv:1608.00281
[43] Bauer F L and Fike C T 1960 Norms and exclusion theorems *Numer. Math.* **2** 137–41
[44] Eckart C and Young G 1936 The approximation of one matrix by another of lower rank *Psychometrika* **1** 211–8
[45] Gottesman D 2009 An introduction to quantum error correction and fault-tolerant quantum computation arXiv:0904.2557
[46] Aaronson S 2015 Read the fine print *Nat. Phys.* **11** 291–3
[47] Harrow A W, Hassidim A and Lloyd S 2009 Quantum algorithm for linear systems of equations *Phys. Rev. Lett.* **103** 150502
[48] Hua Y 1992 Estimating two-dimensional frequencies by matrix enhancement and matrix pencil *IEEE Trans. Signal Process.* **40** 2267–80
[49] Garello R 2013 *Two-Dimensional Signal Analysis* (New York: Wiley)

5 CONCLUSION AND OUTLOOK

Signal processing overlaps with complex quantum systems in many different places. On the one hand, modern signal processing algorithms are vital for handling the large amounts of data that the description of large quantum systems entails. On the other hand, with the rise of quantum computing in sight, the opportunity emerges to accelerate established signal processing routines with a superpolynomial quantum speedup. In this sense, the goal of this undertaking was to harness signal processing for quantum applications and, vice versa, to use quantum systems for the benefit of signal processing.

Using signal processing techniques like compressed sensing, the limits for fully estimating general mixed quantum states can be substantially pushed further. When dealing with experimental data, which is prone to noise, external parameters have to be introduced, leading to ambiguous results. By making use of model selection techniques, these ambiguities could be lifted, as was argued in section 2.3. Still, the curse of dimensionality can only be mitigated because the amount of data is just reduced by a square root factor, as opposed to the exponential increase of the size of the Hilbert space. This is essentially due to the basically *lossless compression* character of compressed sensing. Nevertheless, compressed sensing is of prime importance in handling intermediate-sized quantum systems.

Making further assumptions about the entanglement and purity of the state, yielding higher compression rates for the underlying models, tensor network methods (section 3.1) can be employed for quantum tomography. Continuous systems, however, in principle possess infinitely many degrees of freedom, posing conceptual challenges about how an efficient tomographic procedure should even look like. With the introduction of continuous matrix product states (CMPS), a promising ansatz class appeared, featuring n-point correlation functions that can be expressed in terms of the parameter matrices that determine the respective state. In the translation invariant case, only two finite-dimensional matrices suffice to completely parametrize the state; nevertheless, CMPS can adequately model various quantum systems in the low energy regime.[125,129] The relationship between the CMPS parameters and the correlation functions is not straightforward, however, and a series of reconstruction steps is required to fully determine the state within a tomographic protocol. Importantly, in the initial step, the correlation functions have to be treated with adapted signal processing methods for accurate spectral estimation—*Prony's method* and *matrix pencil methods*. We furthermore extended these methods to handle signals of arbitrary dimension instead of just one-dimensional (time) intervals. The resulting tomographic protocol allowing for the estimation of quantum field states was presented in section 3.2.

Regarding the experimental realization of quantum field tomography procedures, one-dimensional ultracold Bose gases, which represent systems central for experimentally analyzing thermal equilibration in the quantum regime, are perfect candidates. Using lower-order correlation functions as input, CMPS parameters could be determined to robustly make predictions about higher-

order statistics that coincided with the values directly established from the experiment, resulting in a successful modeling of an ultracold Bose gas quantum system with a CMPS (see section 3.3). With increasing time, the agreement between the predicted and the measured statistics deteriorated, which, beyond signal deterioration due to noise, could be attributed to the entanglement growth after sudden quenches and the entailing required increase in model parameters.

An extension of the CMPS protocol to counting probabilities as input with an application in a quantum transport experiment is discussed in the coauthored publication in appendix A.1, allowing for a closer look on the short-time dynamics of the system. So far, the waiting time distributions for two consecutively transported electrons could only be accessed for emission rates in the kilohertz frequency range. By use of the CMPS protocol, determining waiting time distributions would in principle also be possible in the gigahertz range.

A complementary approach for dealing with the inevitable curse of dimensionality—yet not for state estimation, but for ab initio calculations of the electronic structure of quantum systems—is *density functional theory*, which requires certain sets of basis ansatz functions. Following the approach to be published, presented in appendix A.3, these sets can substantially be reduced with compressed sensing methods, while still retaining the necessary accuracy. This enables one to effectively accelerate computations or tackle larger systems with higher precision.

In contrast, classical signal processing algorithms could massively benefit from using physical effects in complex quantum systems to accelerate computations, i.e. from an effective implementation on a future *quantum computer*. Quantum computers, however, have a fundamentally different architecture compared to classical computers and quantum algorithms have to obey entirely different rule sets in order to efficiently solve problems. This makes it challenging to design a quantum analogue for any classical algorithm.

Spectral estimation algorithms are ubiquitous from nuclear magnetic resonance spectroscopy to image processing. It is therefore highly desirable to explore potential quantum speedups. A *matrix pencil method*, similar to the one employed for CMPS tomography, proved to be the right candidate for this, allowing for a speedup from $O(N^3)$ operations to $O(\mathrm{poly}\log N)$ operations (section 4). At its heart lies a generalized eigenvalue problem, which has to be reformulated such that the required quantities can be obtained efficiently via quantum state tomography. Along the way, novel quantum algorithm techniques like concatenated *phase-estimation* and encoding non-positive semidefinite matrices into density matrices for determining their singular spaces were developed.

An important building block for such quantum algorithms is the efficient simulation of the involved matrices as part of a unitary transformation that acts on quantum states; a method that expands the class of simulatable matrices is discussed in the coauthored publication in appendix A.2. This moreover enables one to perform singular value decompositions of non-sparse low-rank matrices with an exponential quantum speedup.

Still, quantum algorithms are not expected to make all computational problems efficiently tractable. It is to this day unknown whether the complexity class BQP (bounded error quantum polynomial time), the class consisting of all problems that can efficiently be solved on a quantum com-

puter, contains the class of NP-*complete* problems, the class consisting of the computationally hardest problems in NP such as the traveling salesman problem.[22] The best known classical methods to solve NP-complete problems boil down to searching algorithms, which, importantly, can be quadratically accelerated by building on *Grover's algorithm* [162].

For future projects, it would be exciting to extend the quantum field tomography protocol to continuous states that are not translation-invariant, allowing for the description of many new systems and a more accurate characterization of systems that are only approximately translation-invariant. Using the CMPS parameter matrices from a complete tomography, the time evolution of the state can accurately be simulated[132]. This again opens new paths of describing physical processes, such as thermal equilibration, and connecting model and experiment. Still, most investigations are conducted on one-dimensional systems, which are considerably easier to handle than high-dimensional ones. Another step forward could lie in the development of recovery protocols for two-dimensional systems that can be captured by continuous PEPS[133].

Tensor network and compressed sensing methods are not mutually exclusive, and there are many starting points for combining both paradigms. In particular, it seems very appealing to introduce notions of compressed sensing to CMPS tomography, unifying both approaches. An interesting ansatz would be to extend *atomic norm denoising* methods[177] to exponentially decaying signals. Other quickly developing fields related to compressed sensing include *biconvex* methods[178] together with *blind deconvolution* [179]. This is connected to the *self-calibration* setting[180] where, e.g., in the case of quantum tomography, neither the state nor the measurement matrices are exactly known (see for example Ref. [181]). Notions of rank minimization and entry-wise sparsity can be combined as well.[178] Moreover, settings with gradually revealed information/measurements are analyzed ("streaming").[182] Already explicitly storing measurement matrices in memory poses substantial challenges—a more implicit approach is presented in Ref. [183]. Going beyond vectors and matrices, *tensor completion* is investigated.[184] All these approaches promise great theoretical and practical progress and many experiments could greatly benefit from including such compressed sensing ideas.

With the introduction of a fully-fledged quantum version of the singular value decomposition (appendix A.2) and together with the new tools developed for the quantum matrix pencil method, various quantum analogues of classical algorithms with SVD as their central component will become easier to realize. An example could be *singular value thresholding* [82], a compressed sensing routine, which is used for efficiently performing matrix completion.

As quantum systems will play a crucial role in future technology, tools such as the ones presented in this thesis will become ever more important. With this work, we believe we have contributed to the understanding and development of efficient and practical methods for system identification and quantum computation.

ACKNOWLEDGMENTS

First and foremost, I want to thank my supervisor Jens Eisert for granting me this truly great opportunity to work in his group in a very exciting field, facilitating particularly productive collaborations, and providing me with numerous scientific prospects.

Furthermore, I would like to thank Matthias Scheffler, Luca Ghiringhelli, and the theory group of the Fritz Haber Institute of the Max Planck Society in Berlin for their close cooperation. My greatest thanks also go to Seth Lloyd from the Massachusetts Institute of Technology in Cambridge, MA, for his generous hosting, many interesting discussions and a very fruitful collaboration. I thankfully acknowledge the Studienstiftung des deutschen Volkes and the Max Planck Society for their generous financial support.

I want to thank Carlos Riofrío for countless insightful and enjoyable discussions, scientific advice, and magnificent teamwork. Moreover, I am grateful for all the enjoyable collaborations that made the presented projects possible, in particular I want to mention Ingo Roth, Patrick Rebentrost, Iman Marvian, Robert Hübener, Tim Langen, Will McCutcheon, Christian Krumnow, Jason Hoelscher-Obermaier, and Niklas Menzel. Nothing is more motivating than the stimulating environment of a terrific group of interesting and interested people and for this I especially thank the QMIO group at the Dahlem Center.

REFERENCES

[1] A. Steffens, C. A. Riofrío, W. McCutcheon, I. Roth, B. A. Bell, A. McMillan, M. S. Tame, J. G. Rarity, and J. Eisert. "Experimentally exploring compressed sensing quantum tomography". *Quantum Science and Technology* 2:025005 (2017).

[2] A. Steffens, C. A. Riofrío, R. Hübener, and J. Eisert. "Quantum field tomography". *New Journal of Physics* 16:123010 (2014).

[3] A. Steffens, M. Friesdorf, T. Langen, B. Rauer, T. Schweigler, R. Hübener, J. Schmiedmayer, C. A. Riofrío, and J. Eisert. "Towards experimental quantum-field tomography with ultracold atoms". *Nature Communications* 6:7663 (2015).

[4] A. Steffens, P. Rebentrost, I. Marvian, J. Eisert, and S. Lloyd. "An efficient quantum algorithm for spectral estimation". *New Journal of Physics* 19:033005 (2017).

[5] G. Haack, A. Steffens, J. Eisert, and R. Hübener. "Continuous matrix product state tomography of quantum transport experiments". *New Journal of Physics* 17:113024 (2015).

[6] P. Rebentrost, A. Steffens, and S. Lloyd. "Quantum singular value decomposition of nonsparse low-rank matrices". *ArXiv e-prints* (2016). arXiv: 1607.05404.

[7] A. Steffens, I. Roth, C. Krumnow, L. Ghiringhelli, M. Scheffler, and J. Eisert (preliminary ordering). "Compressive density functional theory". *To be published* (2018).

[8] J. P. Dowling and G. J. Milburn. "Quantum technology: the second quantum revolution". *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 361:1655–1674 (2003).

[9] Morgan Stanley Research. "Quantum computing – weird science or the next computing revolution?" (2017).

[10] R.E. Bellman. *Adaptive Control Processes: A Guided Tour*. Princeton Legacy Library. Princeton University Press, 1961. ISBN: 97814008874668.

[11] A. Tarantola. *Inverse problem theory and methods for model parameter estimation*. SIAM, 2005. ISBN: 978-0-89871-572-9.

[12] K. Vogel and H. Risken. "Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase". *Physical Review A* 40:2847 (1989).

[13] J. Řeháček and M. Paris. *Quantum State Estimation*. Vol. 649. Lecture Notes in Physics. Springer-Verlag Berlin Heidelberg, 2004. ISBN: 978-3-540-22329-0.

[14] U. Leonhardt. *Measuring the Quantum State of Light*. Cambridge University Press, Cambridge, UK, 1997. ISBN: 0521497302.

[15] I. L. Chuang and M. A. Nielsen. "Prescription for experimental determination of the dynamics of a quantum black box". *Journal of Modern Optics* 44:2455–2467 (1997).

[16] H. Häffner, W. Hänsel, C. F. Roos, J. Benhelm, D. Chek-Al-Kar, M. Chwalla, T. Körber, U. D. Rapol, M. Riebe, P. O. Schmidt, C. Becher, O. Gühne, W. Dür, and R. Blatt. "Scalable multiparticle entanglement of trapped ions". *Nature* 438:643–646 (2005).

[17] C. F. Roos, G. P. T. Lancaster, M. Riebe, H. Häffner, W. Hänsel, S. Gulde, C. Becher, J. Eschner, F. Schmidt-Kaler, and R. Blatt. "Bell States of Atoms with Ultralong Lifetimes and Their Tomographic State Analysis". *Physical Review Letters* 92:220402 (2004).

[18] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White. "Measurement of qubits". *Physical Review A* 64:052312 (2001).

[19] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani. "Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum". *Physical Review Letters* 70:1244–1247 (1993).

[20] R. P. Feynman. "Simulating Physics with Computers". *International Journal of Theoretical Physics* 21:467–488 (1982).

[21] S. Lloyd. "Universal Quantum Simulators". *Science* 273:1073–1078 (1996).

[22] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge Series on Information and the Natural Sciences. Cambridge University Press, 2000. ISBN: 9780521635035.

[23] C. A. Fuchs and J. van de Graaf. "Cryptographic Distinguishability Measures for Quantum-mechanical States". *IEEE Transactions on Information Theory* 45:1216–1227 (1999).

[24] E. Landau. *Handbuch der Lehre von der Verteilung der Primazahlen*. Handbuch der Lehre von der Verteilung der Primzahlen. B. G. Teubner, 1909.

[25] Z. Hradil. "Quantum-state estimation". *Physical Review A* 55:R1561–R1564 (1997).

[26] K. Banaszek. "Maximum-likelihood estimation of photon-number distribution from homodyne statistics". *Physical Review A* 57:5013–5015 (1998).

[27] J. Řeháček, Z. Hradil, E. Knill, and A. I. Lvovsky. "Diluted maximum-likelihood algorithm for quantum tomography". *Physical Review A* 75:042108 (2007).

[28] R. Blume-Kohout. "Optimal, reliable estimation of quantum states". *New Journal of Physics* 12:043034 (2010).

[29] C. W. Helstrom. "Quantum detection and estimation theory". *Journal of Statistical Physics* 1:231–252 (1969).

[30] K. R. W. Jones. "Principles of quantum inference". *Annals of Physics* 207:140–170 (1991).

[31] M. Christandl and R. Renner. "Reliable Quantum State Tomography". *Physical Review Letters* 109:120403 (2012).

[32] C. Schwemmer, L. Knips, D. Richart, H. Weinfurter, T. Moroder, M. Kleinmann, and O. Gühne. "Systematic errors in current quantum state tomography tools". *Physical Review Letters* 114:080403 (2015).

[33] S. T. Flammia and Y.-K. Liu. "Direct fidelity estimation from few Pauli measurements". *Physical review letters* 106:230501 (2011).

[34] M. P. da Silva, O. Landon-Cardinal, and D. Poulin. "Practical Characterization of Quantum Devices without Tomography". *Physical Review Letters* 107:210404 (2011).

[35] O. Gühne, C.-Y. Lu, W.-B. Gao, and J.-W. Pan. "Toolbox for entanglement detection and fidelity estimation". *Physical Review A* 76:030305 (2007).

[36] S. Aaronson. "The learnability of quantum states". *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 463:3089–3114 (2007).

[37] S. Aaronson. "Shadow Tomography of Quantum States". *ArXiv e-prints* (Nov. 2017). arXiv: 1711.01053 [quant-ph].

[38] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland. "Randomized benchmarking of quantum gates". *Physical Review A* 77:012307 (2008).

[39] J. J. Wallman and S. T. Flammia. "Randomized benchmarking with confidence". *New Journal of Physics* 16:103032 (2014).

[40] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert. "Quantum state tomography via compressed sensing". *Physical review letters* 105:150401 (2010).

[41] C. A. Riofrío, D. Gross, S. T. Flammia, T. Monz, D. Nigg, R. Blatt, and J. Eisert. "Experimental quantum compressed sensing for a seven-qubit system". *Nature Communications* 8:15305 (2017).

[42] S. T. Flammia, D. Gross, Y.-K. Liu, and J. Eisert. "Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators". *New Journal of Physics* 14:095022 (2012).

[43] R. Kueng, H. Rauhut, and U. Terstiege. "Low rank matrix recovery from rank one measurements". *Applied and computational harmonic analysis* 42:88–116 (2017).

[44] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu. "Sample-optimal Tomography of Quantum States". *Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing*. STOC '16. Cambridge, MA, USA: ACM, New York, NY, USA, 2016, pp. 913–925. ISBN: 978-1-4503-4132-5.

[45] C. Schwemmer, G. Tóth, A. Niggebaum, T. Moroder, D. Gross, O. Gühne, and H. Weinfurter. "Experimental Comparison of Efficient Tomography Schemes for a Six-Qubit State". *Physical Review Letters* 113:040503 (2014).

[46] M. B. Hastings. "Solving gapped Hamiltonians locally". *Physical Review B* 73:085115 (2006).

[47] J. Eisert. "Entanglement and tensor network states". *ArXiv e-prints* (2013). arXiv: 1308.3318.

[48] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu. "Efficient quantum state tomography". *Nature Communications* 1:149 (2010).

[49] B. K. Natarajan. "Sparse Approximate Solutions to Linear Systems". *SIAM Journal on Computing* 24:227–234 (1995).

[50] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004. ISBN: 0521833787.

[51] J. B. Hiriart-Urruty and C. Lemarechal. *Convex Analysis and Minimization Algorithms I: Fundamentals*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 1996. ISBN: 9783540568506.

[52] D. L. Donoho and M. Elad. "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization". *Proceedings of the National Academy of Sciences* 100:2197–2202 (2003).

[53] E. J. Candès and T. Tao. "Decoding by linear programming". *IEEE Transactions on Information Theory* 51:4203–4215 (2005).

[54] E. J. Candès and T. Tao. "Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?" *IEEE Transactions on Information Theory* 52:5406–5425 (2006).

[55] E. J. Candès, J. Romberg, and T. Tao. "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information". *Information Theory, IEEE Transactions on* 52:489–509 (2006).

[56] D. L. Donoho. "Compressed sensing". *IEEE Transactions on Information Theory* 52:1289–1306 (2006).

[57] M. A. Davenport. "Random observations on random observations: Sparse signal acquisition and processing". PhD thesis. Rice University, 2010.

[58] J. P. Haldar, D. Hernando, and Z. P. Liang. "Compressed-Sensing MRI With Random Encoding". *IEEE Transactions on Medical Imaging* 30:893–903 (2011).

[59] L. Gao, J. Liang, C. Li, and L. V. Wang. "Single-shot compressed ultrafast photography at one hundred billion frames per second". *Nature* 516:74–77 (2014).

[60] X. Andrade, J. N. Sanders, and A. Aspuru-Guzik. "Application of compressed sensing to the simulation of atomic systems". *Proceedings of the National Academy of Science* 109:13928–13933 (2012).

[61] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. "Robust Face Recognition via Sparse Representation". *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31:210–227 (2009).

[62] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski. "Optimization with Sparsity-Inducing Penalties". *Foundations and Trends in Machine Learning* 4:1–106 (2012).

[63] M. A. Herman and T. Strohmer. "High-Resolution Radar via Compressed Sensing". *IEEE Transactions on Signal Processing* 57:2275–2284 (2009).

[64] Y. C. Eldar and G. Kutyniok. *Compressed Sensing: Theory and Applications*. Compressed Sensing: Theory and Applications. Cambridge University Press, 2012. ISBN: 9780511794308.

[65] S. Foucart and H. Rauhut. *A mathematical introduction to compressive sensing*. Applied and Numerical Harmonic Analysis. Springer New York, 2013. ISBN: 978-0-8176-4948-7.

[66] S. Chen and D. Donoho. *Basis Pursuit*. Tech. rep. 1994.

[67] R. Tibshirani. *Journal of the Royal Statistical Society. Series B (Methodological)* 58:267–288 (1996).

[68] B. Logan. "Properties of High-Pass Signals". PhD thesis. Columbia University, New York, 1965.

[69] H. L. Taylor, S. C. Banks, and J. F. McCoy. "Deconvolution with the ℓ_1 norm". *GEOPHYSICS* 44:39–52 (1979).

[70] G. R. de Prony. "Essai éxperimental et analytique: sur les lois de la dilatabilité de fluides élastique et sur celles de la force expansive de la vapeur de l'alkool, à différentes températures". *Journal de l'École Polytechnique Floréal et Plairial* 1:24–76 (1795).

[71] E. J. Candès and B. Recht. "Exact Matrix Completion via Convex Optimization". *Foundations of Computational Mathematics* 9:717 (2009).

[72] E. J. Candès and T. Tao. "The Power of Convex Relaxation: Near-Optimal Matrix Completion". *IEEE Transactions on Information Theory* 56:2053–2080 (2010).

[73] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry. "Using Collaborative Filtering to Weave an Information Tapestry". *Communications of the ACM* 35:61–70 (1992).

[74] E. J. Candès, T. Strohmer, and V. Voroninski. "PhaseLift: Exact and Stable Signal Recovery from Magnitude Measurements via Convex Programming". *Communications on Pure and Applied Mathematics* 66:1241–1274 (2013).

[75] Z. Liu and L. Vandenberghe. "Interior-Point Method for Nuclear Norm Approximation with Application to System Identification". *SIAM Journal on Matrix Analysis and Applications* 31:1235–1256 (2010).

[76] Y. Amit, M. Fink, N. Srebro, and S. Ullman. "Uncovering Shared Structures in Multiclass Classification". *Proceedings of the 24th International Conference on Machine Learning*. ICML '07. ACM, 2007, pp. 17–24. ISBN: 978-1-59593-793-3.

[77] M. Fazel. "Matrix rank minimization with applications". PhD thesis. Stanford University, 2002.

[78] E. J. Candes and Y. Plan. "Tight Oracle Inequalities for Low-Rank Matrix Recovery From a Minimal Number of Noisy Random Measurements". *IEEE Transactions on Information Theory* 57:2342–2359 (2011).

[79] J. A. Tropp. "Convex Recovery of a Structured Signal from Independent Random Linear Measurements". *Sampling Theory, a Renaissance: Compressive Sensing and Other Developments*. Springer International Publishing, 2015, pp. 67–101. ISBN: 978-3-319-19749-4.

[80] D. Gross. "Recovering low-rank matrices from few coefficients in any basis". *IEEE Transactions on Information Theory* 57:1548–1566 (2011).

[81] L. Vandenberghe and S. Boyd. "Semidefinite Programming". *SIAM Review* 38:49–95 (1996).

[82] J.-F. Cai, E. J. Candès, and Z. Shen. "A Singular Value Thresholding Algorithm for Matrix Completion". *SIAM Journal on Optimization* 20:1956–1982 (2010).

[83] Y.-K. Liu. "Universal low-rank matrix recovery from Pauli measurements". *Advances in Neural Information Processing Systems* 24:1638–1646 (2011).

[84] V. Voroninski. "Quantum Tomography From Few Full-Rank Observables". *ArXiv e-prints* (2013). arXiv: 1309.7669.

[85] R. Kueng. "Low rank matrix recovery from few orthonormal basis measurements". *International Conference on Sampling Theory and Applications*. 2015, pp. 402–406.

[86] R. Kueng and D. Gross. "Qubit stabilizer states are complex projective 3-designs". *ArXiv e-prints* (2015). arXiv: 1510.02767.

[87] W.-T. Liu, T. Zhang, J.-Y. Liu, P.-X. Chen, and J.-M. Yuan. "Experimental Quantum State Tomography via Compressed Sampling". *Physical Review Letters* 108:170403 (2012).

[88] A. Smith, C. A. Riofrío, B. E. Anderson, H. Sosa-Martinez, I. H. Deutsch, and P. S. Jessen. "Quantum state tomography by continuous measurement and compressed sensing". *Physical Review A* 87:030102 (2013).

[89] A. Shabani, R. L. Kosut, M. Mohseni, H. Rabitz, M. A. Broome, M. P. Almeida, A. Fedrizzi, and A. G. White. "Efficient Measurement of Quantum Dynamics via Compressive Sensing". *Physical Review Letters* 106:100401 (2011).

[90] A. Carpentier, J. Eisert, D. Gross, and R. Nickl. "Uncertainty quantification for matrix compressed sensing and quantum tomography problems". *ArXiv e-prints* (2015). arXiv: 1504.03234.

[91] K.P. Burnham and D.R. Anderson. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer New York, 2003. ISBN: 9780387953649.

[92] H. Akaike. "A new look at the statistical model identification". *IEEE Transactions on Automatic Control* 19:716–723 (1974).

[93] G. Schwarz. "Estimating the Dimension of a Model". *The Annals of Statistics* 6:461–464 (1978).

[94] D. Mogilevtsev, Z. Hradil, J. Rehacek, and V. S. Shchesnovich. "Cross-Validated Tomography". *Physical Review Letters* 111:120403 (2013).

[95] K. Usami, Yo. Nambu, Y. Tsuda, K. Matsumoto, and K. Nakamura. "Accuracy of quantum-state estimation utilizing Akaike's information criterion". *Physical Review A* 68:022314 (2003).

[96] M. Guţă, T. Kypraios, and I. Dryden. "Rank-based model selection for multiple ions quantum tomography". *New Journal of Physics* 14:105002 (2012).

[97] B. A. Bell, M. S. Tame, A. S. Clark, R. W. Nock, W. J. Wadsworth, and J. G. Rarity. "Experimental characterization of universal one-way quantum computing". *New Journal of Physics* 15:053030 (2013).

[98] E. H. Lieb and D. W. Robinson. "The Finite Group Velocity of Quantum Spin Systems". *Statistical Mechanics: Selecta of Elliott H. Lieb*. Springer Berlin Heidelberg, 1972, pp. 425–431. ISBN: 978-3-662-10018-9.

[99] M. B. Hastings. "Locality in Quantum Systems". *ArXiv e-prints* (2010). arXiv: 1008.5137.

[100] M. B. Hastings and T. Koma. "Spectral Gap and Exponential Decay of Correlations". *Communications in Mathematical Physics* 265:781–804 (2006).

[101] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. "Quantum entanglement". *Reviews of Modern Physics* 81:865–942 (2009).

[102] J. Eisert, M. Cramer, and M. B. Plenio. "Colloquium: Area laws for the entanglement entropy". *Reviews of Modern Physics* 82:277–306 (2010).

[103] M. B. Hastings. "An area law for one-dimensional quantum systems". *Journal of Statistical Mechanics: Theory and Experiment* 8:08024 (2007).

[104] D. Poulin, A. Qarry, R. Somma, and F. Verstraete. "Quantum Simulation of Time-Dependent Hamiltonians and the Convenient Illusion of Hilbert Space". *Physical Review Letters* 106:170501 (2011).

[105] R. Orús. "A practical introduction to tensor networks: Matrix product states and projected entangled pair states". *Annals of Physics* 349:117–158 (2014).

[106] A. Klümper, A. Schadschneider, and J. Zittartz. "Matrix Product Ground States for One-Dimensional Spin-1 Quantum Antiferromagnets". *Europhysics Letters* 24:293 (1993).

[107] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac. "Matrix Product State Representations". *Quantum Information and Computation* 7:401–430 (2007).

[108] F. Verstraete and J. I. Cirac. "Matrix product states represent ground states faithfully". *Physical Review B* 73:094423 (2006).

[109] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki. "Rigorous results on valence-bond ground states in antiferromagnets". *Physical Review Letters* 59:799–802 (1987).

[110] C. K. Majumdar and D. K. Ghosh. "On Next-Nearest-Neighbor Interaction in Linear Chain." *Journal of Mathematical Physics* 10:1388 (1969).

[111] X. Chen, Z.-C. Gu, and X.-G. Wen. "Classification of gapped symmetric phases in one-dimensional spin systems". *Physical Review B* 83:035107 (2011).

[112] M. M. Wolf, G. Ortiz, F. Verstraete, and J. I. Cirac. "Quantum Phase Transitions in Matrix Product Systems". *Physical Review Letters* 97:110403 (2006).

[113] S. R. White. "Density matrix formulation for quantum renormalization groups". *Physical Review Letters* 69:2863–2866 (1992).

[114] G. Vidal. "Efficient Classical Simulation of Slightly Entangled Quantum Computations". *Physical Review Letters* 91:147902 (2003).

[115] U. Schollwöck. "The density-matrix renormalization group in the age of matrix product states". *Annals of Physics* 326:96–192 (2011).

[116] J. C. Bridgeman and C. T. Chubb. "Hand-waving and interpretive dance: an introductory course on tensor networks". *Journal of Physics A: Mathematical and Theoretical* 50:223001 (2017).

[117] M. Bachmayr, R. Schneider, and A. Uschmajew. "Tensor Networks and Hierarchical Tensors for the Solution of High-Dimensional Partial Differential Equations". *Foundations of Computational Mathematics* 16:1423–1472 (2016).

[118] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos. "Tensor Decomposition for Signal Processing and Machine Learning". *IEEE Transactions on Signal Processing* 65:3551–3582 (2017).

[119] I. V. Oseledets. "Tensor-Train Decomposition". *SIAM Journal on Scientific Computing* 33:2295–2317 (2011).

[120] F. Verstraete and J. I. Cirac. "Renormalization algorithms for Quantum-Many Body Systems in two and higher dimensions". *ArXiv e-prints* (2004). arXiv: cond-mat/0407066.

[121] Y.-Y. Shi, L.-M. Duan, and G. Vidal. "Classical simulation of quantum many-body systems with a tree tensor network". *Physical Review A* 74:022320 (2006).

[122] G. Vidal. "Entanglement Renormalization". *Physical Review Letters* 99:220405 (2007).

[123] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac. "Matrix Product Density Operators: Simulation of Finite-Temperature and Dissipative Systems". *Physical Review Letters* 93:207204 (2004).

[124] S. T. Flammia, D. Gross, S. D. Bartlett, and R. Somma. "Heralded Polynomial-Time Quantum State Tomography". *ArXiv e-prints* (2010). arXiv: 1002.3839.

[125] F. Verstraete and J. I. Cirac. "Continuous Matrix Product States for Quantum Fields". *Physical Review Letters* 104:190405 (2010).

[126] T. J. Osborne, J. Eisert, and F. Verstraete. "Holographic Quantum States". *Physical Review Letters* 105:260401 (2010).

[127] E. H. Lieb and W. Liniger. "Exact Analysis of an Interacting Bose Gas. I. The General Solution and the Ground State". *Physical Review* 130:1605–1616 (1963).

[128] M. Gaudin. "Un systeme a une dimension de fermions en interaction". *Physics Letters A* 24:55–56 (1967).

[129] S. S. Chung, S. Bauman, Kuei Sun, and C. J. Bolech. "On the new Continuous Matrix Product Ansatz". *Journal of Physics: Conference Series* 702:012004 (2016).

[130] J. Haegeman, J. I. Cirac, T. J. Osborne, H. Verschelde, and F. Verstraete. "Applying the Variational Principle to $(1+1)$-Dimensional Quantum Field Theories". *Physical Review Letters* 105:251601 (2010).

[131] M Ganahl, J. Rincón, and G. Vidal. "Continuous Matrix Product States for Quantum Fields: An Energy Minimization Algorithm". *Physical Review Letters* 118:220402 (2017).

[132] D. Draxler, J. Haegeman, T. J. Osborne, V. Stojevic, L. Vanderstraeten, and F. Verstraete. "Particles, Holes, and Solitons: A Matrix Product State Approach". *Physical Review Letters* 111:020402 (2013).

[133] D. Jennings, C. Brockt, J. Haegeman, T. J. Osborne, and F. Verstraete. "Continuum tensor network field states, path integral representations and spatial symmetries". *New Journal of Physics* 17:063039 (2015).

[134] J. Haegeman, J. I. Cirac, T. J. Osborne, and F. Verstraete. "Calculus of continuous matrix product states". *Physical Review B* 88:085118 (2013).

[135] R. Hübener, A. Mari, and J. Eisert. "Wick's Theorem for Matrix Product States". *Physical Review Letters* 110:040401 (2013).

[136] M.E. Peskin and D.V. Schroeder. *An Introduction to Quantum Field Theory*. Advanced book classics. Avalon Publishing, 1995. ISBN: 9780201503975.

[137] D. Potts and M. Tasche. "Parameter estimation for nonincreasing exponential sums by Prony-like methods". *Linear Algebra and its Applications* 439:1024–1039 (2013). ISSN: 0024-3795.

[138] Y. Hua and T. K. Sarkar. "Generalized pencil-of-function method for extracting poles of an EM system from its transient response". *IEEE Transactions on Antennas and Propagation* 37:229–234 (1989).

[139] Y. Hua and T. K. Sarkar. "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise". *IEEE Transactions on Acoustics, Speech, and Signal Processing* 38:814–824 (1990).

[140] J. Reichel and V. Vuletic. *Atom Chips*. Wiley, 2010. ISBN: 9783527633364.

[141] S. Hofferberth, I. Lesanovsky, B. Fischer, T. Schumm, and J. Schmiedmayer. "Non-equilibrium coherence dynamics in one-dimensional Bose gases". *Nature* 449:324–327 (2007).

[142] T. Langen, R. Geiger, M. Kuhnert, B. Rauer, and J. Schmiedmayer. "Local emergence of thermal correlations in an isolated quantum many-body system". *Nature Physics* 9:640–643 (2013).

[143] T. Langen. *Non-equilibrium Dynamics of One-Dimensional Bose Gases*. Springer Theses. Springer International Publishing, 2015. ISBN: 9783319185644.

[144] J. Eisert, M. Friesdorf, and C. Gogolin. "Quantum many-body systems out of equilibrium". *Nature Physics* 11:124–130 (2015).

[145] I. Maruyama and H. Katsura. "Continuous Matrix Product Ansatz for the One-Dimensional Bose Gas with Point Interaction". *Journal of the Physical Society of Japan* 79:073002 (2010).

[146] B. Efron. "Bootstrap Methods: Another Look at the Jackknife". *The Annals of Statistics* 7:1–26 (1979).

[147] B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. Monographs on Statistics and Applied Probability. Chapman & Hall/CRC, 1993. ISBN: 9780412042317.

[148] M. Cramer, C. M. Dawson, J. Eisert, and T. J. Osborne. "Exact Relaxation in a Class of Nonequilibrium Quantum Lattice Systems". *Physical Review Letters* 100:030602 (2008).

[149] M. M. Waldrop. "The chips are down for Moore's law". *Nature* 530:144 (2016).

[150] E. Gibney. "The super materials that could trump graphene". *Nature* 522:274–276 (2015).

[151] J. Knechtel, O. Sinanoglu, I. M. Elfadel, J. Lienig, and C. C. N. Sze. "Large-Scale 3D Chips: Challenges and Solutions for Design Automation, Testing, and Trustworthy Integration". *IPSJ Transactions on System LSI Design Methodology* 10:45–62 (2017).

[152] C. Zalka. "Simulating quantum systems on a quantum computer". *Proceedings of the Royal Society of London Series A* 454:313 (1998).

[153] D. S. Abrams and S. Lloyd. "Quantum Algorithm Providing Exponential Speed Increase for Finding Eigenvalues and Eigenvectors". *Physical Review Letters* 83:5162–5165 (1999).

[154] J. I. Cirac and P. Zoller. "Quantum Computations with Cold Trapped Ions". *Physical Review Letters* 74:4091–4094 (1995).

[155] P. G. Kwiat, J. R. Mitchell, P. D. D. Schwindt, and A. G. White. "Grover's search algorithm: an optical approach". *Journal of Modern Optics* 47:257–266 (2000).

[156] Carl Pomerance. "A tale of two sieves". *Notices of the American Mathematical Society* 43:1473–1485 (1996).

[157] J. Preskill. "Quantum computing and the entanglement frontier". *ArXiv e-prints* (2012). arXiv: 1203.5813.

[158] V. Giovannetti, S. Lloyd, and L. Maccone. "Quantum Random Access Memory". *Physical Review Letters* 100:160501 (2008).

[159] V. Giovannetti, S. Lloyd, and L. Maccone. "Architectures for a quantum random access memory". *Physical Review A* 78:052310 (2008).

[160] D. Gottesman. "Class of quantum error-correcting codes saturating the quantum Hamming bound". *Physical Review A* 54:1862–1868 (1996).

[161] D. P. DiVincenzo. "The Physical Implementation of Quantum Computation". *Fortschritte der Physik* 48:771–783 (2000).

[162] L. K. Grover. "A fast quantum mechanical algorithm for database search". *Proceedings, 28th Annual ACM Symposium on the Theory of Computing*. 1996, p. 212.

[163] G. Brassard and P. Hoyer. "An exact quantum polynomial-time algorithm for Simon's problem". *Proceedings of the Fifth Israeli Symposium on Theory of Computing and Systems*. 1997, pp. 12–23.

[164] F. G. S. L. Brandão and K. Svore. "Quantum Speed-ups for Semidefinite Programming". *ArXiv e-prints* (2016). arXiv: 1609.05537.

[165] F. G. S. L. Brandão, A. Kalev, T. Li, C. Yen-Yu Lin, K. M. Svore, and X. Wu. "Exponential Quantum Speed-ups for Semidefinite Programming with Applications to Quantum Learning". *ArXiv e-prints* (Oct. 2017). arXiv: 1710.02581 [quant-ph].

[166] A. W. Harrow, A. Hassidim, and S. Lloyd. "Quantum Algorithm for Linear Systems of Equations". *Physical Review Letters* 103:150502 (2009).

[167] F. Le Gall. "Powers of Tensors and Fast Matrix Multiplication". *Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation*. ISSAC '14. ACM, 2014, pp. 296–303. ISBN: 978-1-4503-2501-1.

[168] N. Wiebe, D. Braun, and S. Lloyd. "Quantum Algorithm for Data Fitting". *Physical Review Letters* 109:050505 (2012).

[169] G. Wang. "Quantum algorithm for linear regression". *Physical Review A* 96:012335 (2017).

[170] P. Rebentrost, M. Mohseni, and S. Lloyd. "Quantum Support Vector Machine for Big Data Classification". *Physical Review Letters* 113:130503 (2014).

[171] S. Lloyd, M. Mohseni, and P. Rebentrost. "Quantum principal component analysis". *Nature Physics* 10:631–633 (2014).

[172] URL: `math.nist.gov/quantum/zoo/`.

[173] A. M. Childs and J. M. Eisenberg. "Quantum Algorithms for Subset Finding". *Quantum Information and Computation* 5:593–604 (2005).

[174] G. Wang. "Efficient quantum algorithms for analyzing large sparse electrical networks". *Quantum Information and Computation* 17:987–1026 (2017).

[175] A. Montanaro. "Quantum algorithms: an overview". *npj Quantum Information* 2:15023 (2016).

[176] S. Aaronson. "Read the fine print". *Nature Physics* 11:291–293 (2015).

[177] B. N. Bhaskar, G. Tang, and B. Recht. "Atomic Norm Denoising With Applications to Line Spectral Estimation". *IEEE Transactions on Signal Processing* 61:5987–5999 (2013).

[178] S. Ling and T. Strohmer. "Self-calibration and biconvex compressive sensing". *Inverse Problems* 31:115002 (2015).

[179] P. Walk, P. Jung, G. E. Pfander, and B. Hassibi. "Blind Deconvolution with Additional Autocorrelations via Convex Programs". *ArXiv e-prints* (2017). arXiv: 1701.04890.

[180] L. Cohen, Y. Pilnyak, D. Istrati, N. M. Studer, J. P. Dowling, and H. S. Eisenberg. "Absolute self-calibration of single-photon and multiplexed photon-number-resolving detectors". *ArXiv e-prints* (2017). arXiv: 1711.03594.

[181] C. Stark. "Self-consistent tomography of the state-measurement Gram matrix". *Physical Review A* 89:052109 (2014).

[182] J. A. Tropp, A. Yurtsever, M. Udell, and V. Cevher. "Fixed-Rank Approximation of a Positive-Semidefinite Matrix from Streaming Data". *ArXiv e-prints* (2017). arXiv: 1706.05736.

[183] A. Yurtsever, M. Udell, J. A. Tropp, and V. Cevher. "Sketchy Decisions: Convex Low-Rank Matrix Optimization with Optimal Storage". *ArXiv e-prints* (2017). arXiv: 1702.06838.

[184] J. Liu, P. Musialski, P. Wonka, and J. Ye. "Tensor Completion for Estimating Missing Values in Visual Data". *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35:208–220 (2013).

[185] N. Ubbelohde, C. Fricke, C. Flindt, F. Hohls, and R. J. Haug. "Measurement of finite-frequency current statistics in a single-electron transistor". *Nature Communications* 3:612 (2012).

[186] D. Aharonov and A. Ta-Shma. "Adiabatic Quantum State Generation and Statistical Zero Knowledge". *Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Computing*. STOC '03. San Diego, CA, USA: ACM, 2003, pp. 20–29. ISBN: 1-58113-674-9.

[187] W. Kohn. "Nobel Lecture: Electronic structure of matter-wave functions and density functionals". *Reviews of Modern Physics* 71:1253 (1999).

[188] R. O. Jones and O. Gunnarsson. "The density functional formalism, its applications and prospects". *Reviews of Modern Physics* 61:689–746 (1989).

[189] P. Hohenberg and W. Kohn. "Inhomogeneous electron gas". *Physical review* 136:B864 (1964).

[190] R.G. Parr and W. Yang. *Density-Functional Theory of Atoms and Molecules*. International Series of Monographs on Chemistry. Oxford University Press, USA, 1994. ISBN: 9780195092769.

[191] R.M. Dreizler and E.K.U. Gross. *Density Functional Theory: An Approach to the Quantum Many-Body Problem*. Springer Berlin Heidelberg, 2012. ISBN: 9783642861055.

[192] N. Schuch and F. Verstraete. "Computational complexity of interacting electrons and fundamental limitations of density functional theory". *Nature Physics* 5:732–735 (2009).

[193] W. Kohn and L. J. Sham. "Self-Consistent Equations Including Exchange and Correlation Effects". *Physical Review* 140:A1133–A1138 (1965).

[194] J. P. Perdew and Y. Wang. "Accurate and simple analytic representation of the electron-gas correlation energy". *Physical Review B* 45:13244–13249 (1992).

[195] J. P. Perdew, K. Burke, and M. Ernzerhof. "Generalized Gradient Approximation Made Simple". *Physical Review Letters* 77:3865–3868 (1996).

[196] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. "Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields". *The Journal of Physical Chemistry* 98:11623–11627 (1994).

[197] A. Tkatchenko and M. Scheffler. "Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data". *Physical Review Letters* 102:073005 (2009).

[198] R. Ditchfield, W. J. Hehre, and J. A. Pople. "Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules". *The Journal of Chemical Physics* 54:724–728 (1971).

[199] T. H. Dunning Jr. "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen". *The Journal of Chemical Physics* 90:1007–1023 (1989).

[200] F. Jensen. "Polarization consistent basis sets: Principles". *The Journal of Chemical Physics* 115:9113–9125 (2001).

[201] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler. "Ab initio molecular simulations with numeric atom-centered orbitals". *Computer Physics Communications* 180:2175–2196 (2009).

[202] K. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, and T. L. Windus. "Basis Set Exchange: A Community Database for Computational Sciences". *Journal of Chemical Information and Modeling* 47:1045–1052 (2007). URL: dblp.org.

[203] S. Goedecker. "Linear scaling electronic structure methods". *Reviews of Modern Physics* 71:1085–1123 (1999).

[204] D. R. Bowler and T. Miyazaki. "O(N) methods in electronic structure calculations". *Reports on Progress in Physics* 75:036503 (2012).

[205] V. Ozolins, R. Lai, R. Caflisch, and S. Osher. "Compressed modes for variational problems in mathematics and physics". *Proceedings of the National Academy of Sciences* 110:18368–18373 (2013).

[206] J. C. Budich, J. Eisert, E. J. Bergholtz, S. Diehl, and P. Zoller. "Search for localized Wannier functions of topological band structures via compressed sensing". *Physical Review B* 90:115110 (2014).

[207] L. M. Bregman. "The Relaxation Method of Finding the Common Point of Convex Sets and Its Application to the Solution of Problems in Convex Programming". *USSR Computational Mathematics and Mathematical Physics* 7:200–217 (1967).

[208] R. Lai and S. Osher. "A Splitting Method for Orthogonality Constrained Problems". *Journal of Scientific Computing* 58:431–449 (2014). ISSN: 0885-7474.

[209] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers". *Foundations and Trends in Machine Learning* 3:1–122 (2011). ISSN: 1935-8237.

[210] R. Lai, J. Lu, and S. Osher. "Density matrix minimization with ell_1 regularization". *Communications in Mathematical Sciences* 13:2097–2117 (2015).

[211] R. Lai and J. Lu. "Localized density matrix minimization and linear-scaling algorithms". *Journal of Computational Physics* 315:194–210 (2016).

[212] J. Lu and K. Thicke. "Orbital minimization method with ℓ_1 regularization". *Journal of Computational Physics* 336:87–103 (2017).

[213] L. M. Ghiringhelli, J. Vybiral, E. Ahmetcik, R. Ouyang, S. V. Levchenko, C. Draxl, and M. Scheffler. "Learning physical descriptors for materials science by compressed sensing". *New Journal of Physics* 19:023017 (2017).

[214] A. Argyriou, T. Evgeniou, and M. Pontil. "Convex multi-task feature learning". *Machine Learning* 73:243–272 (2008).

[215] P.-A. Absil, R. Mahony, and R. Sepulchre. "Optimization algorithms on matrix manifolds". *Foundations of Computational Mathematics* 10:241–244 (2010).

[216] P.-A. Absil, C. G. Baker, and K. A. Gallivan. "Trust-region methods on Riemannian manifolds". *Foundations of Computational Mathematics* 7:303–330 (2007).

[217] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre. "Manopt, a Matlab Toolbox for Optimization on Manifolds". *Journal of Machine Learning Research* 15:1455–1459 (2014).

[218] V. Havu, V. Blum, P. Havu, and M. Scheffler. "Efficient integration for all-electron electronic structure calculation using numeric basis functions". *Journal of Computational Physics* 228:8367–8379 (2009).

[219] MATLAB. *version 8.4.0 (R2014b)*. The MathWorks Inc., 2014.

[220] Jmol: an open-source Java viewer for chemical structures in 3D. http://www.jmol.org/.

[221] S. F. Boys. "Electronic wave functions. I. A general method of calculation for the stationary states of any molecular system". 200:542–554 (1950).

[222] P. Pulay. "Convergence acceleration of iterative sequences. The case of SCF iteration". *Chemical Physics Letters* 73:393–398 (1980).

APPENDIX

A COAUTHORED PUBLICATIONS

A.1 Quantum transport experiments

Continuous matrix product states can be generated using a sequential preparation procedure via continuous measurements, as described in Ref. [126]: A one-dimensional continuous quantum system with Fock space $\mathscr{F}$ in the initial vacuum state $|\Omega\rangle$ is coupled with a finite-dimensional auxiliary system $\mathscr{A} \cong \mathbb{C}^b$ in the initial state $|\varphi_0\rangle$. One can think of $\mathscr{A}$ as a resonating cavity with b internal levels and a particle source that emits a particle every time step ε for a time interval $[0, L]$. Between emitting two particles, the system evolves freely according to a Hamiltonian $K(t) \in \mathbb{C}^{b\times b}$, and the measurement process is modeled by matrices $R(t) \in \mathbb{C}^{b\times b}$, resulting in the total Hamiltonian

$$\hat{H}_\varepsilon(t) = K(t)\otimes\hat{\mathbb{1}} + \varepsilon^{1/2}\sum_{k=1}^{L/\varepsilon}\delta(t-k\varepsilon)\left(\mathrm{i}R(L-k\varepsilon)\otimes\hat{a}^\dagger_{L/\varepsilon-k} - \mathrm{i}R^\dagger(L-k\varepsilon)\otimes\hat{a}_{L/\varepsilon-k}\right). \quad (30)$$

Integrating the Schrödinger equation, setting $Q(t) = -\mathrm{i}K(t) - 1/2\,R^\dagger(t)R(t)$, and decoupling $\mathscr{F}$ and $\mathscr{A}$ by projecting onto $\langle\varphi_L|\otimes\hat{\mathbb{1}}$, we arrive in the limit $\varepsilon\to 0$ at the CMPS definition for $|\Psi_{Q,R}\rangle$ in Eq. (23).

By tracing out the physical system $\mathscr{F}$ instead of the auxiliary system $\mathscr{A}$ and computing the derivative, we obtain a differential equation for the resulting reduced density matrix ϱ in $\mathscr{A}$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho(t) = -\mathrm{i}[K(t),\rho(t)] + R^\dagger(t)\rho(t)R(t) - \frac{1}{2}[R^\dagger(t)R(t),\rho(t)]_+, \quad (31)$$

which is a master equation in *Lindblad* form, implying dissipative dynamics in the auxiliary system. This makes it interesting to connect the auxiliary system to the dynamics of systems that can be characterized by a CMPS, which is done in the following publication [5] in the context of quantum transport experiments.

In the underlying experimental setup of the following publication [5][24], separate electrons tunnel through a single-level quantum dot[185]. We have developed a protocol for estimating the parameters that determine the system based on counting probabilities, extending the protocol presented in section 3.2, which uses spatial correlation functions as input. With the established CMPS model, it is furthermore possible to accurately predict system statistics like higher-order correlation functions and the waiting time distribution (WTD), modeling the statistics of time interval between the transport of two consecutive electrons. When emitting electrons with a rate in the kilohertz frequency range, a direct determination of the WTD from the experiment is still possible with current experimental techniques. This allowed us to compare the WTD with the one estimated by our CMPS

[24] Géraldine Haack, Adrian Steffens, Jens Eisert, and Robert Hübener, "Continuous matrix product state tomography of quantum transport experiments", *New Journal of Physics* 17:113024, 2015 (DOI:10.1088/1367-2630/17/11/113024). Published under a Creative Commons Attribution 3.0 License (creativecommons.org/licenses/by/3.0), © 2015 IOP Publishing.

protocol, revealing high consistency between both signals. When moving on to even higher emission frequencies in the gigahertz range, a direct experimental estimation of the WTD is not feasible, while, however, the CMPS approach still remains valid, which makes it possible to uncover the short-time dynamics of systems in higher frequency regimes.

COAUTHORED PUBLICATIONS

[5] Continuous matrix product state tomography of quantum transport experiments
Géraldine Haack, Adrian Steffens, Jens Eisert, and Robert Hübener,
New Journal of Physics 17:113024, 2015.

[6] Quantum singular value decomposition of non-sparse low-rank matrices
Patrick Rebentrost, Adrian Steffens, and Seth Lloyd,
ArXiv e-prints 1607.05404, 2016.

IOP Publishing New J. Phys. 17 (2015) 113024 doi:10.1088/1367-2630/17/11/113024

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

PAPER

Continuous matrix product state tomography of quantum transport experiments

G Haack[1,2,3,4], A Steffens[3], J Eisert[3] and R Hübener[3]

[1] University of Grenoble Alpes, INAC-SPSMS, F-38000 Grenoble, France
[2] CEA, INAC-SPSMS, F-38000 Grenoble, France
[3] Dahlem Center for Quantum Complex Systems and Fachbereich Physik, Freie Universität Berlin, D-14195 Berlin, Germany
[4] Author to whom any correspondence should be addressed.

E-mail: geraldine.haack@unige.ch

Keywords: tomography, waiting times, correlation functions, transport, condensed matter physics

OPEN ACCESS

RECEIVED
22 June 2015

REVISED
29 September 2015

ACCEPTED FOR PUBLICATION
5 October 2015

PUBLISHED
6 November 2015

Abstract

In recent years, a close connection between the description of open quantum systems, the input–output formalism of quantum optics, and continuous matrix product states (cMPS) in quantum field theory has been established. The latter constitute a variational class of one-dimensional quantum field states and have been shown to provide an efficient ansatz for performing tomography of open quantum systems. So far, however, the connection between cMPS and open quantum systems has not yet been developed for quantum transport experiments in the condensed-matter context. In this work, we first present an extension of the tomographic possibilities of cMPS by demonstrating the validity of reconstruction schemes based on low-order counting probabilities compared to previous schemes based on low-order correlation functions. We then show how fermionic quantum transport settings can be formulated within the cMPS framework. Our procedure, via the measurements of low-order correlation functions only, allows us to gain access to quantities that are not directly measurable with present technology. Emblematic examples are high-order correlations functions and waiting time distributions (WTD). The latter are of particular interest since they offer insights into short-time scale physics. We demonstrate the functioning of the method with actual data, opening up the way to accessing WTD within the quantum regime.

1. Introduction

Continuous matrix product states (cMPS) have recently been recognized as powerful and versatile descriptions of certain one-dimensional quantum field states [1–3]. As continuum limits of the MPS—a well-established type of tensor network states underlying the density-matrix renormalisation group [4]—they introduce the intuition developed in quantum lattice models to the realm of quantum fields, offering similar conceptual and numerical tools. In the cMPS framework, interacting quantum fields such as those described by Lieb–Liniger models have been studied, both in theory [1, 5, 6] and in the context of experiments with ultra-cold atoms [7].

On a formal level, cMPS are intricately related to Markovian *open quantum systems* [1, 2]: the open quantum system takes the role of an ancillary system in a sequential preparation picture of cMPS. Elaborating on this formal analogy, cMPS can capture properties of fields that are coupled to a finite dimensional open quantum system. This connection has been fleshed out already in the description of fermionic quantum fields [8] and of light emitted from cavities in cavity-QED [2, 9] in the quantum optical context, under the keyword of the *input–output formalism* [10].

Another methodological ingredient to this work is that cMPS have been identified as tools to perform efficient quantum state tomography of quantum field systems [7, 11–14], related to other approaches of tensor network quantum tomography [12, 15]. These efforts are in line with the emerging mindset that for quantum many-body and quantum field states, tomography and state reconstruction only make sense within a certain statistical model or a variational class of states. Importantly, in our context at hand, it turns out that cMPS can be

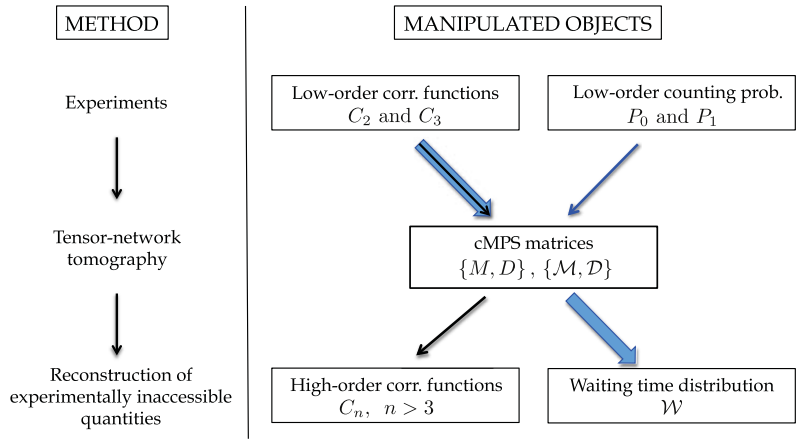

Figure 1. Extension and applications of cMPS based tomography (see sections 3–5). Previous works [7, 11] have shown that measurements of low-order correlation functions C_n with $n = 2, 3$ are sufficient to access higher-order correlations using cMPS tomography based on the cMPS matrices M and D (right side, black arrows). The first achievement of this work is the formal demonstration that measurements of low-order counting probabilities constitute an alternative to measurements of low-order correlation functions for carrying out cMPS tomography (blue arrow). We show that the probabilities to detect zero or one particle (P_0 and P_1, respectively) are sufficient to reconstruct alternative cMPS parameter matrices $\mathcal{M}$ and $\mathcal{D}$, from which higher-order correlation functions can be computed. The second achievement of this work is to extend the applicability of this cMPS framework to quantum transport experiments. As a main illustration, we show that cMPS-based tomography provides an access to the distribution of waiting times $\mathcal{W}$. The according statistics are not directly measurable due to experimental limitations on single-particle detectors. However, we demonstrate with experimental data that they can in fact be reconstructed from the knowledge of low-order correlation functions (broad blue arrows).

reconstructed from the knowledge of low-order correlation functions alone [11, 12]. This is a very attractive feature of cMPS: recently, a reconstruction scheme has successfully been applied to data on quantum fields obtained with ultra-cold Bose gases [7]. In section 2, after introducing the cMPS formulation, we will provide the reader with the key arguments that make this reconstruction scheme possible from low-order correlation functions only. Read in the mindset of open quantum systems, cMPS tomography can be interpreted as open system tomography by monitoring the environment of the open quantum system.

In this work, these methodological components will be put into a different physical context and substantially developed further as illustrated in figure 1. At the heart of the analysis is a tomographic approach, applied to an open quantum system, yet brought to a new level. In section 3, we extend the set of tomographic methods within the cMPS framework, showing that the dynamics of the ancillary system and of the whole open quantum system is not only accessible from low-order correlation functions, but also from low-order counting statistics. Specifically, we prove that for generic systems, the two density functions P_0 and P_1—which express the probability of detecting zero and one particle, respectively, as a function of the time since the last detection—provide sufficient knowledge to successfully perform tomography of the open quantum system.The physical application of the established methods will also be different from the cavity-QED or the quantum field context: here, we treat *fermionic quantum transport experiments* within the cMPS framework.

In a general transport setting, a scatterer is coupled to a left reservoir (the 'source') and a right reservoir (the 'drain'). Fermions (with or without a spin degree of freedom) can be seen as jumping in and out of the scattering region from the source to the drain and can be described by a leaking-out fermionic quantum field. In section 4, we show how the dynamics of the open quantum system (scatterer and leaking-out fermionic field) can be encoded into a cMPS state vector. To provide the reader with an intuition about the equivalence between the cMPS language and a more traditional Hamiltonian formulation, we will consider one of the simplest setups in quantum transport: a single-level quantum dot weakly coupled to two reservoirs. These results are also valid for transport experiments of ultra-cold fermions between a 'hot' and a 'cold' reservoir as recently realised in [16, 17].

This will clear the way for making use of the tomographic possibilities offered by the cMPS formalism to access various quantities in quantum transport that are not yet measurable with current experimental technologies (see figure 1). Emblematic examples are higher-order charge correlation functions and the distribution of waiting times (WTD, see section 5).

The waiting time is defined as the time interval between the arrivals of two *consecutive* electrons. Therefore, the WTD provides a privileged access to short-time physics, short-range interactions and the statistics of the particles. As such, it has gained a lot of attention recently [18–25], but WTDs suffer from their difficulty to be

measured effectively: measuring WTDs requires the detection of single events while ensuring that no events have been missed—for instance, due to the dead time of the detector.

With present technologies, WTDs in transport experiments can be measured when the injection rate of electrons is within the kHz range as in the experiments of [26, 27]. Indeed, at those frequencies, the current trace is resolved in time and the WTD can be directly deduced from it. As we will show in section 5, the WTD reflects the quantum statistics of the electrons. However, quantum coherence and entanglement cannot be detected at those frequencies. To observe these quantum effects, one needs to move to the GHz regime, which can be achieved either with DC sources with a typical bias of tens of meV, or with periodically driven sources at GHz frequencies [28–32]. In the GHz range, the current trace cannot be resolved in time so that the measurement of the WTD is not feasible at present. In contrast, second- and third-order correlation functions have been proven to be feasible [32, 33].

With these experimental constraints in mind, we propose in section 5 an indirect way to access the WTD with methods that are within reach of the experimental state of the art. Namely, the dynamics of the full open quantum system is accessed from measurements of low-order correlation functions (typically second- or third-order). This is made possible with a cMPS formulation of the transport experiments as explained in the following section.

We illustrate this indirect path of accessing the WTD by considering real data obtained in the experiment of [27], where single electrons tunnel through a single-level quantum dot in the kHz regime. Both the current trace resolved in time and the two- and three-point correlation functions have been measured. The data allows us to demonstrate a very good agreement between the WTD deduced directly from the current trace and the WTD obtained via our reconstruction scheme based on the data of the correlation functions. This gives substance to our protocol based on cMPS to access the WTD with present technologies. We claim that this method remains valid in the GHz frequency range and for more complex systems such as a double quantum-dot coupled to two reservoirs—which would exhibit quantum coherence effects—and for quantum transport experiments with fermionic quantum gases.

2. Tomography of cMPS

In order to present a self-contained analysis, we start by reviewing the cMPS formulation of capturing a finite dimensional open quantum system [2] and the tomography procedure of reconstructing the relevant cMPS parameter matrices [11]. Consider an open quantum system (in cMPS terms the *ancillary system*) with dimension d (called *bond dimension* in that context) and interacting with one or more quantum fields that are described by field operators $\hat{\psi}_\alpha$ for different fields α. Its dynamics can in general be represented by different mathematical objects:

(a) The master equation in Lindblad form, which governs the evolution of the ancillary system described by its state vector $|\Psi\rangle$ defined on the Hilbert space $\mathcal{H}$ of dimension $d \times d$. The degrees of freedom of the coupled fields are traced out in this approach.

(b) The set of n-point correlation functions of the coupled fields.

(c) The full counting statistics of the field system, i.e. the complete set of cumulants of the probability distribution of transferred particles. The nth cumulant of the generating function is linked to the n moments of this distribution, which correspond to the n-point correlation function.

(d) The cMPS state vector $|\psi_{\mathrm{cMPS}}\rangle$, which we now introduce.

2.1. Reconstruction of cMPS from correlation functions

An intuitive way of establishing the cMPS state vector $|\psi_{\mathrm{cMPS}}\rangle$ consists in starting from the well-known Lindblad equation. This equation describes the evolution of the state ρ in time via the Liouvillian superoperator $\mathcal{L}$

$$\dot{\rho} = \mathcal{L}[\rho] = -\frac{\mathrm{i}}{\hbar}[K, \rho] - \frac{1}{2}\sum_{\alpha=1}^{p}\left(\left\{R_\alpha^\dagger R_\alpha, \rho\right\} - 2R_\alpha \rho R_\alpha^\dagger\right). \tag{1}$$

The first term relates to the free evolution via a Hamiltonian $K \in \mathbb{C}^{d\times d}$, while the last two terms describe the coupling to the environment (the according operator is known as the dissipator). The matrices $R_\alpha \in \mathbb{C}^{d\times d}$, $\alpha = 1,\ldots, p$, correspond to jump operators between the system and external quantum fields $\{\hat{\psi}_\alpha\}$. The matrices K and $\{R_\alpha\}$ completely characterize the evolution of the system.

Making use of the Choi–Jamiołkowski isomorphism [34] (which maps linear superoperators from $\mathcal{H}_1$ to $\mathcal{H}_2$ to linear operators acting on $\mathcal{H}_1 \otimes \mathcal{H}_2$) the state ρ is mapped to a state vector $|\rho\rangle$ and the Liouvillian $\mathcal{L}$ to the

matrix T [1, 2] with

$$T = Q^* \otimes \mathbb{1} + \mathbb{1} \otimes Q + \sum_\alpha R_\alpha^* \otimes R_\alpha. \tag{2}$$

The matrix $T \in \mathbb{C}^{d^2 \times d^2}$ is known as the *transfer matrix* and the matrix Q is defined as

$$Q = -\mathrm{i}K - \frac{1}{2}\sum_\alpha R_\alpha^\dagger R_\alpha. \tag{3}$$

Formally, the isomorphism introduced above is defined by the following relations for an operator and the product of operators

$$\begin{aligned} \rho &\mapsto |\rho\rangle, \\ A^\dagger \rho B &\mapsto \left(A^* \otimes B\right)|\rho\rangle. \end{aligned} \tag{4}$$

Being closely connected to K and $\{R_\alpha\}$ introduced above, the knowledge of the matrix and T and its components provides access to the dynamics of the open quantum system, and allows to directly derive the according Lindblad equation.

The (translationally invariant) cMPS state vector $|\Psi_{\mathrm{cMPS}}\rangle$ on the interval $[0,\ L]$ is defined in terms of the matrices $Q,\ \{R_\alpha\}$ and the field operators $\hat{\psi}_\alpha^\dagger$ by

$$\left|\psi_{\mathrm{cMPS}}\right\rangle = \mathrm{Tr}_{\mathrm{anc}}\left[\mathcal{P}\exp\int_0^L \mathrm{d}x\left(Q \otimes \hat{\mathbb{1}} + \sum_\alpha R_\alpha \otimes \hat{\psi}_\alpha^\dagger(x)\right)\right]|\Omega\rangle. \tag{5}$$

This expression is related to the path ordered exponential that arises when integrating the Lindblad equation. The embedding of the cMPS state vector $|\psi_{\mathrm{cMPS}}\rangle$ into Fock space becomes clear when expanding the path ordered exponential $\mathcal{P}$ exp. For more details, we refer to [3] where the authors formulate the cMPS in different representations such as the Fock space and a path integral formulation. After integration, the ancillary system is traced out via $\mathrm{Tr}_{\mathrm{anc}}$ and the resulting term is applied to the vacuum state vector $|\Omega\rangle$, where $\hat{\psi}_\alpha|\Omega\rangle = 0$ for each α.

Compared to the Lindblad equation, the main difference is that the degrees of freedom of the ancillary system are traced out such that its dynamics is mapped into the dynamics of the coupled quantum fields $\{\hat{\psi}_\alpha\}$. The evaluation of expectation values of field operators leads to expressions that only contain quantities from the ancillary system, and information about the ancillary system can be inferred from according field operator measurements. For the sake of clarity, we restrict ourselves to the case where a single coupled quantum field, denoted as $\hat{\psi}_\beta$, is measured.

The density-like correlation functions of the measured quantum field $\hat{\psi}_\beta$ then read

$$C_n(\mathbf{x}) = \left\langle \psi_{\mathrm{cMPS}}\right| \hat{n}\left(x_1\right) \ldots \hat{n}\left(x_n\right)\left|\psi_{\mathrm{cMPS}}\right\rangle, \tag{6}$$

where $\mathbf{x} \coloneqq (x_1, \ldots, x_n)$ and $\hat{n} \coloneqq \hat{\psi}_\beta^\dagger \hat{\psi}_\beta$. According to the calculus of expectation values in the cMPS setting [3], inserting equation (5) into equation (6) in the thermodynamic limit $L \to \infty$ leads to the expression

$$C_n(\mathbf{x}) = \lim_{L\to\infty} \mathrm{Tr}\left[\mathrm{e}^{D\left(L-x_n\right)} M \mathrm{e}^{D\left(x_n - x_{n-1}\right)} M \ldots M \mathrm{e}^{D\left(x_1 - 0\right)}\right]. \tag{7}$$

With D we denote the transfer matrix T—introduced in equation (2)—in its diagonal basis

$$D = X^{-1} T X, \tag{8}$$

where the columns of X represent the eigenvectors of T. Analogously, the matrix M denotes $R_\beta^* \otimes R_\beta$ in the diagonal basis of T

$$M = X^{-1}\left(R_\beta^* \otimes R_\beta\right)X. \tag{9}$$

Let us mention that the knowledge of X is in principle not necessary to reconstruct the matrices Q and R and hence the according Lindblad equation [11].

Specifically, the second- and third-order correlation functions take the form

$$C_2(x) = \mathrm{Tr}\left[\mathrm{e}^{D\infty} M \mathrm{e}^{Dx} M\right] = \sum_{j=1}^{d^2} M_{1,j} M_{j,1} \mathrm{e}^{\lambda_j x} \tag{10}$$

and

$$C_3(x,\, x') = \sum_{j,k=1}^{d^2} M_{1,k} M_{k,j} M_{j,1} \mathrm{e}^{\lambda_j x} \mathrm{e}^{\lambda_k (x'-x)}, \tag{11}$$

with $\{\lambda_j\}$ being the eigenvalues of T. Due to the translation invariance of the system, we can set $x_1 = 0$. The tomographic possibilities of the cMPS formalism can be understood from equations (10)–(11): if the products $\{M_{1,k}M_{k,j}M_{j,1}\}$ are known, we can—using gauge arguments [12]—require each of the matrix elements $\{M_{1,j}\}$ to be equal to one, which enables us to access each $M_{k,j}$ by dividing the appropriate terms:

$$\frac{M_{1,k}M_{k,j}M_{j,1}}{M_{1,1}M_{1,j}M_{j,1}} = M_{k,j}. \tag{12}$$

Both numerator and denominator appear as coefficients in C_3 and can be determined with spectral estimation procedures. This means that in principle we just need to analyse a three-point function in order to obtain the building elements M and D of arbitrary-order correlation functions.

This reconstruction scheme demonstrates the central role of the matrices M and D to derive the different equivalent objects that describe the dynamics of an open quantum system: the Lindblad equation, the set of n-point correlation functions, the full counting statistics of the number of transferred particles and the cMPS state vector. These matrices M and D can therefore be considered as the central quantities on which our reconstruction procedure is based; this is illustrated in figure 1.

2.2. Use of the thermodynamic limit

Intuitively, it is clear that the reconstruction of the matrices M and D should gain precision by increasing the number of correlation functions C_n on which the reconstruction scheme is based. The same statement is valid when increasing the size of the set of available counting probabilities P_n. But in general, experiments will only provide us measurements of low-order correlation functions, typically those of the second- and third-order [26, 27, 33]. *A priori*, this might render the reconstruction of the matrices M and D infeasible, but the work in [12] proved that this limitation can be circumvented by making use of the structure of the cMPS state vector combined with the thermodynamic limit.

For a given finite region I and a fixed bond dimension d, all expectation values can be computed from all correlation functions $C_n(\mathbf{x})$ taking values in the finite range I, $\mathbf{x} = (x_1,...,x_n) \subset I^{\times n}$. This contrasts with the situation of having access to correlation functions $C_n(\mathbf{x})$ for *arbitrary* values of $\mathbf{x} = (x_1,...,x_n)$, but for low n. Here, arbitrary values $\mathbf{x}$ imply the thermodynamic limit, i.e. the finite region I tends to infinity. Then indeed, low order correlation functions (typically C_1, C_2, C_3) are sufficient to reconstruct an arbitrary expectation value of an observable supported on I.

3. Reconstruction of cMPS from low-order counting probabilities

In this section, we extend the central role played by the matrices M and D for tomographic purposes by showing that they (more precisely: their equivalents $\mathcal{M}$ and $\mathcal{D}$) are also accessible from low-number detector-click statistics, i.e. the idle time probability density function P_0 and the density function P_1, which correspond to the detection of zero and one particle, respectively, within a certain time interval τ.

It is well-known that correlators and counting statistics are closely related. When assuming perfect detectors, the probability to observe n events in the time interval between t and $t+\tau$ is given [35] by the expression

$$P_n(t,\, t+\tau) = \frac{1}{n!}\sum_{m=n}^{\infty}\frac{(-1)^{m-n}}{(m-n)!}\int_t^{t+\tau} \mathrm{d}t_1 \cdots \int_t^{t+\tau} \mathrm{d}t_m\; C_m\Big(t_1,\, t_2,\, \ldots,\, t_m\Big), \tag{13}$$

where the correlation function C_m has been introduced in equation (6). For a translationally invariant system, we can without loss of generality set $t = 0$. Furthermore, when changing the integration bounds and performing the limit $L \to \infty$, we obtain

$$P_n(\tau) \coloneqq P_n(0,\, 0+\tau) = \int_0^{\tau} \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \cdots \int_0^{t_{n-1}} \mathrm{d}t_n\; \tilde{C}_n\Big(\tau,\, t_1,\, t_2,\, \ldots,\, t_n\Big), \tag{14}$$

with

$$\tilde{C}_n\Big(\tau,\, t_1,\, \ldots,\, t_n\Big) \coloneqq \mathrm{e}_1^T Z^{-1}\mathrm{e}^{\mathcal{D}t_n}\mathcal{M}_n \cdots \mathcal{M}_2\mathrm{e}^{\mathcal{D}\left(t_1-t_2\right)}\mathcal{M}_1\mathrm{e}^{\mathcal{D}\left(\tau-t_1\right)}Z\mathrm{e}_1, \tag{15}$$

the canonical unit vector e_1, the diagonal matrix $\mathcal{D}$ of $Q^* \otimes \mathbb{1} + \mathbb{1} \otimes Q$ with basis transformation matrix Y, $\mathcal{M}_j \coloneqq Y^{-1}(R_j^* \otimes R_j)Y$, and $Z \coloneqq Y^{-1}X$, where X diagonalizes T as defined in equation (8).With equations (13)–(15), the low-order counting probabilities $P_0(\tau)$ and $P_1(\tau)$ within a cMPS formulation are given by similar expressions to equations (10) and (11), namely

$$P_0(\tau) = \mathrm{e}_1^T Z^{-1}\mathrm{e}^{\mathcal{D}\tau}Z\mathrm{e}_1 = \sum_{j=1}^{d^2}\hat{z}_{1,j}z_{j,1}\mathrm{e}^{\mu_j\tau}, \tag{16}$$

$$P_1(\tau) = \sum_{j,k=1}^{d^2} \hat{z}_{1,j}\mathcal{M}_{j,k}z_{k,1}\left(\left(1-\delta_{j,k}\right)\frac{\mathrm{e}^{\mu_k\tau}-\mathrm{e}^{\mu_j\tau}}{\mu_k-\mu_j}+\delta_{j,k}\tau\mathrm{e}^{\mu_j\tau}\right), \tag{17}$$

with $\{\mu_j\}$ being the diagonal values of $\mathcal{D}$ and $\{z_{j,k}\}$ being the elements of the matrix Z (with inverse $Z^{-1} =: (\hat{z}_{j,k})$). See appendix for details.

As a first step, we can extract from P_0 and P_1 the coefficients $\{\hat{z}_{1,j}z_{j,1}\}$ and the eigenvalues $\{\mu_i\}$, which give rise to $\mathcal{D}$. The matrix elements of $\mathcal{M}$ can then in principle be determined using gauge arguments and under the assumption that the additive components of P_n are linearly independent. From $\mathcal{M}$ and $\mathcal{D}$, the cMPS matrices Q, R and K describing the dynamics of the open quantum system can be determined in a straightforward way (see appendix for details).

Let us comment on the feasibility of this reconstruction scheme with present technology. In order to measure P_0 and P_1, efficient single-particle detectors without *dark-counting* and *tiny dead-time* are necessary. *Dark-counting* leads to detector output pulses in the absence of any incident photons while the *dead-time* is the time interval after a detection event during which the detector cannot detect another particle. Although significant experimental efforts have been made in order to improve single-photon [36] and single-electron detectors [37, 38], the state-of-the-art for single-particle detection is not yet sufficient to perform a reliable measurement of P_1. For the moment, these experimental constraints make the reconstruction scheme based on P_1 only valid on a formal, mathematical level. In the light of the recent experimental progress towards the reliable detection of single particles, we believe that this idea will become relevant in the future.

4. Application to fermionic quantum transport experiments

Very recent works have successfully formulated experimental setups in cavity QED and ultra-cold Bose gases as well as the corresponding measurements in terms of cMPS [7, 9]. This allowed them to make predictions for higher-order correlation functions that are not accessible experimentally and to investigate the ground-state entanglement.

Here, we tackle the problem of formulating quantum transport experiments and the corresponding measurements (average charge current, charge noise) in cMPS terms To this end, we demonstrate that the field that is leaking out and is measured in a quantum transport experiment belongs to the cMPS variational class. We then provide an example to illustrate the equivalence between an Hamiltonian and a cMPS formulation by considering one of the simplest transport experiment, namely single electrons tunnelling through a single-level quantum dot. We derive the first-order and second-order correlation functions in cMPS terms, and show that we recover the well-known expression of the average current and charge noise, when writing the cMPS state equation (5) in terms of the parameters of the quantum system.

4.1. Quantum transport experiments in terms of cMPS

We now turn to a description of the physical setting under consideration. We assume here transport experiments, where single electrons transit through a scatterer coupled to fermionic reservoirs. The reservoirs, considered at equilibrium, are characterized by their chemical potential and their temperature via the Fermi distribution. The bias energy and the bias temperature between the different reservoirs will set the direction of the charge current. For the sake of simplicity, we restrict ourselves to two reservoirs, the source and the drain. This transport setting can be described by the Hamiltonian

$$\hat{H}_{\mathrm{T}} = \hat{H}_{\mathrm{sys}} + \hat{H}_{\mathrm{res}} + \hat{H}_{\mathrm{int}}, \tag{18}$$

where $\hat{H}_{\mathrm{sys}}$ relates to the quantum system under investigation, which acts as scatterer. It is characterized by discrete energy levels ε_i with occupation number operators given by $\hat{d}^{\dagger}_{i,\sigma}\hat{d}_{i,\sigma}$ ($\hat{d}_{i,\sigma}$ and $\hat{d}^{\dagger}_{i,\sigma}$ denote the fermionic annihilation and creation operators for an electron on the energy level i and spin degree of freedom $\sigma = \uparrow, \downarrow$). The Hamiltonian $\hat{H}_{\mathrm{res}}$ relates to the left and right reservoirs, and $\hat{H}_{\mathrm{int}}$ describes the interaction between the quantum system and the reservoirs,

$$\hat{H}_{\mathrm{res}} = \sum_{\alpha=L,R}\sum_{\sigma=\uparrow,\downarrow}\int_0^{E_\alpha} \mathrm{d}E\; E\; \hat{c}^{\dagger}_{\alpha,\sigma}(E)\,\hat{c}_{\alpha,\sigma}(E), \tag{19}$$

$$\hat{H}_{\mathrm{int}} = \sum_{\alpha=L,R}\sum_{i,\sigma}\int \mathrm{d}E\Big(t_{\alpha,i,\sigma}(E)\,\hat{d}_{i,\sigma}\otimes\hat{c}^{\dagger}_{\alpha,\sigma}(E) + \mathrm{h.c.}\Big). \tag{20}$$

The creation and annihilation operators of the reservoirs, $\hat{c}_\alpha$ and $\hat{c}^{\dagger}_\alpha$, satisfy the canonical anti-commutation relations and $\alpha = L,\ R$ denotes the left and right reservoirs, respectively. The amplitude $t_{\alpha,i,\sigma}$ sets the interaction between the quantum system and its environments.

In order to model a DC source, the energy levels in the left and right reservoirs are assumed to be densely filled up to the energies $E_F + eV$ and E_F, respectively. Here, E_F is the Fermi energy and V is the bias potential applied on the 'source' reservoir. At zero temperature, the bias energy eV enables uni-directional transport of electrons between the left and right reservoirs. It plays a similar role to the frequency bandwidth when, e.g., considering cavity QED setups, and fixes the energy domain over which electronic transport takes place.

With this assumption about the direction of propagation of the electrons (from left to right), we will see that equation (18) is equivalent to a generalized version of the cMPS Hamiltonian introduced in [1, 2],

$$\hat{H}_{\mathrm{cMPS}} = Q \otimes \hat{\mathbb{1}} + \left(R_{\mathrm{L}} \otimes \hat{\psi}_{\mathrm{L}} + R_{\mathrm{R}} \otimes \hat{\psi}_{\mathrm{R}}^{\dagger} \right), \tag{21}$$

where the matrices Q and $\{R_\alpha\}$ and the quantum fields $\{\hat{\psi}_\alpha\}$ have been introduced in section 2. The cMPS Hamiltonian for quantum transport experiment reflects the direction of the current: a fermionic excitation present on the left of the scatterer is annihilated at the scatterer as described by the quantum field $\hat{\psi}_{\mathrm{L}}$ (an electron jumps into the scatterer). Similarly, a fermionic excitation present on the right of the scatterer is created at the scatterer as described by the quantum field $\hat{\psi}_{\mathrm{R}}^{\dagger}$ (an electron jumps out of the scatterer). The case of a multi-terminal setup can be considered in a similar way. Showing that equations (18) and (21) are equivalent implicates that there is a fermionic quantum field leaking out of the scatterer to be measured and that it belongs to the cMPS variational class. Such a description of the transport experiment corresponds to a fermionic version of the input-output formalism of cavity-QED setups.

Using equation (20), the quantum field leaking out of the quantum system, $\hat{\psi}_{\mathrm{R}}(t)$, can be written in terms of the creation operator in the right reservoir $\hat{c}_{\mathrm{R}}$; the incoming quantum field can be written in a similar way in terms of the creation operator in the left reservoir $\hat{c}_{\mathrm{L}}$

$$\hat{\psi}_{\alpha,\sigma}^{\dagger}(t) = \int_E \mathrm{d}E \; \mathrm{e}^{-\mathrm{i}Et/\hbar} \hat{c}_{\alpha,\sigma}^{\dagger}(E), \qquad \alpha = \mathrm{L},\, \mathrm{R}. \tag{22}$$

The Fermi sea for the electrons is taken into account in the following way: on the right side of the scatterer, the quantum field satisfies $\hat{\psi}_{\mathrm{R}}(t)|E_F\rangle = 0$, where $|E_F\rangle$ denotes the state of the Fermi sea at energy E_F, whereas on the left side of the scatterer, $\hat{\psi}_{\mathrm{L}}(t)|E_F + eV\rangle = 0$, where the state vector $|E_F + eV\rangle$ defines the state of a Fermi sea at energy $E_F + eV$.

Assuming that the energy levels ε_i of the quantum system are well inside the bias energy window eV, we can rewrite the integration over the energy domain E as $\int_E \mathrm{d}E = \int_0^{eV} \mathrm{d}E$.

This assumption is the so-called *large-bias limit*, which is considered in order to derive the master equation corresponding to the tight-binding Hamiltonian. In quantum optics, it corresponds to a finite frequency bandwidth, which allows the use of the rotating wave approximation [9, 10]. In the following, we assume that the interaction amplitude is spin- and energy-independent within the interval $[E_F,\, E_F + eV]$: $t_{\alpha,i,\sigma}(E) = t_\alpha$. Let us remark that the demonstration remains valid with an interaction amplitude that depends on spin and energy. Importantly, no assumption about the coupling strength is required here.

In a rotating frame with respect to the energies of the reservoirs and after a Jordan–Wigner transformation using the definitions of the quantum fields $\hat{\psi}_{\mathrm{R,L}}$ given in equation (22), the Hamiltonian in equation (18) can be rewritten as

$$\hat{H}_{\mathrm{T}} = \hat{H}_{\mathrm{sys}} \otimes \hat{\mathbb{1}} + \sum_{\alpha=\mathrm{L,R}} \sum_{i,\sigma} \left(t_\alpha \hat{d}_i \otimes \psi_{\alpha,\sigma}^{\dagger}(t) + \mathrm{h.c.} \right). \tag{23}$$

Following quantum optics calculations—which remain valid in this case because $\hat{H}_{\mathrm{T}}$ is a transport version of the spin-boson model—we finally arrive at an effective non-Hermitian Hamiltonian

$$\hat{H}_{\mathrm{eff}} = \left(\hat{H}_{\mathrm{sys}} - \frac{\mathrm{i}\hbar}{2} \sum_{\alpha=\mathrm{L,R}} \sum_{i,\sigma} \Gamma_\alpha \hat{d}_{i,\sigma}^{\dagger} \hat{d}_{i,\sigma} \right) \otimes \mathbb{1} + \sum_{i,\sigma} \left(\sqrt{\Gamma_{\mathrm{R}}}\, \hat{d}_{i,\sigma} \otimes \hat{\psi}_{\mathrm{R},\sigma}^{\dagger}(t) + \sqrt{\Gamma_{\mathrm{L}}}\, \hat{d}_{i,\sigma}^{\dagger} \otimes \hat{\psi}_{\mathrm{L},\sigma}(t) \right) \tag{24}$$

with $t_\alpha := \sqrt{\Gamma_\alpha}$. Expressed in the eigenbasis of $\hat{H}_{\mathrm{sys}}$, the operators $\sqrt{\Gamma_{\mathrm{R}}}\, \hat{d}_{i,\sigma}$ and $\sqrt{\Gamma_{\mathrm{L}}}\, \hat{d}_{i,\sigma}^{\dagger}$ take the form of matrices labelled $R_{\mathrm{R},i,\sigma}$ and $R_{\mathrm{L},i,\sigma}$, respectively. The effective non-Hermitian Hamiltonian can then be rewritten in a compact form

$$\hat{H}_{\mathrm{eff}} = Q \otimes \hat{\mathbb{1}} + \sum_{i,\sigma} \left(R_{\mathrm{L},i,\sigma}^{\dagger} \otimes \hat{\psi}_{\mathrm{L},\sigma}(t) + R_{\mathrm{R},i,\sigma} \otimes \hat{\psi}_{\mathrm{R},\sigma}^{\dagger}(t) \right). \tag{25}$$

When comparing this effective Hamiltonian with equation (21), the identification of the matrix Q and the matrices $\{R_\alpha\}$ is direct. For spin-less fermions, the matrices R verify $R_{\alpha,i,\sigma}^2 = 0$ in order to satisfy the Pauli principle. Equation (25) demonstrates that transport settings can be adequately formulated within the cMPS framework. This result is important as it clears the way for applying methods from cMPS tomography to fermionic quantum transport experiments.

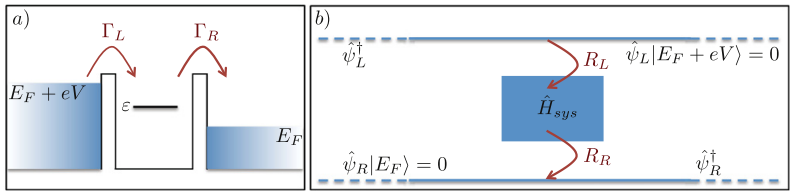

Figure 2. Scheme of a transport experiment through a single-level quantum dot. (a) The single-level quantum dot with energy ε is tunnel-coupled to two biased reservoirs with coupling strengths $\Gamma_{\rm L}$ and $\Gamma_{\rm R}$. Spin-less single-electron tunnelling events take place in the energy window eV above the Fermi sea with energy $E_{\rm F}$. (b) The same transport experiment from the open quantum system perspective for a cMPS formulation. The single-level dot is described by $\hat{H}_{\rm sys}$ and coupled to fermionic quantum fields $\hat{\psi}_{\rm L}$ and $\hat{\psi}_{\rm R}$. The coupling matrices $R_{\rm L/R}$ depend on the parameters $\Gamma_{\rm L/R}$, see equations (24)–(25). The transport direction fixed by the biased energy between the left and right reservoirs is ensured in the cMPS formulation by imposing $\hat{\psi}_{\rm L}|E_{\rm F} + eV\rangle = 0$ and $\hat{\psi}_{\rm R}|E_{\rm F}\rangle = 0$.

4.2. Single energy-level quantum dot

To illustrate the input–output formalism and the cMPS formulation of quantum transport experiments, we consider one of the simplest setups, namely a single energy-level quantum dot, without spin-degree of freedom, weakly coupled to two fermionic reservoirs. Even though this experiment is characterized by Markovian dynamics, this example is of particular interest for this work as it has been widely investigated experimentally. In section 5, we will use real data obtained in [27] for this setup to show that cMPS tomography allows us to access the electronic distribution of waiting times.

This simple transport experiment is sketched in figure 2 and the corresponding Hamiltonian reads

$$\hat{H}_T = \varepsilon\, \hat{d}^\dagger \hat{d} + \sum_{\alpha=L,R} \int \mathrm{d}E \left(\sqrt{\Gamma_\alpha}\, \hat{d} \otimes \hat{c}_\alpha^\dagger(E) + \mathrm{h.c.} \right) + \hat{H}_{\rm res}. \tag{26}$$

Assuming that we perform a measurement on the right of the scatterer, the first two correlation functions of the right quantum field $\hat{\psi}_{\rm R}(t)$ read in terms of cMPS matrices

$$\langle \hat{\psi}_{\rm R}^\dagger \hat{\psi}_{\rm R} \rangle = \lim_{L\to\infty} \mathrm{Tr}\left[\mathrm{e}^{TL} \left(R_{\rm R}^* \otimes R_{\rm R} \right) \right] \tag{27}$$

and

$$\langle \hat{\psi}_{\rm R}^\dagger(0)\, \hat{\psi}_{\rm R}^\dagger(\tau)\, \hat{\psi}_{\rm R}(\tau)\, \hat{\psi}_{\rm R}(0) \rangle = \lim_{L\to\infty} \mathrm{Tr}\left[\mathrm{e}^{T(L-\tau)} \left(R_{\rm R}^* \otimes R_{\rm R} \right) \mathrm{e}^{T\tau} \left(R_{\rm R}^* \otimes R_{\rm R} \right) \right]. \tag{28}$$

The matrices $R_{\rm R/L}$ correspond to the operators $\sqrt{\Gamma_{\rm R}}\,\hat{d}$ and $\sqrt{\Gamma_{\rm L}}\,\hat{d}^\dagger$ expressed in the eigenbasis of the single-level quantum dot, $\{|0\rangle, |1\rangle\}$ (empty and occupied state)

$$R_{\rm L} = \begin{pmatrix} 0 & 0 \\ \sqrt{\Gamma_{\rm L}} & 0 \end{pmatrix}, \quad R_{\rm R} = \begin{pmatrix} 0 & \sqrt{\Gamma_{\rm R}} \\ 0 & 0 \end{pmatrix}. \tag{29}$$

Inserting these expressions into equation (27), we recover the well-known expression for the steady-state current of a single-level QD coupled to biased reservoirs [39, 40]

$$\langle \hat{\psi}_{\rm R}^\dagger \hat{\psi}_{\rm R} \rangle = \frac{\Gamma_{\rm L}\Gamma_{\rm R}}{\Gamma_{\rm R} + \Gamma_{\rm L}} =: \langle \hat{I} \rangle_{\rm ss}. \tag{30}$$

Furthermore, we can derive the noise spectrum from equation (28) via the MacDonald formula [41–43]

$$S(\omega) = 2\langle \hat{I} \rangle_{\rm ss} \left(1 - \frac{2\Gamma_{\rm L}\Gamma_{\rm R}}{\left(\Gamma_{\rm L} + \Gamma_{\rm R}\right)^2 + \omega^2} \right). \tag{31}$$

This example aims at bridging the gap between a more traditional Hamiltonian and the cMPS formulation, which allows to write these well-known expressions in terms of the parameter matrices Q, T, and $\{R_\alpha\}$.

5. Reconstruction of waiting time statistics

In this section, we address the problem of accessing the distribution of waiting times in electronic transport experiments. As mentioned in the introduction, a direct measurement of the WTD in the GHz range is not yet possible due to the lack of single-particle detectors with sufficient accuracy at those frequencies. Here, we propose to reconstruct the WTD based on the experimental measurements of low-order correlation functions.

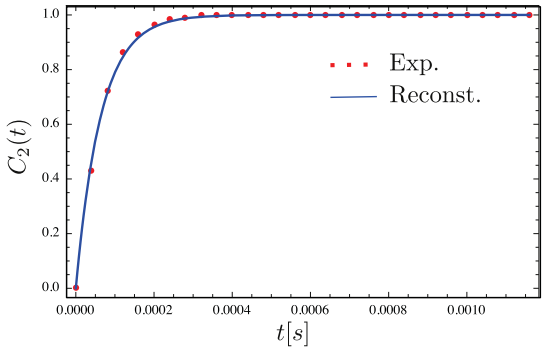

Figure 3. Two-point correlation function C_2 for single fermions tunnelling through a single-level quantum dot with data from [27] (red-dotted curve). The blue curve, obtained from the reconstructed values of the parameters $\Gamma_{L,R}$ using a cMPS formulation of the quantum experiment, agrees well with the experimental measurement of C_2. The deviation for small times is due to experimental limitations in the time bin with respect to which the current trace is resolved.

The reconstruction is carried out using the cMPS framework presented in section 2 and the formulation of transport experiments in terms of cMPS as exposed in section 4.

5.1. Definitions

The statistics of waiting times can be expressed in terms of the probability density function P_0, which—as a function of τ—expresses the probability of having detected zero particles in the interval $[0,\ \tau]$. In terms of P_0, the WTD has first been derived in the context of quantum transport experiments in [20],

$$\mathcal{W}(\tau) = \langle\tau\rangle \frac{\partial^2 P_0(\tau)}{\partial \tau^2}. \tag{32}$$

Here, $\langle\tau\rangle$ denotes the mean waiting time. Inserting $P_0(\tau)$ in cMPS terms (equation (16)), we arrive at an expression for $\mathcal{W}$ in terms of the cMPS matrices D, $\mathcal{D}$ and Z defined in equations (8) and (15),

$$\mathcal{W}(\tau) = \frac{1}{c} e_1^T \left(D^2 Z^{-1} - 2DZ^{-1}\mathcal{D} + Z^{-1}\mathcal{D}^2 \right) e^{\mathcal{D}\tau} Z e_1. \tag{33}$$

The normalization factor $c > 0$ ensures that $\int_0^\infty \mathcal{W}(\tau)\mathrm{d}\tau = 1$. Equation (33) allows us to access the WTD from the measurements of the low-order correlation functions only via the use of the cMPS framework to reconstruct the cMPS matrices D, $\mathcal{D}$ and Z.

5.2. Results based on experimental data

We demonstrate our novel approach to derive the WTD from the measurement of correlation functions using experimental data obtained in [27] for spinless electrons tunnelling through a single-level quantum dot. This system is also known as a single-electron transistor at the nanoscale and has been discussed in section 4.2. The experiment in [27] has been carried out in the kHZ frequency range, where a time-resolved measurement of the current trace is possible. Although all the statistics—including correlation functions of arbitrary order as well as the WTD—can directly be computed from this time-resolved current trace, this experiment provides an ideal test-bed for our proposal. We can compare the WTD obtained from our reconstruction scheme based on cMPS with the WTD directly deduced from the experimental current trace.

Due to the simplicity of the setup, our proposed method to access the WTD only requires the two-point function C_2. This one can directly be derived from the experimental spike train I (the time-resolved current trace) and is shown in figure 3 (red dots). The rates $\Gamma_L = 13.23$ kHz and $\Gamma_R = 4.81$ kHz have been determined experimentally and the corresponding C_2-function agrees very well with the analytical expression when the detector rate is taken into account [27]

$$C_2(\tau) = \frac{\Gamma_L \Gamma_R}{\Gamma_L + \Gamma_R} \left(1 - e^{-\tau(\Gamma_L + \Gamma_R)} \right). \tag{34}$$

In our reconstruction scheme, the quantity $\Gamma_L + \Gamma_R$ can be determined from the current spike train autocorrelation function $I \star I$ by least squares methods or spectral estimation procedures analogous to the procedure described in [11]. By requiring $\Gamma_L > \Gamma_R$ and using the expression of the steady-state current (see

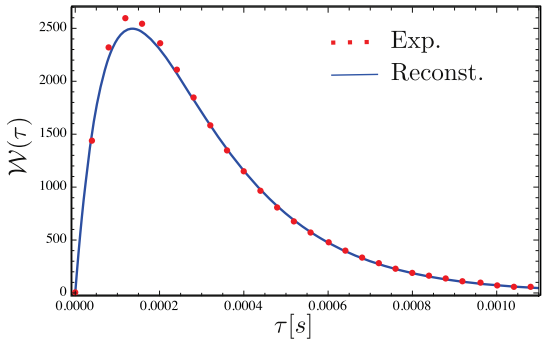

Figure 4. WTD obtained from state-of-the-art experimental measurements with data from [27]. The reconstructed WTD using equation (33) is shown in blue. It matches the WTD obtained directly from the time resolved experimental current trace (red dots) well. The deviation is due to the finite-sized time bin corresponding to the resolution of the current trace. A more accurate reconstruction of the WTD is expected by increasing the time resolution of the current trace or of the measurement of C_2.

equation (30)), Γ_L and Γ_R can be uniquely identified. The reconstructed values for the rates are

$$\Gamma_L^{\text{recon}} = 10.80\ \text{kHz}, \tag{35}$$

$$\Gamma_R^{\text{recon}} = 4.76\ \text{kHz}. \tag{36}$$

The differences to the values from [27] are well within the range we would expect, regarding the time-resolution in the spike train data. The curve plotted from these reconstructed values of the parameters $\Gamma_{L,R}$ is shown in figure 3 in blue. The slight deviation between the experimental points and this reconstructed C_2-function is due to the discretization of the counting time intervals used in the experiment: the size of each time bin is not much smaller than the time scale on which C_2 changes mostly. This leads to an error in the estimation of the damping factor $\Gamma_L + \Gamma_R$ and explains the difference of the blue and the red dotted curves. Naturally, one could expect a more accurate reconstruction of the parameters Γ_L and Γ_R when increasing the time resolution of the current trace or of the measurement of C_2.

From Γ_L and Γ_R, the corresponding cMPS matrices R_L and R_R can be constructed, as well as the matrices M and D. In this simple case, we did not need to employ the whole reconstruction procedure from [11]. Indeed, it is clear from equation (34) that only two out of the four parameters that characterize the system appear: C_n only depends on the tunnelling rates Γ_L and Γ_R—the eigenenergies 0 and ε of $\hat{H}_{\text{sys}}$ do not contribute[5]. This will in general not be the case.

The matrices R_L and R_R give access to the matrices D and $\mathcal{D}$ by direct computation. Inserting the latter into equation (33), the WTD can be reconstructed and the result is plotted in figure 4 (blue curve). In order to build confidence in our procedure, we compare this result with the experimentally accessible WTD (red dots). Let us recall that the transport rate is in the kHz range, hence the WTD can directly be extracted from the current spike train I: by sorting, counting all (discrete) waiting times between two consecutive incidents, and subsequently normalizing the resulting histogram, one obtains the red-dotted WTD in figure 4. The slight deviation between the WTD reconstructed via our proposal and the experimental one is again due to the discretization of the counting time intervals. One could expect a more accurate reconstruction of the WTD when increasing the time resolution of the current trace.

The WTD in figure 4 shows elementary transport properties of single independent fermions that cannot tunnel at the same time through the single-level quantum dot, which is consistent with the fact that $\mathcal{W}(\tau) \to 0$ for $\tau \to 0$. It is important to emphasize that it is the first time a WTD is extracted from experimental data, therefore bridging the gap between theoretical predictions and experiments. The good agreement of the two curves demonstrates the potential of our cMPS-based reconstruction procedure to access the WTD from the measurements of low-order correlation functions. This opens the route to access the WTDs in the high-frequency domain from low-order correlation-functions measurements.

[5] The eigenvectors of T only depend on Γ_L and Γ_R, this applies to M_L, M_R and all residues as well. Accordingly, the two non-real poles are the only quantities that depend on ϵ, however, only the residues connected to the two real poles do not vanish. When adding off-diagonal elements to K, the terms mix and a dependency on K arises.

6. Conclusion

In this work, we have taken an approach motivated by cMPS to perform tomographic reconstructions of quantum transport experiments. On a formal level, we have extended this formalism to perform a reconstruction of unknown dissipative processes based on the knowledge of low-order counting probabilities. We then demonstrated that cMPS is an adequate formalism to describe quantum transport experiments based on tight-binding Hamiltonians.

This work advocates a paradigm change in the analysis of transport experiments. The traditional method is to make explicit use of a model to put the estimated quantities into context, a model that may or may not precisely reflect the physical situation at hand. The cMPS approach is to not assume the form of the model, with the exception that the quantum state can be described by a cMPS. Such an approach is of particular interest as it opens the way to the access of quantities that are not measurable experimentally with current technologies, high-order correlation functions and distributions of waiting times.

To convincingly demonstrate the functioning of cMPS tomographic tools applied to quantum transport experiments, we presented a simple example that consists of electrons tunnelling through a single-level quantum dot. Making use of experimental data, we showed that we could successfully reconstruct the distribution of waiting times from the measurement of the two-point correlation function only. This work constitutes therefore a significant step towards accessing the waiting time distribution in the quantum regime experimentally, a challenge present for several years now. Importantly, the application of our reconstruction procedure goes beyond the interest in WTD: it also provides an access to higher-order correlation functions, which are key quantities to better understand interacting quantum systems.

In subsequent research, it would be desirable to further flesh out the statistical aspects of the problem. After all, the description in terms of cMPS constitutes a statistical model. It would constitute an exciting enterprise in its own right to identify region estimators that provide efficiently computable and reliable confidence regions [44] when considering the problem as a statistical estimation problem, related to the framework put forth in [45–47]. We hope that the present work inspires such further studies of transport problems in the mindset of quantum tomography.

Acknowledgments

We acknowledge the group of R Haug, and especially N Ubbelohde, for sharing with us the experimental data. We also acknowledge valuable comments and discussions with M Albert, D Dasenbrook and C Flindt. GH acknowledges support from the A von Humboldt foundation, from the Swiss NCCR QSIT and from the ERC grant MesoQMC, AS from the Studienstiftung des deutschen Volkes, JE from the BMBF (Q com), the EU (RAQUEL, SIQS, AQuS, COST), the DFG (SPP 1798), and the ERC grant TAQ.

Appendix. Reconstruction method from low-order counting probabilities

In this appendix, we provide further technical details on the reconstruction scheme based on the measurement of low-order counting probabilities, P_0 and P_1. The goal is to access the central cMPS parameter matrices $\mathcal{M}$ and $\mathcal{D}$. We refer to figure 1 for a general view of the reconstructible items. We start from equation (13) in the main text. By changing the integration bounds, we obtain equation (14),

$$P_n(\tau) = \int_0^{\tau} \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \cdots \int_0^{t_{n-1}} \mathrm{d}t_n \; \tilde{C}_n\big(\tau, t_1, t_2, \ldots, t_n\big),$$

where the integrand C_n is altered to

$$\tilde{C}_n\big(\tau, t_1, \ldots, t_n\big) \coloneqq \lim_{L\to\infty} \mathrm{Tr}\Big[\mathrm{e}^{T(L-t_n-\tau)}\mathrm{e}^{St_n}R_n^* \otimes R_n \cdots R_2^* \otimes R_2\mathrm{e}^{S(t_1-t_2)}R_1^* \otimes R_1\mathrm{e}^{S(\tau-t_1)}\Big]. \tag{A.1}$$

Note that in contrast to equation (7) where the propagating matrix is the transfer matrix T defined by equation (2) (or equivalently its diagonal representation D), the propagating matrix in the exponential terms between two measurement points now is the matrix S, which is defined by

$$S \coloneqq Q^* \otimes \mathbb{1} + \mathbb{1} \otimes Q = T - \sum_j R_j^* \otimes R_j. \tag{A.2}$$

We can further simplify equation (A.1) by performing the thermodynamic limit $L \to \infty$. The spectrum of T for a generic system consists of complex values with negative real part and only one eigenvalue being equal to zero. When taking the limit $L \to \infty$, all eigenvalue contributions to $\mathrm{e}^{T(L-\tau)}$ vanish, except the one corresponding to the zero eigenvalue. Hence

$$\lim_{L\to\infty} \mathrm{e}^{T(L-\tau)} = Xe_1 e_1^T X^{-1}, \tag{A.3}$$

with the first canonical unit vector denoted by e_1 and the basis transformation matrix X to the diagonal basis of T. Similar to D and T, we define the matrix $\mathcal{D}$ as the diagonal matrix of S with basis transformation matrix Y:

$$S = Y\mathcal{D}Y^{-1}. \tag{A.4}$$

By defining the matrices $\mathcal{M}_j := Y^{-1}(R_j^* \otimes R_j)\,Y$ for $j = 1,\ldots,n$, and setting $Z := (z_{j,k}) := Y^{-1}X$ (with inverse $Z^{-1} =: (\hat{z}_{j,k})$), we arrive at equation (15),

$$\tilde{C}_n\left(\tau, t_1, \ldots, t_n\right) = e_1^T\, Z^{-1}\, \mathrm{e}^{\mathcal{D}t_n}\, \mathcal{M}_n \cdots \mathrm{e}^{\mathcal{D}(t_2-t_3)}\, \mathcal{M}_2\, \mathrm{e}^{\mathcal{D}(t_1-t_2)}\, \mathcal{M}_1\, \mathrm{e}^{\mathcal{D}(\tau-t_1)}\, Ze_1.$$

Equations (15) and (7) have a close structural resemblance: the matrices M and $\mathcal{M}$ are *similar* in the linear algebra sense, i.e., there exists a basis transformation from $\mathcal{M}$ to M. The matrices D and $\mathcal{D}$ are the diagonal matrices of the transfer matrix T and the matrix S, respectively. It is straightforward to transform M and D into $\mathcal{M}$ and $\mathcal{D}$ and vice versa: by subtracting M from D, we obtain S (up to similarity/basis transformation), whose diagonal matrix is $\mathcal{D}$. Applying the same basis transformation (from $(D - M)$ to $\mathcal{D}$) to the matrix M results in the matrix $\mathcal{M}$.

For $n = 0$, the counting probability function then reads

$$P_0(\tau) = e_1^T\, X^{-1}Y\, \mathrm{e}^{D_S\tau}\, Y^{-1}X\, e_1, \tag{A.5}$$

which can be rewritten as a sum of complex exponential terms, with $\{\mu_j\}$ being the eigenvalues of S as

$$P_0(\tau) = \sum_{j=1}^{d^2} \hat{z}_{1,j} z_{j,1}\, \mathrm{e}^{\mu_j\tau}. \tag{A.6}$$

This expression corresponds to the analogue of equation (10) in the main text. Since S is by definition a *Kronecker sum* of Q^* and Q with eigenvalues $\{q_j^*\}$ and $\{q_j\}$ respectively, the spectrum of S consists of the sums $q_j^* + q_k$ with $j, k = 1,\ldots,d$. It is closed under complex conjugation (for each element of the set its complex conjugate is also element of the set), as well as the coefficient set $\{\hat{z}_{1,j} z_{j,1}\}$. This ensures that P_0 is real-valued. Being related to Q (which consists of a skew-hermitian matrix (with imaginary spectrum) and negative definite matrices), we have that $\mathrm{Re}\,\mu_j < 0$ for each j, such that all summands vanish sufficiently fast and P_0 is normalizable. Furthermore, the dominance of the damping factors over the oscillatory components ensures the positivity of P_0 (in particular, the μ_j with the least damping is always real-valued). Analogously, for $P_1(\tau)$ we obtain

$$\sum_{j,k=1}^{d^2} \hat{z}_{1,j}\mathcal{M}_{j,k} z_{k,1} \left(\left(1 - \delta_{j,k}\right) \frac{\mathrm{e}^{\mu_k\tau} - \mathrm{e}^{\mu_j\tau}}{\mu_k - \mu_j} + \delta_{j,k}\tau \mathrm{e}^{\mu_j\tau} \right) \tag{A.7}$$

with the Kronecker delta $\delta_{j,k}$,

$$f(\tau) = \sum_{m,n} c_{m,n}\tau^m \mathrm{e}^{\mu_n\tau}. \tag{A.8}$$

Assuming that the terms $\tau^m \mathrm{e}^{\mu_n\tau}$ are linearly independent, in principle one can always single out these contributions as well as their corresponding prefactors $c_{m,n}$. This gives us the chance to extract the coefficients $\{\hat{z}_{1,j} z_{j,1}\}$ and the eigenvalues $\{\mu_i\}$ from P_0, provided that no coefficient is identical to zero. Rearranging the values $\{\mu_i\}$ to a diagonal matrix in Kronecker sum form results in the matrix $\mathcal{D}$. One should note, however, that efficient spectral recovery algorithms like the *matrix pencil method* do not straightforwardly work for functions such as P_n, $n \geqslant 2$, where the exponential functions are multiplied with powers of τ.

In order to reconstruct the elements of the matrix $\mathcal{M}$ together with the off-diagonal elements of Z, we use a gauge argument: All probability functions P_n are invariant under scaling and permutation of the eigenvectors in the matrices X and Y (except for the eigenvector of T corresponding to eigenvalue zero). This allows us to require all but one $\hat{z}_{1,j}$ to be equal to one, and immediately obtain the according number $z_{j,1}$. The remaining coefficient can then be determined via the normalization constraint

$$\sum_{j=1}^{d^2} \hat{z}_{1,j} z_{j,1} = 1, \tag{A.9}$$

so that all $z_{j,1}$ are known. This can be used to obtain the diagonal elements $\mathcal{M}_{j,j}$ from $\mathcal{M}$. For the remaining matrix elements, only the symmetric elements $\mathcal{M}_{j,k} + \mathcal{M}_{k,j}$ (but not their constituents) are directly accessible since

$$\sum_{\substack{j,k=1 \\ j\neq k}}^{d^2} \hat{z}_{1,j}\mathcal{M}_{j,k} z_{k,1} \frac{\mathrm{e}^{\mu_k\tau} - \mathrm{e}^{\mu_j\tau}}{\mu_k - \mu_j} = \sum_{j<k}^{d^2} \left(\hat{z}_{1,j}\mathcal{M}_{j,k} z_{k,1} + \hat{z}_{1,k}\mathcal{M}_{k,j} z_{j,1} \right) \frac{\mathrm{e}^{\mu_k\tau} - \mathrm{e}^{\mu_j\tau}}{\mu_k - \mu_j}. \tag{A.10}$$

However, this does not constitute a limitation for the reconstruction of the matrices Q and R of the ancillary system. To this end, we make use of the inner structure of $\mathcal{M}$. The diagonal matrix $\mathcal{D}$ with eigenvalues μ_j can be

reordered such that it has the form $\mathcal{D} = D_Q^* \otimes \mathbb{1} + \mathbb{1} \otimes D_Q$ with diagonal D_Q consisting of the eigenvalues of Q. Reordering the eigenvectors in Y accordingly, we can assume that the matrix Y and hence the matrix $\mathcal{M}$ have the form of a Kronecker product

$$\mathcal{M} = R_{\mathrm{rec}}^* \otimes R_{\mathrm{rec}}. \tag{A.11}$$

Here $R_{\mathrm{rec}} = (r_{j,k}) \in \mathbb{C}^{d\times d}$ is in general not diagonal. The symmetrized components of $\mathcal{M}$ can then be written as $r_{j,k}^* r_{l,m} + r_{k,j}^* r_{m,l}$ and the constituents r_j can be determined (up to a phase factor) by equating them with the coefficients in equation (A.10). The according equation system can then be solved.

The important point is that R_{rec} and $Q_{\mathrm{rec}} := D_Q$ are valid cMPS parameter matrices in the same gauge and hence are sufficient for reconstruction with the same argument as in [11, III.E]. Let us note that concrete values of the basis transformation matrices X and Y are in fact never used or needed in the reconstruction procedure. From R_{rec} and Q_{rec}, we can compute all quantities we need to establish the correlation and counting probability functions, in particular $\mathcal{M}$ and $\mathcal{D}$. Regauging R_{rec} and Q_{rec} such that the orthonormalization condition [1] is fulfilled, yields a reconstruction of the free Hamiltonian K_{rec} of the ancillary system.

References

[1] Verstraete F and Cirac J I 2010 *Phys. Rev. Lett.* **104** 190405
[2] Osborne T J, Eisert J and Verstraete F 2010 *Phys. Rev. Lett.* **105** 260401
[3] Haegeman J, Cirac J I, Osborne T J and Verstraete F 2013 *Phys. Rev.* B **88** 085118
[4] Schollwöck U 2011 *Ann. Phys.* **326** 96
[5] Draxler D, Haegeman J, Osborne T J, Stojevic V, Vanderstraeten L and Verstraete F 2013 *Phys. Rev. Lett.* **111** 020402
[6] Quijandría F, García-Ripoll J J and Zueco D 2014 *Phys. Rev.* B **90** 235142
[7] Steffens A, Friesdorf M, Langen T, Rauer B, Schweigler T, Hübener R, Schmiedmayer J, Riofrío C A and Eisert J 2015 *Nat. Commun.* **6** 7663
[8] Sangwoo K S, Chung S and Bolech C J 2015 *Phys. Rev.* B **91** 121108(R)
[9] Barrett S, Hammerer K, Harrison S, Northup T E and Osborne T J 2013 *Phys. Rev. Lett.* **110** 090501
[10] Gardiner C W and Zoller P 2004 *Quantum Noise* (Berlin: Springer)
[11] Steffens A, Riofrío C, Hübener R and Eisert J 2014 *New J. Phys.* **16** 123010
[12] Hübener R, Mari A and Eisert J 2013 *Phys. Rev. Lett.* **110** 040401
[13] Bruderer M, Contreras-Pulido L D, Thaller M, Sironi L, Obreschkow D and Plenio M B 2014 *New J. Phys.* **16** 033030
[14] Kiukas J, Guta M, Lesanovsky I and Garrahan J P 2015 *Phys. Rev.* E **92** 012132
[15] Cramer M, Plenio M B, Flammia S T, Somma R, Gross D, Bartlett S D, Landon-Cardinal O, Poulin D and Liu Y-K 2010 *Nat. Commun.* **1** 149
[16] Brantut J-P, Meineke J, Stadler D, Krinner S and Esslinger T 2012 *Science* **337** 1069
[17] Brantut J P, Grenier C, Meineke J, Stadler D, Krinner S, Kollath C, Esslinger T and Georges A 2013 *Science* **342** 713
[18] Brandes T 2008 *Ann. Phys.* **17** 477
[19] Albert M, Flindt C and Büttiker M 2011 *Phys. Rev. Lett.* **107** 086805
[20] Albert M, Haack G, Flindt C and Büttiker M 2012 *Phys. Rev. Lett.* **108** 186806
[21] Thomas K H and Flindt C 2013 *Phys. Rev.* B **87** 121405(R)
[22] Rajabi L, Pötl C and Governale M 2013 *Phys. Rev. Lett.* **111** 067002
[23] Dasenbrook D, Flindt C and Büttiker M 2014 *Phys. Rev. Lett.* **112** 146801
[24] Albert M and Devillard P 2014 *Phys. Rev.* B **90** 035431
[25] Tang G-M, Xu F and Wang J 2014 *Phys. Rev.* B **89** 205310
[26] Gustavsson S, Leturcq R, Simovic B, Schleser R, Ihn T, Studerus P, Ensslin K, Driscoll D C and Gossard A C 2006 *Phys. Rev. Lett.* **96** 076605
[27] Ubbelohde N, Fricke C, Flindt C, Hohls F and Haug R J 2012 *Nat. Commun.* **3** 612
[28] Fève G, Mahé A, Berroir J-M, Kontos T, Plaçais B, Glattli C, Cavanna A, Etienne B and Jin Y 2007 *Science* **316** 1169
[29] Blumenthal M D, Kaestner B, Li L, Giblin S, Janssen T J B M, Pepper M, Anderson D, Jones G and Ritchie D A 2007 *Nat. Phys.* **3** 343
[30] Fujiwara A, Nishiguchi K and Ono Y 2008 *Appl. Phys. Lett.* **92** 042102
[31] Maire N, Hohls F, Kaestner B, Pierz K, Schumacher H W and Haug R J 2008 *Appl. Phys. Lett.* **92**
[32] Dubois J, Jullien T, Portier F, Roche P, Cavanna A, Jin Y, Wegscheider W, Roulleau P and Glattli D C 2013 *Nature* **502** 659
[33] Gabelli J, Spietz L, Aumentado J and Reulet B 2013 *New J. Phys.* **15** 113045
[34] Jamiolkowski A 1972 *Rep. Math. Phys.* **3** 275
[35] Kelley P L and Kleiner W H 1964 *Phys. Rev.* **136** 316
[36] Eisaman M D, Fan J, Migdall A and Polyakov S V 2011 *Rev. Sci. Instrum.* **82**
[37] Thalineau R, Wieck A, Bäuerle C and Meunier T 2014 arXiv:1403.7770
[38] Gasparinetti S, Viisanen K L, Saira O-P, Faivre T, Arzeo M, Meschke M and Pekola J P 2015 *Phys. Rev. Appl.* **3** 014007
[39] Stoof T H and Nazarov Y V 1995 *Phys. Rev.* B **53** 1050
[40] Blanter Y M and Büttiker M 2000 *Phys. Rep.* **336** 1
[41] MacDonald D K C 1949 *Rep. Prog. Phys.* **12** 56
[42] Flindt C, Novotny T and Jauho A-P 2005 *Phys.* E **29** 411
[43] Lambert N, Aguado R and Brandes T 2007 *Phys. Rev.* B **75** 045340
[44] Shao J 2003 *Mathematical Statistics* (Berlin: Springer)
[45] Blume-Kohout R 2012 arXiv:1202.5270
[46] Christandl M and Renner R 2012 *Phys. Rev. Lett.* **109** 120403
[47] Carpentier A, Eisert J, Gross D and Nickl R 2015 (arXiv:1504.03234)

A.2 Quantum singular value decomposition

An indispensable building block for many *quantum algorithms* (see section 4) is *quantum phase estimation*: Given a unitary operator $U = \mathrm{e}^{-\mathrm{i}H\Delta t}$, exponentiating the Hermitian H, and an eigenvector $|\psi_j\rangle$, it provides an estimate for the corresponding eigenvalue $\mathrm{e}^{-\mathrm{i}2\pi\varphi_j}$. More generally, using an initial state $|\chi\rangle$ with eigenvector overlaps $\langle\psi_j|\chi\rangle$, the operation

$$|\chi\rangle\,|0,\ldots,0\rangle_{\mathrm{A}} \;\mapsto\; \sum_j \langle\psi_j|\chi\rangle\,|\psi_j\rangle\,|\mathrm{e}^{-\mathrm{i}2\pi\tilde{\varphi}_j}\rangle_{\mathrm{A}} \tag{32}$$

is performed. $|\cdot\rangle_{\mathrm{A}}$ is a quantum register consisting of m qubits, each of which are initially in the state $|0\rangle$, and ultimately provide a binary representation of $\mathrm{e}^{-\mathrm{i}2\pi\varphi_j}$ with m binary digits. Each state $|\mathrm{e}^{-\mathrm{i}2\pi\tilde{\varphi}_j}\rangle_{\mathrm{A}}$ is a product state and can be read out efficiently. The operation makes use of *Hadamard* quantum gates and the (inverse) quantum Fourier transform; the latter involves the successive controlled application of U on $|\chi\rangle$. For this, U needs to be efficiently simulatable, which is generally not the case for an arbitrary unitary. Simulating $U = \mathrm{e}^{-\mathrm{i}H\Delta t}$ was shown to be efficient first for Hamiltonians with local interactions[21]. More classes of Hamiltonians followed, including sparse Hamiltonians[186] and positive-semidefinite, non-sparse, low-rank matrices[171].

For practical quantum algorithms, however, it is also desirable to exponentiate general, non-positive, even non-quadratic matrices A. Hamiltonians of the form

$$\widetilde{H} \coloneqq \begin{bmatrix} 0 & A \\ A^\dagger & 0 \end{bmatrix}, \tag{33}$$

as in Ref. [166], there however for sparse matrices, have positive and negative eigenvalues by construction and hence cannot simply be used in the algorithm in Ref. [171] without losing relative phase information of the eigenspaces. The following publication [6][25] helps closing this gap by providing a prescription for efficiently exponentiating Hamiltonians of the form (33) with non-sparse, low-rank matrices A, which need not be quadratic. The eigenvectors of $\widetilde{H}$ comprise the singular vectors of A and the eigenvalues the corresponding singular values. Together with the preservation of the phase relations between the singular spaces, this paves the way for a complete quantum analogue of the singular value decomposition with all its entailing applications. Specifically discussed in the publication is the *Procrustes problem* of finding an *isometry* (a matrix with orthonormal columns) that is in a least squares sense closest to a given linear map.

As an example for a potential embedding into a larger quantum algorithm, note that the SVD also plays a central role in compressed sensing: A classical fast iterative algorithm that solves the matrix completion problem Eq. (13) and also the compressive quantum state tomography problem

[25] Patrick Rebentrost, Adrian Steffens, and Seth Lloyd,"Quantum singular value decomposition of non-sparse low-rank matrices", *ArXiv e-prints* 1607.05404, 2016 (arxiv.org/abs/1607.05404). After submission of this thesis published as Patrick Rebentrost, Adrian Steffens, Iman Marvian, and Seth Lloyd, *Physical Review A* 97:012327, 2018 (© 2018 American Physical Society).

Eq. (15) (see section 2.2) alternates between fitting the ansatz matrix to the constraints and truncating its singular spectrum (*singular value thresholding*).[82] The quantum SVD could be used as a subroutine in a quantum matrix completion algorithm, interfacing with a quantum version of the fitting operation, and is expected to be faster than a more general quantum semidefinite programming algorithm.[164]

COAUTHORED PUBLICATIONS

[5] Continuous matrix product state tomography of quantum transport experiments
Géraldine Haack, Adrian Steffens, Jens Eisert, and Robert Hübener,
New Journal of Physics 17:113024, 2015.

[6] Quantum singular value decomposition of non-sparse low-rank matrices
Patrick Rebentrost, Adrian Steffens, and Seth Lloyd,
ArXiv e-prints 1607.05404, 2016.

Quantum singular value decomposition of non-sparse low-rank matrices

Patrick Rebentrost,[1,*] Adrian Steffens,[2,1] and Seth Lloyd[1,3,†]

[1]*Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139*
[2]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin*
[3]*Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139*

In this work, we present a method to exponentiate non-sparse indefinite low-rank matrices on a quantum computer. Given an operation for accessing the elements of the matrix, our method allows singular values and associated singular vectors to be found quantum mechanically in a time exponentially faster in the dimension of the matrix than known classical algorithms. The method extends to non-Hermitian and non-square matrices via embedding matrices. In the context of the generic singular value decomposition of a matrix, we discuss the Procrustes problem of finding a closest isometry to a given matrix.

Matrix computations are central to many algorithms in optimization and machine learning [1–3]. At the heart of these algorithms regularly lies an eigenvalue or a singular value decomposition of a matrix, or a matrix inversion. Such tasks could be performed efficiently via phase estimation on a universal quantum computer [4], as long as the matrix can be simulated (exponentiated) efficiently and controllably as a Hamiltonian acting on a quantum state. Almost exactly twenty years ago, Ref. [5] paved the way for such a simulation of quantum systems by introducing an efficient algorithm for exponentiating Hamiltonians with tensor product structure—enabling applications such as in quantum computing for quantum chemistry [6]. Step by step, more general types of quantum systems were tackled and performance increased: Aharonov and Ta-Shma [7] showed a method for simulating quantum systems described by sparse Hamiltonians, while Childs *et al.* [8] demonstrated the simulation of a quantum walk on a sparse graph. Berry *et al.* [9] reduced the temporal scaling to approximately linear via higher-order Suzuki integrators. Further improvements in the sparsity scaling were presented in Ref. [10]. Beyond sparse Hamiltonians, quantum principal component analysis (qPCA) was shown to handle non-sparse positive semidefinite low-rank Hamiltonians [11] when given multiple copies of the Hamiltonian as a quantum density matrix. This method has applications in quantum process tomography and state discrimination [11], as well as in quantum machine learning [12–18], specifically in curve fitting [19] and support vector machines [20]. In an oracular setting, Ref. [10, 21, 22] showed the simulation of non-sparse Hamiltonians via discrete quantum walks. The scaling in terms of the simulated time t is $t^{3/2}$ or even linear in t.

In the spirit of Ref. [11], we provide an alternative method for non-sparse matrices in an oracular setting which requires only one-sparse simulation techniques. We achieve a run time in terms of the matrix maximum element and a t^2 scaling. We discuss a class of matrices with low-rank properties that make the non-sparse methods efficient. Compared to Ref. [11] the matrices need not be positive semidefinite. In order to effectively treat a general non-Hermitian non-quadratic matrix, we make use of an indefinite "extended Hermitian matrix" that incorporates the original matrix. With such an extended matrix, we are able to efficiently determine the singular value decomposition of dense non-square, low-rank matrices. As one possible application of our method, we discuss the Procrustes problem [1] of finding a closest isometric matrix.

Method. We have been given an $N \times N$ dense (non-sparse) Hermitian indefinite matrix $A \in \mathbb{C}^{N\times N}$ via efficient oracle access to the elements of A. The oracle either performs an efficient computation of the matrix elements or provides access to a storage medium for the elements such as quantum RAM [23, 24]. Our new method simulates $\mathrm{e}^{-\mathrm{i}\,(A/N)t}$ on an arbitrary quantum state for arbitrary times t. Note that the eigenvalues of A/N are bounded by $\pm\|A\|_{\max}$, where $\|A\|_{\max}$ is the maximal absolute value of the matrix elements of A. This means that there exist matrices A for which the unitary $\mathrm{e}^{-\mathrm{i}\,(A/N)t}$ can be far from the identity operator for a time of order $\|A\|^{-1}_{\max}$, i.e. an initial state can evolve to a perfectly distinguishable state. For such times, the unitary $\mathrm{e}^{-\mathrm{i}\,(A/N)t}$ can be well approximated by a unitary generated by a low-rank matrix.

Let σ and ρ be N-dimensional density matrices. The state σ is the target state on which the matrix exponential of A/N is applied to, while multiple copies of ρ are used as ancillary states. Our method embeds the N^2 elements of A into a Hermitian sparse matrix $S_A \in \mathbb{C}^{N^2\times N^2}$, which we call "modified swap matrix" because of its close relation to the usual swap matrix. Each column of S_A contains a single element of A. The modified swap matrix between the registers for a single copy of ρ and σ is

$$S_A = \sum_{j,k=1}^{N} A_{jk}|k\rangle\langle j| \otimes |j\rangle\langle k| \in \mathbb{C}^{N^2\times N^2}. \qquad (1)$$

This matrix is one-sparse in a quadratically bigger space and reduces to the usual swap matrix for $A_{jk} = 1$ and $j, k = 1, \ldots, N$. Given efficient oracle access to the elements, we can simulate a one-sparse matrix such as S_A with a constant number of oracle calls and negligible error [7–9, 25]. We discuss the oracle access below. This matrix exponential of S_A is applied to a tensor product of a uniform superposition and an arbitrary state. Performing S_A for small Δt leads to a reduced dynamics of σ when expanded to terms of second order in Δt

as

$$\mathrm{tr}_1\{\mathrm{e}^{-\mathrm{i}\,S_A\Delta t}\,\rho\otimes\sigma\,\mathrm{e}^{\mathrm{i}\,S_A\Delta t}\} = \sigma - \mathrm{i}\,\mathrm{tr}_1\{S_A\,\rho\otimes\sigma\}\Delta t + \mathrm{i}\,\mathrm{tr}_1\{\rho\otimes\sigma\,S_A\}\Delta t + O(\Delta t^2). \quad (2)$$

Here, tr_1 denotes the partial trace over the first register containing ρ. The first $O(\Delta t)$ term is $\mathrm{tr}_1\{S_A\,\rho\otimes\sigma\} = \sum_{j,k=1}^{N} A_{jk}\langle j|\rho|k\rangle|j\rangle\langle k|\sigma$. Choosing $\rho = |\vec{1}\rangle\langle\vec{1}|$, with $|\vec{1}\rangle \coloneqq \frac{1}{\sqrt{N}}\sum_k |k\rangle$ the uniform superposition, leads to $\mathrm{tr}_1\{S_A\,\rho\otimes\sigma\} = \frac{A}{N}\,\sigma$. This choice for ρ contrasts with qPCA, where ρ is proportional to the simulated matrix [11]. Analogously, the second $O(\Delta t)$ term becomes $\mathrm{tr}_1\{\rho\otimes\sigma\,S_A\} = \sigma\,\frac{A}{N}$. Thus for small times, evolving with the modified swap matrix S_A on the bigger system is equivalent to evolving with A/N on the σ subsystem,

$$\mathrm{tr}_1\{\mathrm{e}^{-\mathrm{i}\,S_A\Delta t}\,\rho\otimes\sigma\,\mathrm{e}^{\mathrm{i}\,S_A\Delta t}\} = \sigma - \mathrm{i}\,\frac{\Delta t}{N}\,[A,\sigma] + O(\Delta t^2) \approx \mathrm{e}^{-\mathrm{i}\,\frac{A}{N}\Delta t}\,\sigma\,\mathrm{e}^{\mathrm{i}\,\frac{A}{N}\Delta t}. \quad (3)$$

Let ϵ_0 be the trace norm of the error term $O(\Delta t^2)$. We can bound this error by $\epsilon_0 \leq 2\|A\|_{\max}^2\Delta t^2$ (see Appendix). Here, $\|A\|_{\max} = \max_{mn}|A_{mn}|$ denotes the maximal absolute element of A. Note that $\|A\|_{\max}$ coincides with the largest absolute eigenvalue of S_A. The operation in Eq. (3) can be performed multiple times in a forward Euler fashion using multiple copies of ρ. For n steps the resulting error is $\epsilon = n\,\epsilon_0$. The simulated time is $t = n\,\Delta t$. Hence, fixing ϵ and t,

$$n = O\left(\frac{t^2}{\epsilon}\|A\|_{\max}^2\right) \quad (4)$$

steps are required to simulate $\mathrm{e}^{-i\frac{A}{N}t}$. The total run time of our method is nT_A, the number steps n is multiplied with the matrix oracle access time T_A (see below).

We discuss for which matrices the algorithm runs efficiently. Note that an upper bound for the eigenvalues of A/N in terms of the maximal matrix element is $|\lambda_j|/N \leq \|A\|_{\max}$. At a simulation time t only the eigenvalues of A/N with $|\lambda_j|/N = \Omega(1/t)$ matter. Let the number of these eigenvalues be r. Thus, effectively a matrix A_r/N is simulated with $\mathrm{tr}\{A_r^2/N^2\} = \sum_{j=1}^{r}\lambda_j^2/N^2 = \Omega(r/t^2)$. It also holds that $\mathrm{tr}\{A_r^2/N^2\} \leq \|A\|_{\max}^2$. Thus, the rank of the effectively simulated matrix is $r = O(\|A\|_{\max}^2 t^2)$.

Concretely, for the algorithm to be efficient in terms of matrix oracle calls, we require that the number of simulation steps n is $O(\mathrm{poly}\log N)$. Let the desired error be $1/\epsilon = O(\mathrm{poly}\log N)$. Assuming $\|A\|_{\max} = \Theta(1)$, meaning a constant independent of N, we have from Eq. (4) that we can only exponentiate for a time $t = O(\mathrm{poly}\log N)$. For such times, only the large eigenvalues of A/N with $|\lambda_j|/N = \Omega(1/\mathrm{poly}\log N)$ matter. Such eigenvalues can be achieved when the matrix is dense enough, for example A/N has $\Theta(N)$ non-zeros of size $\Theta(1/N)$ per row. For the rank of the simulated matrix in this case we find that $r = O(\mathrm{poly}\log N)$, thus effectively a low-rank matrix is simulated. To summarize, we expect the method to work well for low rank matrices A that are dense with relatively small matrix elements.

A large class of matrices satisfies these criteria. Sample a random unitary $U \in \mathbb{C}^{N\times N}$ and r suitable eigenvalues of size $|\lambda_j| = \Theta(N)$ and multiply them as $U\,\mathrm{diag}_r(\lambda_j)\,U^\dagger$ to construct A. Here, $\mathrm{diag}_r(\lambda_j)$ is the diagonal matrix with the r eigenvalues on the diagonal and zero otherwise. A typical random normalized vector has absolute matrix elements of size $O(1/\sqrt{N})$. The outer product of such a vector with itself has absolute matrix elements of size $O(1/N)$. Each eigenvalue of absolute size $\Theta(N)$ is multiplied with such an outer product and the r terms are summed up. Thus, a typical matrix element of A will be of size $O(\sqrt{r})$ and $\|A\|_{\max} = O(r)$.

Phase estimation. Phase estimation provides a gateway from unitary simulation to many interesting applications. For the use in phase estimation, we extend our method such that the matrix exponentiation of A/N can be performed conditioned on additional control qubits. With our method, the eigenvalues λ_j/N of A/N can be both positive and negative. The modified swap operator S_A for a Hermitian matrix A with eigendecomposition $A = \sum_j \lambda_j|u_j\rangle\langle u_j|$ is augmented as $|1\rangle\langle 1|\otimes S_A$, which still is a one-sparse Hermitian operator. The resulting unitary $\mathrm{e}^{-\mathrm{i}\,|1\rangle\langle 1|\otimes S_A\Delta t} = |0\rangle\langle 0|\otimes\mathbb{1} + |1\rangle\langle 1|\otimes\mathrm{e}^{-\mathrm{i}\,S_A\Delta t}$ is efficiently simulatable. This operator is applied to a state $|c\rangle\langle c|\otimes\rho\otimes\sigma$ where $|c\rangle$ is an arbitrary control qubit state. Sequential application of such controlled operations allows the use phase estimation to prepare the state [25]

$$|\phi\rangle = \frac{1}{\sqrt{\sum_j|\beta_j|^2}}\sum_{\frac{|\lambda_j|}{N}\geq\epsilon}\beta_j|u_j\rangle|\frac{\lambda_j}{N}\rangle \quad (5)$$

from an initial state $|\psi\rangle|0\ldots0\rangle$ with $O(\lceil\log(1/\epsilon)\rceil)$ control qubits forming an eigenvalue value register. Here, $\beta_j = \langle u_j|\psi\rangle$ and ϵ is the accuracy for resolving eigenvalues. To achieve this accuracy, phase estimation is run for a total time $t = O(1/\epsilon)$. Thus, $O(\|A\|_{\max}^2/\epsilon^3)$ queries of the oracle for A are required, which is of order $O(\mathrm{poly}\log N)$ under the low-rank assumption for A discussed above.

Matrix oracle and resource requirements. To simulate the modified swap matrix, we employ the methods developed in Refs. [8, 9]. First, we assume access to the original matrix A,

$$|j\,k\rangle|0\cdots0\rangle \mapsto |j\,k\rangle|A_{jk}\rangle. \quad (6)$$

This operation can be provided by quantum random access memory (qRAM) [23, 24] using $O(N^2)$ storage space and quantum switches for accessing the data in $T_\mathrm{A} = O(\log^2 N)$ operations. Alternatively, there matrices whose elements are efficiently computable, i.e. $T_\mathrm{A} = O(\mathrm{poly}\log N)$. For the one-sparse matrix S_A, the unitary operation for the sparse simulation methods [8, 9] can be simply constructed from the oracle in Eq. (6) and is given by

$$|(j,k)\rangle|0\cdots0\rangle \mapsto |(j,k)\rangle|(k,j),(S_A)_{(k,j),(j,k)}\rangle. \quad (7)$$

Here, we use (j,k) as label for the column/row index of the modified swap matrix.

We compare the required resources with those of other methods for sparse and non-sparse matrices. For a general $N \times N$ and s-sparse matrix, $O(sN)$ elements need to be stored. In certain cases, the sparse matrix features more structure and its elements can be computed efficiently [9, 25]. For non-sparse matrices and the qPCA method in Ref. [11], only multiple copies of the density matrix as opposed to an operation as in Eq. (6) are required for applications such as state tomography. For machine learning via qPCA [11, 20], the density matrix is prepared from a classical source via quantum RAM [23, 24] and requires $O(N^2)$ storage. In comparison, the requirements of the method in this work are in principle not higher than these sparse and non-sparse methods, both in the case of qRAM access and in the case when matrix elements are computed instead of stored.

Non-square matrices. Our method enables us also to determine properties of general non-square low-rank matrices effectively. To determine the singular value decomposition of a matrix $A = U\Sigma V^\dagger \in \mathbb{C}^{M\times N}$ with rank r, simulating the positive semidefinite matrices $AA^\dagger$ and $A^\dagger A$ via qPCA yields the correct singular values and vectors. However, essential information is missing, leading to ambiguities in the singular vectors that become evident when inserting diagonal matrices into the singular value decomposition of $AA^\dagger$ that change the relative phases of the singular vectors,

$$AA^\dagger = U\Sigma^2 U^\dagger = U\Sigma D^\dagger V^\dagger \; VD\Sigma U^\dagger =: \hat{A}\hat{A}^\dagger, \quad (8)$$

with $D \coloneqq \mathrm{diag}(\mathrm{e}^{-\mathrm{i}\vartheta_j})$, ϑ_j being arbitrary phases. If $Av_j = \sigma_j u_j$ for each $j = 1, \ldots, r$, then

$$\hat{A}v_j = U\Sigma D^\dagger V^\dagger v_j = \sigma_j \mathrm{e}^{\mathrm{i}\vartheta_j} u_j \coloneqq \sigma_j \hat{u}_j, \quad (9)$$

which means different phase relations between left and right singular vectors in $\hat{A}$ from those in A. Although A and $\hat{A}$ still share the same singular values and even the same singular vectors up to phase factors, $\|A - \hat{A}\|_F$ will in general (with the exception of positive semidefinite matrices, where $U = V$) not be zero or even be small: The matrix A cannot be reproduced this way—a singular value decomposition is more than a set of singular values and normalized singular vectors. This affects all kinds of algorithms that require the appropriate phase relations between each left singular vector u_j and the according right singular vector v_j. Such applications are determining the best low-rank approximation of a matrix, signal processing algorithms discussed in Ref. [26], or determining the nearest isometric matrix, related to the unitary Procrustes problem, of a non-Hermitian matrix.

In order to overcome this issue, consider the "extended matrix"

$$\tilde{A} \coloneqq \begin{bmatrix} 0 & A \\ A^\dagger & 0 \end{bmatrix}, \quad (10)$$

which was introduced for singular value computations in Ref. [27] and recently in sparse quantum matrix inversion in [25]. The eigenvalues of $\tilde{A}$ correspond to $\{\pm\sigma_j\}$ with $\{\sigma_j\}$ being the singular values of A for $j = 1, \ldots, r$. The corresponding eigenvectors are proportional to $(u_j, \pm v_j) \in \mathbb{C}^{M+N}$, see Appendix. The left and right singular vectors of A can be extracted from the first M and last N entries, respectively. Since $\tilde{A}$ is Hermitian, its eigenvectors can assumed to be orthonormal: $\|(u_j, v_j)\|^2 = \|u_j\|^2 + \|v_j\|^2 = 1$, and $(u_j, v_j) \cdot (u_j, -v_j)^\dagger = \|u_j\|^2 - \|v_j\|^2 = 0$, from which follows that the norm of each of the subvectors u_j and v_j is $1/\sqrt{2}$, independent of their respective lengths M and N. The important point is that the eigenvectors of the extended matrix preserve the correct phase relations between the left and right singular vectors since $(\mathrm{e}^{\mathrm{i}\vartheta_j} u_j, v_j)$ is only an eigenvector of $\tilde{A}$ for the correct phase $\mathrm{e}^{\mathrm{i}\vartheta_j} = 1$.

The requirements for our quantum algorithm can be satisfied also for the extended matrix. For randomly sampled left and right singular vectors, the matrix elements have maximal size of $O(\sum_{j=1}^r \sigma_j/\sqrt{MN})$, thus $\sigma_j = O(\sqrt{MN})$. In addition, an $1/(M+N)$ factor arises in the simulation of the extended matrix from the ancillary state $\rho = |\vec{1}\rangle\langle\vec{1}|$ as before, which leads to the requirement $\sigma_j = \Theta(M+N)$. These two conditions for σ_j can be satisfied if the matrix A is not too skewed, i.e. $M = \Theta(N)$. In summary, by simulating the corresponding Hermitian extended matrices, general complex matrices of low rank can be simulated efficiently, yielding the correct singular value decomposition.

Procrustes problem. The unitary Procrustes problem is to find the unitary matrix that most accurately transforms one matrix into another. It has many applications, such as in shape/factor/image analysis and statistics [1]. We consider non-square matrices thus consider the Procrustes problem to find the *isometry* that most accurately transforms one matrix into another. Formally, minimize $\|WB - C\|_F$ among all isometries $W \in \mathbb{C}^{M\times N}$, $W^\dagger W = \mathbb{1}$, with $B \in \mathbb{C}^{N\times K}$ and $C \in \mathbb{C}^{M\times K}$, where $M > N$. The problem is equivalent to the general problem of finding the nearest isometric matrix $W \in \mathbb{C}^{M\times N}$ to a matrix $A \in \mathbb{C}^{M\times N}$ by taking $A = CB^\dagger$. Since our quantum algorithm is restricted to low rank matrices, let $A = CB^\dagger$ be low-rank with rank r and singular value decomposition $A = U\,\Sigma\,V^\dagger$ with $U \in \mathbb{C}^{M\times r}$, $\Sigma \in \mathbb{R}^{r\times r}$, and $V \in \mathbb{C}^{N\times r}$. The optimal solution to the Procrustes problem is $W = U\,V^\dagger$ [1], setting all singular values to one, in both the low-rank and the full-rank situation. Since A is assumed to be low rank, we find a *partial* isometry with $W^\dagger W = \mathbb{P}_{\mathrm{col}(V)}$, with $\mathbb{P}_{\mathrm{col}(V)}$ the projector into the subspace spanned by the columns of V. Thus, W acts as an isometry for vectors in that subspace (see Appendix).

In a quantum algorithm, we want to apply the nearest low-rank isometry to a quantum state $|\psi\rangle$. The state $|\psi\rangle$ is assumed to be in or close to the subspace spanned by the columns of V. We assume that the extended matrix for A in Eq. (10) is given in oracular form and that A is not too skewed such that $\sigma_j/(M+N) = \Theta(1)$ and $\|A\|_{\max} = \Theta(1)$. We perform phase estimation on the input state $|0, \psi\rangle|0\ldots0\rangle$ and, analogous to Eq. (5), obtain a state proportional to

$$\sum_{\frac{\sigma_j}{M+N}\geq\epsilon} \beta_j^\pm |u_j, \pm v_j\rangle| \pm \frac{\sigma_j}{M+N}\rangle \quad (11)$$

with $\beta_j^{\pm} = \langle u_j, \pm v_j | 0, \psi \rangle = \pm \langle v_j | \psi \rangle / \sqrt{2}$. The sum has $2r$ terms corresponding to the eigenvalues of the extended matrix with absolute value greater than $(M+N)\epsilon$. Performing a σ_z operation on the qubit encoding the sign of the respective eigenvalue an uncomputing the eigenvalue register yields a state proportional to $\sum_j \beta_j |u_j, \pm v_j\rangle$. Projecting onto the u_j part (success probability $1/2$) results in a state proportional to

$$\sum_{\frac{\sigma_j}{M+N} \geq \epsilon} |u_j\rangle\langle v_j|\psi\rangle \propto U V^{\dagger} |\psi\rangle. \tag{12}$$

This prepares the desired state for the non-square low-rank Procrustes problem with accuracy ϵ in runtime $O(\|A\|_{\max}^2 \log^2(N+M)/\epsilon^3)$. Classically, performing the singular value decomposition of a low-rank A without further structural assumptions takes generally $O(N^3)$.

Conclusion. The method presented here allows non-sparse low-rank non-positive Hermitian $N \times N$ matrices A/N to be exponentiated for a time t with accuracy ϵ in run time $O\left(\frac{t^2}{\epsilon}\|A\|_{\max}^2 T_A\right)$, where $\|A\|_{\max}$ is the maximal absolute element of A. The data access time is T_A. If the matrix elements are accessed via quantum RAM or computed efficiently and the significant eigenvalues of A are $\Theta(N)$, our method can achieve a run time of $O(\text{poly} \log N)$ for a large class of matrices. Our method allows non-Hermitian and non-square matrices to be exponentiated via extended Hermitian matrices.

We have shown how compute the singular value decomposition of a non-Hermitian non-sparse matrix on a quantum computer directly while keeping all the correct relative phase information. As one of the many potential applications of the singular value decomposition, we can find the pseudoinverse of a matrix and the closest isometry exponentially faster than any known classical algorithm. It remains to be seen if the time complexity of our method can be improved from $O(t^2)$ to an approximately linear scaling via higher-order Suzuki-Trotter steps or other techniques. In addition, by using a (possibly unknown) ancillary state other than the uniform superposition, the oracular setting of the present work and the tomography setting of [11] could be combined.

We are grateful to Iman Marvian for insightful discussions. We acknowledge support from DARPA, NSF, and AFOSR. AS thanks the German National Academic Foundation (Studienstiftung des deutschen Volkes) and the Fritz Haber Institute of the Max Planck Society for support.

* rebentr@mit.edu
† slloyd@mit.edu

[1] G. H. Golub and C. F. Van Loan, *Matrix computations*, Vol. 3 (JHU Press, 2012).
[2] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, 2004).
[3] K. P. Murphy, *Machine Learning: A Probabilistic Perspective* (MIT Press, 2012).
[4] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge university press, 2010).
[5] S. Lloyd, Science **273**, 1073 (1996).
[6] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Science **309**, 1704 (2005).
[7] D. Aharonov and A. Ta-Shma, in *Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Compu* STOC '03 (ACM, New York, NY, USA, 2003) pp. 20–29.
[8] A. M. Childs, R. Cleve, E. Deotto, E. Farhi, S. Gutmann, and D. A. Spielman, in *Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Compu* STOC '03 (ACM, New York, NY, USA, 2003) pp. 59–68.
[9] D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders, Comm. Math. Phys. **270**, 359 (2007).
[10] D. W. Berry and A. M. Childs, Quantum Info. Comput. **12**, 29 (2012).
[11] S. Lloyd, M. Mohseni, and P. Rebentrost, Nature Physics **10**, 631 (2014).
[12] N. Wiebe, A. Kapoor, and K. M. Svore, arXiv preprint arXiv:1401.2142 (2014).
[13] N. Wiebe, A. Kapoor, and K. M. Svore, arXiv preprint arXiv:1412.3489 (2014).
[14] M. Benedetti, J. Realpe-Gómez, R. Biswas, and A. Perdomo-Ortiz, arXiv preprint arXiv:1510.07611 (2015).
[15] M. Schuld, I. Sinayskiy, and F. Petruccione, Physics Letters A **379**, 660 (2015).
[16] M. Schuld, I. Sinayskiy, and F. Petruccione, arXiv preprint arXiv:1601.07823 (2016).
[17] I. Kerenidis and A. Prakash, arXiv preprint arXiv:1603.08675 (2016).
[18] H.-K. Lau, R. Pooser, G. Siopsis, and C. Weedbrook, arXiv preprint arXiv:1603.06222 (2016).
[19] G. Wang, arXiv preprint arXiv:1402.0660 (2014).
[20] P. Rebentrost, M. Mohseni, and S. Lloyd, Physical Review Letters **113**, 130503 (2014).
[21] A. Childs, Comm. Math. Phys. **294**, 581 (2010).
[22] A. Childs and R. Kothari, in Theory of Quantum Computation, Communication, and Cryptography, Lecture Notes in Computer Science Vol. 6519, edited by W. van Dam, V. Kendon, and S. Severini (Springer, Berlin, Heidelberg, 2011), p. 94, ISBN 978-3-642-18072-9 (2011).
[23] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **100**, 160501 (2008).
[24] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. A **78**, 052310 (2008).
[25] A. W. Harrow, A. Hassidim, and S. Lloyd, Phys. Rev. Lett. **103**, 150502 (2009).
[26] A. Steffens, P. Rebentrost, I. Marvian, J. Eisert, and S. Lloyd, to be submitted (2016).
[27] G. Golub and W. Kahan, Journal of the Society for Industrial and Applied Mathematics, Series B: Numerical Analysis **2**, 205 (1965).

Appendix

Norms. Denote the maximum absolute element of a matrix $A \in \mathbb{C}^{N \times N}$ with $\|A\|_{\max} := \max_{j,k} |A_{jk}|$. The Frobenius or Hilbert-Schmidt norm is given by $\|A\|_F := \sqrt{\sum_{j,k} |A_{jk}|^2}$ and its nuclear norm by $\|A\|_* := \sum_{i=1}^{r} \sigma_i$, where r is the rank and σ_j are the singular values.

Modified swap matrix. The modified swap matrix is defined as

$$S_A = \sum_{j,k=1}^{N} A_{jk} |k\rangle\langle j| \otimes |j\rangle\langle k| \in \mathbb{C}^{N^2 \times N^2}. \quad (13)$$

Taking $A_{jk} \to 1$ leads to the original swap matrix $S = \sum_{j,k=1}^{N} |k\rangle\langle j| \otimes |j\rangle\langle k| \in \mathbb{C}^{N^2 \times N^2}$. The N^2 eigenvalues of S_A are

$$A_{11}, A_{22}, \ldots, A_{NN}, A_{12}, -A_{12}, \ldots, A_{j,k>j}, -A_{j,k>j}, \ldots, \quad (14)$$

where $k > j$ denotes an index k greater than j. The maximal absolute eigenvalue of S_A is thus $\max_{j,k} |A_{jk}| \equiv \|A\|_{\max}$, corresponding to the maximal absolute matrix element of A. The square of the modified swap matrix is

$$(S_A)^2 = \sum_{j,k=1}^{N} |A_{jk}|^2 |k\rangle\langle k| \otimes |j\rangle\langle j| \leq \|A\|_{\max}^2 \mathbb{1}. \quad (15)$$

Its eigenvalues are $|A_{jk}|^2$ and the maximal eigenvalue is $\|A\|_{\max}^2$. This already points to the result that the second order error of our method naturally scales with $\|A\|_{\max}^2$, which we will now derive.

Error analysis. In the following, we estimate the error from the second-order term in Δt in the expansion Eq. (2). The nuclear norm of the operator part of the second order error is

$$\epsilon_{\rho,\sigma} = \|\mathrm{tr}_1\{S_A\, \rho \otimes \sigma\, S_A\} - \frac{1}{2}\mathrm{tr}_1\{(S_A)^2\, \rho \otimes \sigma\} - \frac{1}{2}\mathrm{tr}_1\{\rho \otimes \sigma\, (S_A)^2\}\|_*. \quad (16)$$

In Ref. [11], this error was equal to $\epsilon_{\rho,\sigma}^{\mathrm{qPCA}} = \|\rho - \sigma\|_* \leq 2$, which is achieved in the present algorithm by choosing A such that $A_{jk} = 1$ for each j, k. Here, our algorithm coincides with the qPCA method for ρ chosen as the uniform superposition. For general low-rank A, we bound Eq. (16) via the triangle inequality. Taking the nuclear norm of the first term results in

$$\begin{aligned}\|\mathrm{tr}_1\{S_A \rho \otimes \sigma S_A\}\|_* \leq& \|S_A \rho \otimes \sigma S_A\|_* \\ \leq& \|\rho \otimes \sigma\|_* \|S_A^2\|_* \leq \|A\|_{\max}^2. \end{aligned} \quad (17)$$

The second and third term can be treated similarly. We obtain $\|\mathrm{tr}_1\{(S_A)^2 \rho \otimes \sigma\}\|_* \leq \|A\|_{\max}^2$. Combining all terms yields the bound

$$\epsilon_{\rho,\sigma} \leq 2\|A\|_{\max}^2. \quad (18)$$

Extended matrices. We define the Hermitian extended matrix $\tilde{A}$ of a complex-valued, not necessarily square matrix $A \in \mathbb{C}^{M \times N}$ as

$$\tilde{A} = \begin{bmatrix} 0 & A \\ A^\dagger & 0 \end{bmatrix} \in \mathbb{C}^{(M+N)\times(M+N)}. \quad (19)$$

Using block matrix identities for the determinant, we obtain its characteristic polynomial

$$\chi_{\tilde{A}}(\lambda) = \lambda^{|M-N|} \det(\lambda \mathbb{1} + \sqrt{AA^\dagger})(\lambda \mathbb{1} - \sqrt{AA^\dagger}). \quad (20)$$

The eigenvalues of $\tilde{A}$ are either zero or correspond to $\{\pm\sigma_j\}$, the singular values of A for $j = 1, \ldots, r$ with an additional sign. Hence, if A has low rank r, then $\tilde{A}$ has low rank $2r$. The corresponding eigenvectors are proportional to $(u_j, \pm v_j) \in \mathbb{C}^{M+N}$ since

$$\begin{bmatrix} \mp\sigma_j \mathbb{1} & A \\ A^\dagger & \mp\sigma_j \mathbb{1} \end{bmatrix} \cdot \begin{bmatrix} u_j \\ \pm v_j \end{bmatrix} = 0, \quad (21)$$

where u_j and v_j are the jth left and right singular vector of A, respectively. The important point is that the eigenvectors of the extended matrix preserve the correct phase relations between the left and right singular vectors since $(\mathrm{e}^{\mathrm{i}\vartheta_j} u_j, \pm v_j)$ is only an eigenvector of $\tilde{A}$ for the correct phase $\mathrm{e}^{\mathrm{i}\vartheta_j} = 1$,

$$\begin{aligned}\begin{bmatrix} \mp\sigma_j \mathbb{1} & A \\ A^\dagger & \mp\sigma_j \mathbb{1} \end{bmatrix} \cdot \begin{bmatrix} \mathrm{e}^{\mathrm{i}\vartheta_j} u_j \\ \pm v_j \end{bmatrix} &= \begin{bmatrix} \mp\sigma_j \mathrm{e}^{\mathrm{i}\vartheta_j} u_j \pm A v_j \\ \mathrm{e}^{\mathrm{i}\vartheta_j} A^\dagger u_j - \sigma_j v_j \end{bmatrix} \\ &= (\mathrm{e}^{\mathrm{i}\vartheta_j} - 1)\sigma_j \begin{bmatrix} \mp u_j \\ v_j \end{bmatrix}. \end{aligned} \quad (22)$$

The right hand side is only equal to zero for the correct phase $\mathrm{e}^{\mathrm{i}\vartheta_j} = 1$.

Low-rank Procrustes. Let the isometry be $W = UV^\dagger$ with $U \in \mathbb{C}^{M \times r}$ and $V \in \mathbb{C}^{N \times r}$. Assume that $M > N$, giving orthogonal columns in the full-rank Procrustes problem ($r = N$). We find for the low-rank (partial) isometry that

$$W^\dagger W = VU^\dagger UV^\dagger = VV^\dagger = \sum_{j=1}^{r} \vec{v}_j \vec{v}_j^\dagger. \quad (23)$$

Pick an arbitrary vector $\vec{x} = \sum_{j=1}^{r} \alpha_j \vec{v}_j + \vec{x}^\perp = \vec{x}^\parallel + \vec{x}^\perp$. where $\vec{x}^\perp$ denotes the part orthogonal to the orthonormal vectors $\vec{v}_j$. Then,

$$W^\dagger W \vec{x} = \sum_{j=1}^{r} \alpha_j \vec{v}_j = \vec{x}^\parallel. \quad (24)$$

Thus, $W^\dagger W$ acts as the identity operator in the low-rank subspace, and projects out the space perpendicular to that subspace.

A.3 Compressive density functional theory

Density functional theory[187,188] (DFT) is a very successful method for determining the electronic structure, especially the ground state properties, of a quantum many-body system, allowing to accurately compute quantities such as the ionization and atomization energies or the vibrational frequencies of a molecule. Complementarily to tensor network states, the key point of DFT is to treat a N-body quantum system just in terms of its electronic density $\varrho(r) \in \mathbb{R}_+$ at position $r \in \mathbb{R}^3$, thus reducing the N-particle problem with the entailing curse of dimensionality (see section 2.1) to a one-particle problem in the three spatial coordinates of ϱ. DFT works astonishingly well for many applications and has its theoretical foundation in the *Hohenberg-Kohn* theorems[189]. For an overview, see, e.g., Refs. [190, 191].

Since atomic nuclei are much heavier and hence move slower than electrons, the overall Hamiltonian can be split into an electronic and an atomic part that can be treated consecutively (*Born-Oppenheimer approximation*). The electronic Hamiltonian acting on an N-electron wave function $\Psi(r_1, \ldots, r_N)$ consists of a kinetic term T, the electron-electron interactions V_{ee} depending on the distances $r_{j,k} = |r_j - r_k|$ and the electron-nucleus interactions V_{ext} depending on $\tilde{r}_{j,A} = |r_j - \tilde{r}_A|$, and the nuclear charges Z_A:

$$H = -\frac{1}{2}\sum_{j=1}^{N}\nabla_j^2 + \sum_{1\le j<k\le N}\frac{1}{r_{j,k}} - \sum_{j,A=1}^{N}\frac{Z_A}{\tilde{r}_{j,A}} = T + V_{\text{ee}} + V_{\text{ext}}. \tag{34}$$

The ground state of H with ground state energy E_0 could in principle be determined using the *Rayleigh-Ritz* variational principle

$$E_0 = \min_{|\Psi\rangle}\frac{\langle\Psi|T + V_{\text{ee}} + V_{\text{ext}}|\Psi\rangle}{\langle\Psi|\Psi\rangle}. \tag{35}$$

For high-dimensional $|\Psi\rangle$, this becomes computational infeasible. The first Hohenberg-Kohn theorem, however, states that the external potential is uniquely determined by the electronic density: V_{ext} is a unique functional of ϱ. Since V_{ext} also fixes the Hamiltonian H and hence the state vector $|\Psi\rangle$, there is a one-to-one correspondence between the electronic density and the multi-particle wave function for each external potential. The correspondence is not constructive, however, and the explicit form of the universal, system-independent functional $F : \varrho \mapsto T[\varrho] + E_{\text{ee}}[\varrho]$ is unknown. In this sense, the computational complexity of the original problem of finding the ground state has been passed over to the specification of the universal functional F. In fact, it has been shown that determining F is computationally hard even on a hypothetical quantum computer[192].

Nevertheless, the functional can be suitably approximated in many practically relevant situations. Introducing a noninteracting reference system with electronic orbitals $\{\psi_j(r)\}$, yielding the same electronic density as in the original system, $\varrho(r) = \sum_{j=1}^{N}|\psi_j(r)|^2$, and with approximative

kinetic energy

$$T_{\mathrm{KS}} = -\frac{1}{2}\sum_{j=1}^{N}\langle\psi_j|\nabla^2|\psi_j\rangle, \tag{36}$$

followed by applying the variational principle on the formal expression of the energy functional, results in the famous *Kohn-Sham equations*[193]:

$$\left(-\frac{1}{2}\nabla^2 + \int \mathrm{d}r'\,\frac{\varrho(r')}{|r-r'|} - \sum_{A=1}^{N}\frac{Z_A}{|r-\tilde{r}_A|} + V_{\mathrm{XC}}(\varrho(r))\right)\psi_j(r) = \varepsilon_j\psi_j(r), \quad j=1,\ldots,N \tag{37}$$

Since the electronic density ϱ depends on the Kohn-Sham orbitals ψ_j, Eq. (37) has to be solved iteratively until the eigenvalues and the density converge (self-consistency). All remaining energetic contributions without explicit analytical form are summarized in the exchange-correlation term V_{XC}. Well-known heuristics for V_{XC} include *local density approximation* (LDA) functionals such as the *Perdew-Wang* functional[194], *generalized gradient approximation* (GGA) functionals—additionally taking gradient information of the electronic density into account—such as the *Perdew–Burke–Ernzerhof* (PBE) functional[195], and linear combinations of LDA and GGA functionals with the *Hartree-Fock* exchange functional (so-called hybrid functionals) such as *B3LYP* (*Becke* functional, three parameters, *Lee-Yang-Parr* functional)[196]. The standard functionals are not suitable for systems where the density varies rapidly, such as systems with Van der Waals interactions, where further extensions have to be included.[197]

For practical computations, each Kohn-Sham orbital is expanded into a linear combination of basis functions out of a dictionary of size $d \geq N$,

$$\psi_j = \sum_{k=1}^{d} c_k^{(j)}\varphi_k, \quad j=1,\ldots,N. \tag{38}$$

Inserting this expansion into Eq. (37) results in the $d\times d$ real-valued generalized eigenvalue problem

$$H^{\mathrm{KS}}c^{(j)} = \varepsilon_j S c^{(j)}, \quad j=1,\ldots,N \tag{39}$$

with Kohn-Sham Hamiltonian H^{KS} with components $H^{\mathrm{KS}}_{j,k} = \int\overline{\psi}_j(r)\widetilde{H}^{\mathrm{KS}}\psi_k(r)\,\mathrm{d}^3r$ ($\widetilde{H}^{\mathrm{KS}}$ denoting the operator on the left hand side of Eq. (37)) and overlap matrix S with components $S_{j,k} = \int\overline{\psi}_j(r)\psi_k(r)\,\mathrm{d}^3r$. H^{KS} depends via the electronic density ϱ on the coefficients $c_k^{(j)}$ corresponding to the occupied orbitals (the ones with lowest energy eigenvalues ε_j) and is updated after each iteration until the required self-consistency criteria are met. The accuracy of the DFT calculations critically relies on the employed basis set and great effort has been put into establishing suitable basis functions[198–201] resulting in large basis set databases[202].

In general, the computational effort for self-consistently solving Eq. (39) scales asymptotically with $O(d^3)$, which becomes forbidding for large systems.[26] On the other hand, discarding basis functions in general amounts to lower computational effort, but also lower accuracy.

The goal of the present project is to use concepts from compressed sensing in order to identify basis subsets that due to their considerably smaller size allow for faster computations while still retaining sufficient accuracy for the physical properties of interest. In the following, the principles of the to be published project [7] are outlined.

Similar approaches outside the DFT context for selecting localized *Wannier functions* out of a dictionary of functions have been pursued[205,206] using an ℓ_1-regularization term (see Eq. (11)). The optimization routine, which has to take orthogonality constraints for the orbitals ψ_j into account, was based on an *Bregman-iterations* algorithm[207,208] that is closely related to the *Alternating Direction Method of Multipliers* (ADMM)[209]. Building on this, ℓ_1-regularized Bregman optimization algorithms for density matrix minimization have been proposed[210,211], as well as an ℓ_1-regularized orbital minimization method[212]—the computational advantage over methods such as in Ref. [201] has yet to be worked out, however. Beyond ground state calculations, DFT and compressed sensing have, e.g., been combined for learning physical descriptors in materials science[213].

By arranging the generalized eigenvectors $\{c^{(j)}\}$ that correspond to the N occupied orbitals (ignoring electronic spin degeneracy for clarity) as column vectors of a matrix $C \in \mathbb{C}^{D\times N}$, Eq. (39) can be reformulated as an iterative minimization problem

$$\begin{aligned} C^{l+1} = \; & \underset{C\in\mathbb{R}^{D\times N}}{\arg\min} \; \mathrm{tr}(H^{\mathrm{KS}}(C^l)\, C C^T) \\ & \text{subject to} \quad C^T S C = \mathbb{1}_N, \end{aligned} \tag{40}$$

where the constraint represents the orthonormality of the columns of C (corresponding to the electronic orbitals) with respect to the overlap matrix S. The minimizer of the l-th iteration C^l determines H^{KS} for the subsequent iteration until self-consistency is achieved.

Using a large dictionary of basis functions (φ_k), many basis functions will not contribute much to the minimizing function, i.e., the corresponding *rows* in C will comprise entries close to zero, which could be facilitated by introducing a regularization term, as in Eq. (11), consisting of the ℓ_1-norm of the ℓ_2-norms of the row vectors of C,

$$\|C\|_{2,1} \coloneqq \sum_{k=1}^{D} \left(\sum_{j=1}^{N} C_{k,j}^2 \right)^{1/2}, \tag{41}$$

[26] Making certain physical assumptions such as locality, however, one can also design linear scaling methods.[203,204]

which is also used, e.g., in *multi-task feature learning*[214]. Due to the constraint, the resulting minimization problem

$$\min_{C\in\mathbb{R}^{n\times N}} f_{\text{obj}}(C) \coloneqq \min_{C\in\mathbb{R}^{D\times N}} \operatorname{tr}(H^{\text{KS}}CC^T) + \lambda||C||_{2,1} \quad \text{subject to} \quad C^TSC = \mathbb{1}_N. \tag{42}$$

is non-convex and, thus, cannot easily be treated with generic convex solvers. However, the structure at hand allows for the use of techniques that make use of the manifold geometry and provide efficient optimization routines on matrix manifolds[215]. Matrices C that fulfill the constraint in Eq. (42) are elements of the *generalized Stiefel manifold*

$$V_N^S(\mathbb{R}^n) = \{C \in \mathbb{R}^{n\times N} : C^TSC = \mathbb{1}_N\}. \tag{43}$$

Using second order derivatives, a trust region algorithm was employed that acts on generalized Stiefel manifolds[216] and is based on the *Manopt*[217] toolbox implementation. The map $C^k \mapsto H^{\text{KS}}(C^k)$ was provided by accessing the *ab initio molecular simulations* package *FHI-aims*[201,218], whose input/output functions were modified for interfacing with *Matlab*[219] and *Manopt*. The sparsity of the minimizer C^* of Eq. (42) depends on the size of the regularization parameter $\lambda > 0$: In general, larger values of λ result in a higher concentration of the entries of C^* within a few rows (cf. Fig. 3), while for smaller λ more basis functions will contribute and the energy term, the sum of the lowest N generalized eigenvalues, will be more accurate. This corresponds to the fact that adding more basis functions to a dictionary necessarily increases the possibilities of describing the system, however often only marginally. This requires to carefully tune λ for trading off energetic accuracy and effective basis size (see Fig. 7).

The matrix-valued gradient of the objective function f_{obj} can be written as

$$\left.\frac{\partial}{\partial C_{i,j}}\right|_C f_{\text{obj}} = 2(HC)_{i,j} + \lambda\frac{C_{i,j}}{||C_{i,:}||_2}, \tag{44}$$

where $||C_{i,:}||_2 \coloneqq \left(\sum_{l=1}^N C_{i,l}^2\right)^{1/2}$ is the ℓ_2-norm of the i-th row of the matrix C. The fourth-order Hessian tensor with the components

$$\left.\frac{\partial^2}{\partial C_{k,l}\,\partial C_{i,j}}\right|_C f_{\text{obj}} = H_{i,k}\delta_{j,l} + 2\lambda\,\delta_{i,k}\left(\frac{\delta_{l,j}}{||C_{i,:}||_2} - \frac{C_{i,j}C_{i,l}}{||C_{i,:}||_2^3}\right) \tag{45}$$

can be treated as a matrix by grouping the indices i, j and k, l, respectively. The orbitals in *FHI-aims* with orbital quantum number ℓ arise in groups of $2\ell + 1$ elements comprising the subshells with magnetic quantum numbers $m = -\ell, \ldots, +\ell$. Accordingly, the objective function and its derivatives have to be modified by taking the ℓ_2-norm of all rows corresponding to one orbital group.

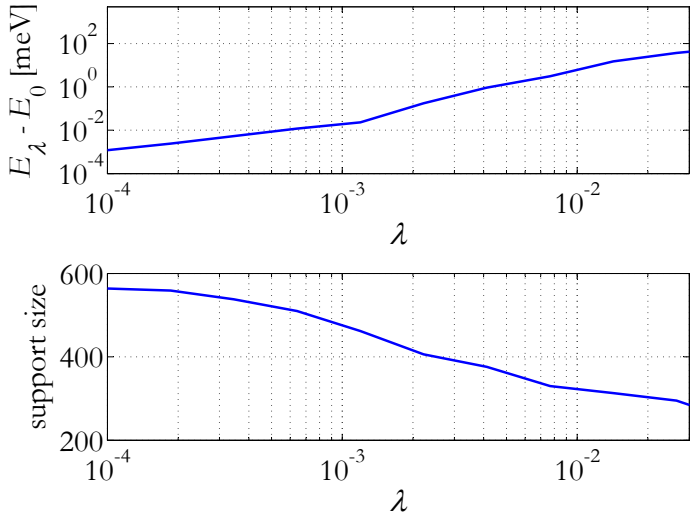

Figure 7: *Buying sparsity with energy at the example of the C_2H_4 molecule with $D = 796$ Gaussian basis functions with logarithmically even-spaced exponents: The difference between the generalized eigenvalue sum of the occupied orbitals E_λ with the one without regularization $E_{\lambda=0}$ is plotted as a function of the regularization parameter λ (above) after the initial SCF iteration. The corresponding size of the distilled basis set (support size) as a function of λ is shown below. The support size is taken to be the number of basis indices whose ℓ_2-norm is larger than 10^{-6}. By increasing λ, as the energy becomes less accurate, the effective basis set size becomes smaller.*

The resulting minimizer C^* of Eq. (42) is row-wise thresholded, i.e., rows $C^*_{i,:}$ with $||C^*_{i,:}||_2$ below a λ-dependent threshold η are considered noncontributing and discarded. The remaining basis functions constitute the distilled basis set, which, while substantially smaller, provides sufficient accuracy in describing the system.

In order to demonstrate the principle, the procedure is applied to an ethylene (C_2H_4) molecule, using *uncontracted Gaussian-type orbitals*[221] [27] with logarithmically even-spaced exponents and orbital quantum number $\ell = 0, \ldots, 4$ as an ansatz for a generic basis set.[28] As exchange-correlation functional, the PBE functional was employed. The geometric data of the nuclei was obtained from the *NOMAD repository*. *Pulay mixing*[222] was used to increase convergence speed towards self-consistency.

As can be seen by Fig. 7, larger values for λ yield smaller effective bases and larger deviations

[27] *Gaussian type orbitals* (GTO) are centered around the nuclei of the constituting atoms with radial part consisting of linear combinations of Gaussian functions with varying exponents and coefficients. Uncontracted GTOs possess only one Gaussian function.

[28] Handling more competitive basis functions such as *numeric atom-centered orbital* (NAO) basis functions[201] is less straightforward due to the underlying parameter structure, but has been implemented as well.

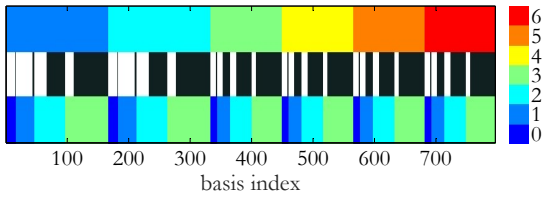

Figure 8: *Ethylene molecule C_2H_4: Selected basis functions (middle row, light). Discarded indices are marked with black. The basis set is ordered by atom number (upper row, the first two atoms are carbon, the remaining four hydrogen) and orbital angular quantum number ℓ (lower row).*

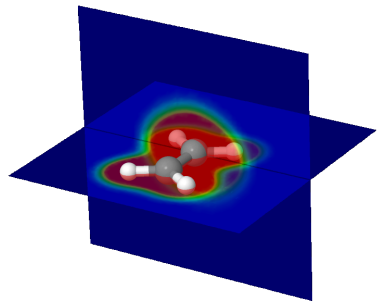

Figure 9: *Resulting electronic density of the ethylene molecule from the distilled basis set. The carbon nuclei are marked in gray, the hydrogen nuclei in white. A higher electronic density corresponds to a warmer color.* Created with the rendering software Jmol [220].

from the more accurate energy of the original, larger dictionary without regularization term $E_{\lambda=0}$. When exemplarily setting $\lambda = 0.01$, the basis functions are selected as shown in Fig. 8. There remain more basis functions per atom for the carbon atoms. Furthermore, basis functions with low orbital quantum number are preferred. The resulting basis set has a size of 292 basis functions out of originally 796 basis functions and an total energy deviation after achieving self-consistency of 12 meV (corresponding to a relative difference of $6 \cdot 10^{-6}$). This means that cutting the original basis by more than one half affects the relative error in terms of energy with less than 10^{-5}. For obtaining a desired target basis size, the regularization parameter λ can be tuned accordingly. Naturally, the procedure is not limited to ethylene, but has been successfully tested on other systems. The electronic density of the molecule $\varrho(r)$ corresponding to the reduced basis set is depicted in Fig. 9.

This illustrates how discarding larger parts of the models of molecules by use of compressed sensing methods still can preserve their descriptive power, paving the way for smarter Kohn-Sham orbital basis sets that allow for efficiently tackling larger quantum systems. The entire procedure scales with $O(D^3)$. This has to be compared to existing greedy strategies, such as the one in Ref. [201] for NAO basis functions, which scales linearly in the size of the dictionary and can be interpreted as a matching pursuit algorithm related to compressed sensing, but suggests worse performance in identifying the relevant basis functions because of its nature of only making a locally optimal choice at each optimization step. For large dictionaries, the overlap matrix S will become singular, which can be addressed by essentially separating the Stiefel manifold into a null space component below a certain eigenvalue threshold of S and its complement[7].

B MISCELLANEOUS

B.1 Abstract

The present thesis represents a cumulative dissertation and is based on four publications that highlight the connections of signal processing with complex quantum systems from different angles.

On one side, large quantum systems require exponentially large parameter spaces for a complete description, calling for sophisticated techniques for handling the multitude of data. The compressed sensing paradigm and how to adapt it for practical quantum state estimation from experimental measurement data involving model selection methods constitutes one main topic. A further topic pertains to tensor networks in combination with spectral estimation protocols. This yields a less general procedure than compressed sensing regarding the range of applicable quantum states, however, substantially larger systems can be handled. In the course of the underlying publications of this thesis, both approaches have successfully been applied within concrete experiments.

Complementarily, quantum systems can be used to massively accelerate classical algorithms by means of future quantum computers. Due to their fundamentally different architecture, it is not straightforward to devise quantum analogues of classical algorithms. By designing novel algorithmic building blocks, a quantum algorithm for spectral estimation with superpolynomial speedup towards classical algorithms could be developed. The four major publications are supplemented by coauthored publications in the appendix.

Die vorliegende Arbeit ist kumulativ angelegt und basiert auf vier Veröffentlichungen, die die Verbindungen von Signalverarbeitung und komplexen Quantensystemen aus verschiedenen Blickwinkeln beleuchten.

Einerseits benötigen große Quantensysteme exponentiell große Parameterräume für eine vollständige Beschreibung, was ausgefeilte Techniken zur Handhabung der Datenmengen erfordert. Ein Schwerpunkt liegt auf dem Compressed-Sensing-Paradigma und wie dieses für die praktische Rekonstruktion von Quantenzuständen aus experimentellen Daten unter Einsatz von Model-Selection-Methoden angepasst zu werden hat. Ein weiterer Schwerpunkt liegt auf Tensornetzwerken in Verbindung mit Spektralanalyse-Methoden. Was die Breite der handhabbaren Quantenzustände betrifft, ist dieser Ansatz weniger allgemein als Compressed Sensing, allerdings kann er weitaus größere Systeme verarbeiten. Im Rahmen der dieser Arbeit zugrunde liegenden Veröffentlichungen wurden beide Ansätze erfolgreich innerhalb konkreter Experimente angewendet.

Komplementär hierzu können Quantensysteme auch eingesetzt werden, um klassische Algorithmen erheblich zu beschleunigen – mithilfe zukünftiger Quantencomputer. Durch ihren grundsätzlich andersartigen Aufbau ist es nicht einfach, analoge Quantenfassungen von Algorithmen zu entwickeln. Durch die Nutzbarmachung neuartiger algorithmischer Bausteine konnte ein Quantenalgorithmus zur Spektralanalyse mit superpolynomieller Beschleunigung gegenüber klassischen

Algorithmen ausgearbeitet werden. Die vier Hauptveröffentlichungen werden im Appendix ergänzt durch Publikationen, die in Koautorschaft entstanden sind.

B.2 The publications and the author's contributions

First author publications

[1] A. Steffens, C. A. Riofrío, W. McCutcheon, I. Roth, B. A. Bell, A. McMillan, M. S. Tame, J. G. Rarity, and J. Eisert, Experimentally exploring compressed sensing quantum tomography, *Quantum Science and Technology* 2:025005, 2017. See p. 17.

The author is the main contributor of this publication. He substantially conceived the structure of the project, carried out the underlying numerics and statistical analysis, significantly contributed to the manuscript, and created the figures.

[2] A. Steffens, C. A. Riofrío, R. Hübener, and J. Eisert, Quantum field tomography, *New Journal of Physics* 16:123010, 2014. See p. 38.

The author is the main contributor of this publication. He developed large parts of the tomographical protocol, implemented and tested the algorithm, wrote significant parts of the manuscript, and created the figures.

[3] A. Steffens, M. Friesdorf, T. Langen, B. Rauer, T. Schweigler, R. Hübener, J. Schmiedmayer, C. A. Riofrío, and J. Eisert, Towards experimental quantum-field tomography with ultracold atoms, *Nature Communications* 6:7663, 2015. See p. 73.

As the main contributor of this publication, the author substantially developed the underlying tomographical protocol and its numerical implementation, and considerably contributed to the creation of the manuscript and the figures.

[4] A. Steffens, P. Rebentrost, I. Marvian, J. Eisert, and S. Lloyd, An efficient quantum algorithm for spectral estimation, *New Journal of Physics* 19:033005, 2017. See p. 83.

As the main contributor of this publication, the author conceived a quantum algorithm for solving an established classical problem, and significantly contributed to the development of the necessary quantum primitives and the manuscript.

Coauthored publications

[5] G. Haack, A. Steffens, J. Eisert, and R. Hübener,
Continuous matrix product state tomography of quantum transport experiments,
New Journal of Physics 17:113024, 2015. See p. 120.

The author took part in the develepment of the method, carried out the numerical analysis and contributed to the manuscript and the creation of the figures.

[6] P. Rebentrost, A. Steffens, and S. Lloyd,
Quantum singular value decomposition of non-sparse low-rank matrices,
ArXiv e-prints 1607.05404, 2016. See p. 136.

The author contributed to the development of technical parts and the applications of the underlying method, as well as the manuscript.

Erstautorenschaften

[1] Der Autor war federführend bei dieser Publikation. Er hat in erheblicher Weise die Struktur dieses Projekts konzipiert, die Numerik und statistische Auswertung ausgeführt und substantiell zum Manuskript beigetragen sowie die Abbildungen erstellt.

[2] Der Autor war federführend bei dieser Publikation. Er hat große Teile der Tomographie-Prozedur entwickelt, den Algorithmus implementiert und getestet, erhebliche Teile des Manuskripts verfasst und die Abbildungen erstellt.

[3] Als federführend Mitwirkender dieser Publikation hat der Autor substantiell zur Entwicklung und numerischen Implementation der tomographischen Prozedur beigetragen, ebenso wie zur Erstellung des Manuskripts sowie der Abbildungen.

[4] Als federführend Mitwirkender dieser Publikation hat der Autor einen Quantenalgorithmus zur Lösung eines etablierten klassischen Problems konzipiert und substantiell zur Entwicklung der nötigen Methoden und des Manuskripts beigetragen.

Koautorenschaften

[5] Der Autor war an der Entwicklung der Methode beteiligt, führe die numerische Analyse durch und trug zum Manuskript und der Erstellung der Abbildungen bei.

[6] Der Autor trug sowohl zur Entwicklung von technischen Teilen als auch Anwendungen der Methode bei, sowie zur Erstellung des Manuskripts.

B.3 List of figures

Selbstständigkeitserklärung

Hiermit bestätige ich, dass ich die vorliegende Arbeit selbstständig und nur mit Hilfe der angegebenen Hilfsmittel angefertigt habe. Alle Stellen der Arbeit, die wörtlich oder sinngemäß aus Veröffentlichungen oder aus anderweitigen fremden Quellen entnommen wurden, sind als solche kenntlich gemacht. Ich habe die Arbeit nicht in einem früheren Promotionsverfahren eingereicht.

Adrian Steffens