The reconstruction of the state of a multipartite quantum mechanical system represents a fundamental task in quantum information science. At its most basic, it concerns a state of a bipartite quantum system whose subsystems are subjected to local operations. We compare two different methods for obtaining the original state from the state resulting from the action of these operations. The first method involves quantum operations called Petz recovery maps, acting locally on the two subsystems. The second method is called matrix (or state) reconstruction and involves local, linear maps which are not necessarily completely positive. Moreover, we compare the quantities on which the maps employed in the two methods depend. We show that any state which admits Petz recovery also admits state reconstruction. However, the latter is successful for a strictly larger set of states. We also compare these methods in the context of a finite spin chain. Here, the state of a finite spin chain is reconstructed from the reduced states of a few neighbouring spins. In this setting, state reconstruction is the same as the MPO (i.e. matrix product operator) reconstruction proposed by Baumgratz et al. [1]. Finally, we generalize both these methods so that they employ long-range measurements instead of relying solely on short-range correlations embodied in such local reduced states. Long-range measurements enable the reconstruction of states which cannot be reconstructed from measurements of local few-body observables alone and hereby we improve existing methods for quantum state tomography of quantum many-body systems.
1. Introduction

Consider a bipartite quantum state $\rho_{X Y}$ which is transformed to a state $\tau_{X' Y'}$ under the action of local quantum operations $N_X : X \to X'$ and $N_Y : Y \to Y'$. These local operations could either correspond to (i) undesirable noise (resulting from unavoidable interactions of the quantum
system $X$Y with its environment) or they could correspond to (ii) local measurements made by an experimenter doing quantum state tomography. We are interested in determining the conditions under which the state $\tau_{X^\prime Y^\prime}$ can be transformed back to the original state $\rho_{XY}$ with maps which act locally on $X^\prime$ and $Y^\prime$. In the case (i), these would be the conditions under which the effect of the noise can be reversed, whereas in the case (ii) these would be the conditions under which reconstruction of the original state from the outcome of the experimenter’s chosen measurements is possible.

The question whether $\tau_{X^\prime Y^\prime}$ can be transformed back to $\rho_{XY}$ can be answered with different methods. If the transformation is to be achieved with quantum operations, an answer is provided by the Petz recovery map [2, 3] under a condition on the mutual information of the two states. If general linear (not necessary completely positive) maps are allowed in the transformation, one can use a matrix reconstruction method. This matrix reconstruction method is related to MPO (i.e. matrix product operator) reconstruction and so-called pseudoskeleton (or CUR) matrix decompositions [1, 4, 5]. In either case, the construction of the maps which transform $\tau_{X^\prime Y^\prime}$ into $\rho_{XY}$ does not require complete information of $\rho_{XY}$ if suitable maps $\mathcal{N}_X$ and $\mathcal{N}_Y$ are used. In this case, the transformation can be used for efficient quantum state tomography of $\rho_{XY}$ with less measurements than necessary for standard quantum state tomography.

A fundamental quantity in quantum information theory is the quantum relative entropy $D(\rho\|\sigma)$ between a state $\rho$ and a positive semi-definite operator $\sigma$ (see Section 2.3 for its definition). It acts as a parent quantity for other entropic quantities arising in quantum information theory, e.g. von Neumann entropy, conditional entropy and mutual information. When $\rho$ and $\sigma$ are both states, $D(\rho\|\sigma)$ also has an operational interpretation as a measure of distinguishability between the two states [6, 7]. One of its most important properties is its monotonicity under the joint action of a quantum operation (say, $\mathcal{N}$). This is also called the data processing inequality (DPI) and is given by,

$$D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)) \leq D(\rho\|\sigma).$$

The condition under which the above inequality is saturated was obtained by Petz [8, 9] and has found important applications in quantum information theory. Petz proved that equality in the DPI holds if and only if there exists a recovery map, given by a quantum operation $\mathcal{R}$ which reverses the action of $\mathcal{N}$ on both $\rho$ and $\sigma$, i.e. $\mathcal{R}(\mathcal{N}(\rho)) = \rho$ and $\mathcal{R}(\mathcal{N}(\sigma)) = \sigma$. Petz also obtained an explicit form of such a recovery map, which is often called the Petz recovery map. Petz’s condition on the equality in the DPI immediately yields a necessary and sufficient under which the conditional mutual information $I(A : C|B)$ of a tripartite state $\rho_{ABC}$ is zero [3], which in turn is the condition under which strong subadditivity of the von Neumann entropy (arguably the most powerful entropic inequality in quantum information theory) is saturated. Petz’s result, when applied to the problem studied in this paper, implies that the original state $\rho_{XY}$ can be recovered from the transformed state $\tau_{X^\prime Y^\prime}$ if and only if the mutual
information $I(X : Y)_\rho$ of $\rho_{XY}$ is equal to the mutual information $I(X' : Y')_\tau$ of $\tau_{X'Y'}$ [3, 10]. Moreover, a valid recovery map is a tensor product of maps acting locally on $X'$ and $Y'$, each having the structure of a Petz recovery map. A detailed discussion of Petz’s result and of the quantities on which the Petz recovery maps depend is given in Section 2.3.

The data processing inequality of the relative entropy implies a DPI for the mutual information,

$$I(X' : Y')_\tau \leq I(X : Y)_\rho.$$

The mutual information quantifies the amount of correlations that exist between the two subsystems of a bipartite quantum state. Another measure of such correlations is the operator Schmidt rank [11, 12] of the state, which we denote as $\text{OSR}(X : Y)_\rho$ for a bipartite state $\rho_{XY}$ (see Eq. (2.8) for its definition).

In the following, we discuss the main results of this paper. We show that the operator Schmidt rank also satisfies a DPI:

$$\text{OSR}(X' : Y')_\tau \leq \text{OSR}(X : Y)_\rho,$$

where $\tau_{X'Y'}$ is the state obtained from $\rho_{XY}$ via the local quantum operations $N_X$ and $N_Y$, as discussed above. The DPI for the operator Schmidt rank is directly implied by the fact that the matrix rank satisfies $\text{rk}(MN) \leq \text{rk}(M) \text{rk}(N)$ for any two matrices $M$ and $N$ (see Corollary 3.9 for details). We show that $\tau_{X'Y'}$ can be transformed into $\rho_{XY}$ with local maps if and only if the DPI of the operator Schmidt rank is saturated. Our proof does not guarantee that the maps which transform $\tau_{X'Y'}$ into $\rho_{XY}$ are completely positive but it also does not require that $\rho_{XY}$ and $\tau_{X'Y'}$ are positive semidefinite or that $N_X$ and $N_Y$ are completely positive. The proof proceeds by transforming the reconstruction problem into a reconstruction problem for a general, rectangular matrix. Here, we provide an extension of the known pseudoskeleton decomposition [4, 5], which is also known as CUR decomposition and which can reconstruct a low-rank matrix from few of its rows and columns. Our method reconstructs a matrix $M$ from the matrix products $LM$ and $MR$ if the rank of $M$ equals the rank of $LMR$; $L$, $M$ and $R$ are general rectangular matrices.

We explore the relation between Petz recovery and state/MPO reconstruction for the case of 2, 4 and $n$ parties. State/MPO reconstruction, when compared to Petz recovery, is shown to be possible for a strictly larger set of states but requires more information.

The state of an $n$-partite quantum system, such as $n$ spins in a linear chain, can be represented as a matrix product operator (MPO) with MPO bond dimensions given by the operator Schmidt ranks $\text{OSR}(1 \ldots k : k + 1 \ldots n)$ (between the sites $1, \ldots, k$ and $k + 1, \ldots, n$; [13, 14]). If the operator Schmidt ranks are all bounded by a constant $D$, the MPO representation is given in terms of $\sim nD^2$ complex numbers, which is much less than the number of entries of the density matrix of the $n$-partite quantum system. Baumgratz et al. [1] presented a condition
under which an MPO representation of the state of an $n$-partite quantum system can be reconstructed from the reduced states of few neighbouring systems (MPO reconstruction). We will demonstrate that their work implies, for the case where the local operations $N_X$ and $N_Y$ are partial traces, that $\tau_{X'Y'}$ can be transformed into $\rho_{XY}$ if the two states have equal operator Schmidt rank.

The ability to reconstruct the state of an $n$-partite quantum system from reduced states of $l < n$ systems, as provided e.g. by MPO reconstruction, is advantageous for quantum state tomography of many-body systems. Standard quantum state tomography requires the expectation values of a number of observables which grows exponentially with $n$. If the full state can be reconstructed from $l$-body reduced states, then the number of observables grows exponentially with $l$ but only linearly with the number of reduced states. MPO reconstruction uses the reduced states of blocks of $l$ neighbouring sites on a linear chain. As the number of such blocks increases linearly with $n$, MPO reconstruction enables quantum state tomography with a number of observables which increases only linearly with $n$.

We call a method for quantum state tomography efficient if it requires only polynomially many (in $n$) sufficiently simple observables (more details on permitted observables are given in Section 6.1, Remark 6.3). Here we assume that exact expectation values are available. For a given method to be useful in practice, it is however necessary that the quantum state can be estimated up to a fixed estimation error using approximate expectation values from measurements on at most polynomially many (in $n$) copies of the state. In this paper, we discuss only the number of necessary observables but not the number of necessary copies of the state. Numerical simulations indicate that e.g. MPO reconstruction and similar methods are efficient also in the number of necessary copies [1, 15, 16, 17].

There are multipartite quantum states (e.g. states of a spin chain) which admit an efficient matrix product state (MPS) or MPO representation but which cannot be reconstructed from reduced states of a few of its parties (e.g. a few neighbouring sites of the spin chain). The $n$-qubit GHZ state is an example of such a state (Section 5.1). However, it has been shown that the GHZ state can be reconstructed from a number of observables linear in $n$, provided global observables (i.e. those which act on the whole system) are allowed [15, 16]. The necessary observables are given by simple tensor products [16] or simple tensor products and unitary control of few neighbouring sites [15]. We generalize MPO reconstruction and a similar technique based on the Petz recovery map [18] to use a certain class of long-range measurements which includes those just mentioned as special cases (Section 6). We represent a long-range measurement as a sequence of local quantum operations followed by the measurement of a local observable. However, a tensor product of single-party observables, whose expectation value can be obtained by a simple, sequential measurement of the single-party observables, already constitutes an allowed long-range measurement.
The example of the GHZ state shows that long-range measurements enable the recovery or reconstruction of a larger set of states than those obtained by local few-body observables. Our reconstruction and recovery methods provide a representation of the reconstructed state in terms of a sequence of local linear maps which is equivalent to an MPO representation. For methods based on the Petz recovery map, the local linear maps are quantum operations and, because of this, a PMPS (locally purified MPS) representation can be obtained ([19], Appendix A.5). A PMPS representation is advantageous because it can be computationally demanding to determine whether a given MPO representation represents a positive semidefinite operator [19] whereas a PMPS representation always represents a positive semidefinite operator. Our work on the reconstruction of spin chain states is partially based on similar ideas developed in the context of tensor train representations [20] and there is also related work on Tucker and hierarchical Tucker representations [5, 21, 22].

The remainder of the paper is structured as follows: Section 2 introduces notation, definitions, MPS/MPO representations and known results on the Petz recovery map. Section 3 shows how a low-rank matrix reconstruction technique enables bipartite state reconstruction, i.e. a transformation of $\tau_{XY}$ into $\rho_{XY}$. We also prove that approximate matrix reconstruction is possible if a low-rank matrix is perturbed by a small high-rank component (Section 3.2). We apply the Petz recovery map to the bipartite setting in Section 4 and investigate the relation between Petz recovery and state reconstruction in Section 5. Any state which admits Petz recovery is found to also admit state reconstruction. In Section 6, we discuss reconstruction of spin chain states with Petz recovery maps and state reconstruction. If reconstruction is performed with local reduced states (Section 6.1), a known application of the Petz recovery map [18] and the known MPO reconstruction technique [1] are obtained. In Section 6.2, we reconstruct spin chain states from recursively defined long-range measurements. We show that successful recovery of a given spin chain state implies successful reconstruction both for local reduced states and for long-range measurements. The set of states which can be reconstructed with long-range measurements is seen to be strictly larger than the set of states which can be reconstructed with measurements on local reduced states. Long-range measurements were used in earlier work on the reconstruction of pure states [15] and we show that our methods can recover or reconstruct these states if the same long-range measurements are used.

2. Preliminaries

2.1. Notation and basic definitions

In this paper, all Hilbert spaces $\mathcal{H}$ are finite-dimensional. We use capital letters $A, B, C, \ldots$ to denote quantum systems with Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_C, \ldots$, and set $d_A = \dim(\mathcal{H}_A)$. 
For notational simplicity, we often use $A$ to denote both the system and its associated Hilbert space, when there is no cause for confusion. If $n$ systems are involved, we denote their Hilbert spaces by $\mathcal{H}_1, \ldots, \mathcal{H}_n$ and tensor products of the latter by $\mathcal{H}_{1 \ldots n} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$.

We denote the set of linear maps from $A$ to $B$ by $\mathcal{B}(A; B) = \mathcal{B}(\mathcal{H}_A; \mathcal{H}_B)$, and the set of linear operators on $A$ by $\mathcal{B}(A) \equiv \mathcal{B}(A; A)$. If tensor products are involved, we use the notation $\mathcal{B}(AB; CD) \equiv \mathcal{B}(A, B; C, D) \equiv \mathcal{B}(A \otimes B; C \otimes D)$. The trace of a linear operator $F \in \mathcal{B}(A)$ (or a square matrix $F \in \mathbb{C}^{m \times m}$) is denoted by $\text{Tr}(F)$. A quantum state (or density matrix) of a system $A$ is a positive semi-definite operator $\rho \in \mathcal{B}(A)$ with unit trace. Let $\mathcal{D}(A)$ denote the set of quantum states in $\mathcal{B}(A)$. In case of a pure quantum state $\rho = |\psi\rangle \langle \psi| \in \mathcal{H}_A$, we refer to both $\rho = |\psi\rangle \langle \psi|$ and $|\psi\rangle$ as the pure state. For any state $\rho \in \mathcal{D}(A)$, its von Neumann entropy is defined as $S(\rho) = -\text{Tr}(\rho \log(\rho))$. In this paper, all logarithms are taken to base 2.

For any $F \in \mathcal{B}(A)$, let $F^\dagger$ denote its Hermitian adjoint, $\text{supp}(F)$ its support, $\text{rk}(F)$ its rank, $\|F\|$ its operator norm (largest singular value) and $\sigma_{\text{min}}(F)$ its smallest non-zero singular value. A linear operator $F \in \mathcal{B}(A)$ is an observable if it is Hermitian. The Hilbert–Schmidt inner product on $\mathcal{B}(A)$ is denoted by

$$\langle F, G \rangle = \text{Tr}(F^\dagger G) \quad \forall F, G \in \mathcal{B}(A). \quad (2.1)$$

The vector space $\mathcal{B}(A)$ becomes a Hilbert space when equipped with this inner product.

The notation $\mathcal{B}(\mathcal{B}(A); \mathcal{B}(B))$ denotes the set of linear maps from $\mathcal{B}(A)$ to $\mathcal{B}(B)$. This includes the set of quantum operations (or superoperators) from $A$ to $B$ which are given by linear, completely positive, trace-preserving (CPTP) linear maps $N \in \mathcal{B}(\mathcal{B}(A); \mathcal{B}(B))$. We use the shorthand notation $\mathcal{N} : A \rightarrow B$ to indicate such a quantum operation. Given any linear map $N \in \mathcal{B}(\mathcal{B}(A); \mathcal{B}(B))$, its Hermitian adjoint (with respect to the Hilbert–Schmidt inner product) is denoted by $N^\dagger \in \mathcal{B}(\mathcal{B}(B); \mathcal{B}(A))$, i.e., $\langle F, N(G) \rangle = \langle N^\dagger(F), G \rangle$, for all $F \in \mathcal{B}(B)$, $G \in \mathcal{B}(A)$.

Since we are dealing with finite-dimensional Hilbert spaces, all linear operators and maps are represented by matrices. Given a matrix $M \in \mathbb{C}^{m \times n}$ (or a linear map $M$), let $M^\dagger$ denote its conjugate transpose matrix, $\overline{M}$ denotes its (element-wise) complex conjugate matrix, and $M^*$ denotes its Moore–Penrose pseudoinverse. The following four properties of the pseudoinverse also define it uniquely [23, 24]:

$$MM^*M = M, \quad M^*MM^* = M^*, \quad (MM^*)^* = MM^*, \quad (M^*M)^* = M^*M. \quad (2.2)$$

Given a real number $t \geq 0$, we define $M_t^+ = (M_t)^*$ where $M_t$ is obtained from $M$ by replacing its singular values which are smaller than or equal to $t$ by zero.

For a system $A$, we choose an operator basis $\{F_i^{(A)}\}_{i=1}^{d_A^2}$ which is orthonormal in the Hilbert–Schmidt inner product:

$$F_i^{(A)} \in \mathcal{B}(A), \quad \langle F_i^{(A)}, F_j^{(A)} \rangle = \delta_{ij}, \quad i, j \in \{1, \ldots, d_A^2\}. \quad (2.3)$$
Given a basis element $F_i^{(A)}$, we denote its dual element (in the Hilbert–Schmidt inner product) by $\tilde{F}_i^{(A)}$:

$$\tilde{F}_i^{(A)} \in \mathcal{B}(\mathcal{B}(A); \mathbb{C}): \quad G \mapsto \langle \tilde{F}_i^{(A)}, G \rangle = \langle F_i^{(A)}, G \rangle.$$ (2.4)

The identity map, $\text{id}$, on $\mathcal{B}(A)$ can then be expressed as

$$\text{id} = \sum_{i=1}^{d^2_A} F_i^{(A)} \tilde{F}_i^{(A)}.$$ (2.5)

This is nothing but the resolution of the identity operator for the Hilbert space $\mathcal{B}(A)$.

Consider a linear map $M \in \mathcal{B}(\mathcal{B}(Y); \mathcal{B}(X))$. Since the vector spaces $\mathcal{B}(XY)$ and $\mathcal{B}(\mathcal{B}(Y); \mathcal{B}(X))$ have the same finite dimension, $(d_X d_Y)^2 = d^2_X d^2_Y$, it is possible to define a bijective linear map between the two spaces. To do so, we define the components of $\rho \in \mathcal{B}(XY)$ and $M$ in terms of the operator bases from above:

$$[\rho]_{ij} = \langle F_i^{(X)} \otimes F_j^{(Y)}, \rho \rangle, \quad [M]_{ij} = \langle F_i^{(X)}, M(F_j^{(Y)}) \rangle.$$ (2.6)

Given a linear operator $\rho \in \mathcal{B}(XY)$, we define a linear map $M_\rho$ by

$$[M_\rho]_{ij} = [\rho]_{ij}, \quad M_\rho \in \mathcal{B}(\mathcal{B}(Y); \mathcal{B}(X)).$$ (2.7)

We denote the matrix representation of $M_\rho$ in the operator basis chosen above by $M_\rho$. The maps $\rho \mapsto M_\rho$ and $\rho \mapsto M_\rho$ defined by Eq. (2.7) are of course bijective. Note that $\rho$ can be represented by a matrix of size $d_X d_Y \times d_X d_Y$ while $M_\rho$ can be represented by the matrix $M_\rho \in \mathbb{C}^{d^2_X \times d^2_Y}$. The transpose map $M^\top$ is defined in the same operator basis, i.e. $[M^\top]_{ij} = [M]_{ji}$.

Given a linear operator $\rho \in \mathcal{B}(XY)$, its operator Schmidt rank is given by

$$\text{OSR}(X : Y)_\rho = \min \left\{ r : \rho = \sum_{k=1}^{r} G_k' \otimes G_k'' \otimes G_k' \in \mathcal{B}(X), G_k'' \in \mathcal{B}(Y) \right\}.$$ (2.8)

The operator Schmidt rank is equal to

$$\text{OSR}(X : Y)_\rho = \text{rk}(M_\rho),$$ (2.9)

which can be shown as follows: The matrix representation $M_\rho$ of $M_\rho$ can be written as

$$M_\rho = GH \quad \text{where} \quad G \in \mathbb{C}^{d_X \times s}, \quad H \in \mathbb{C}^{s \times d_Y} \quad \text{and} \quad s = \text{rk}(M_\rho) = \text{rk}(M_\rho).$$ (2.10)

Since the components of $M_\rho$ and $\rho$ are related by $[M_\rho]_{ij} = [\rho]_{ij}$, we have

$$\rho = \sum_{ij} [\rho]_{ij} F_i^{(X)} \otimes F_j^{(Y)} = \sum_{k=1}^{s} \left( \sum_i G_{ik} F_i^{(X)} \otimes H_{kj} F_j^{(Y)} \right).$$ (2.11)
where $G_{ik}$ and $H_{kj}$ are the components of the matrices $G$ and $H$. This shows that the operator Schmidt rank cannot exceed $s = \text{rk}(M_{\rho})$. Now suppose that the operator Schmidt rank was less than that, i.e. $r = \text{OSR}(X : Y)_{\rho} < s = \text{rk}(M_{\rho})$. Then, a decomposition of $\rho$ as in Eq. (2.8) implies that

$$[M]_{ij} = [\rho]_{ij} = \sum_{k=1}^{r} G_{ik} H_{kj}, \quad G_{ik} = \langle f^{(X)}_i, G'_{k} \rangle, \quad H_{kj} = \langle f^{(Y)}_j, G''_{k} \rangle,$$

(2.12)

i.e. that $\text{rk}(M_{\rho}) \leq r < \text{rk}(M_{\rho})$. This contradiction shows that the operator Schmidt rank must equal $\text{rk}(M_{\rho}) = \text{rk}(M_{\rho})$. The operator Schmidt rank is also equal to the (smallest possible) bond dimension of a matrix product operator (MPO) representation of the linear operator [13]. This is discussed in Section 2.2.

### 2.2. MPS, MPO and PMPS representations

In this section, we introduce frequently-used efficient representations of pure and mixed quantum states on $n$ systems. We call a representation efficient if it describes a state with a number of parameters (i.e. complex numbers) which increases at most polynomially with $n$. The number of parameters of a particular representation of a state is accordingly given by the total number of entries of all involved vectors, matrices and tensors. For example, a pure state $|\psi\rangle \in \mathbb{C}^{d^n}$ of $n$ quantum systems of dimension $d$ has $d^n$ parameters and is not an efficient representation. To discuss whether a given representation is efficient or not, we use the following notation: For a function $f(n)$, we write $f = O(\text{poly}(n))$ or $f(n) = O(\text{poly}(n))$ if there is a polynomial $g(n)$ such that $f(n) \leq g(n)$. We write $f = O(\exp(n))$ if there are constants $c_1, c_2$ such that $f(n) \leq c_1 \exp(c_2 n)$.

First, we introduce the matrix product state (MPS) representation (see e.g. [14]), which is also known as tensor train (TT) representation [25]. Consider $n$ quantum systems of dimensions $d_1, \ldots, d_n$ respectively, and let $\{|\phi_{i_k}^{(k)}\rangle\}_{i_k=1}^{d_k}$ be an orthonormal basis of the $k$-th system. An MPS representation of a pure state on $n$ systems is given by

$$\langle \phi_{i_1}^{(1)} \ldots \phi_{i_n}^{(n)} | \psi \rangle = G_1(i_1) G_2(i_2) \ldots G_n(i_n)$$

(2.13)

where $D_0 = D_n = 1$, $G_k(i_k) \in \mathbb{C}^{D_{k-1} \times D_k}$ and $i_k \in \{1, \ldots, d_k\}$. The condition $D_0 = D_n = 1$ ensures that $G_1(i_1)$ and $G_n(i_n)$ are row and column vectors, while the $G_k(i_k)$ for $k$ between 1 and $n$ can be matrices. The matrix sizes $D_k$ are called the bond dimensions of the representation. The maximal local dimension and the maximal bond dimension are indicated by $d = \max_k d_k$ and $D = \max_k D_k$. For $d = O(\text{poly}(n))$ and $D = O(\text{poly}(n))$, the total number of parameters of the MPS representation is $ndD^2 = O(\text{poly}(n))$ and the representation is efficient. The bond dimension $D_k$ of any MPS representation of $|\psi\rangle$ is larger than or equal to the Schmidt rank of $|\psi\rangle$ for the bipartition $1 \ldots k|k+1 \ldots n$ and a representation with all bond dimensions equal
to the corresponding Schmidt ranks can always be determined (see, for example, [14]). We discuss the analogous property of the matrix product operator (MPO) representation in more detail.

A matrix product operator (MPO) representation [26, 27] of a mixed state on \( n \) systems is given by

\[
\langle \phi^{(1)}_{i_1} \ldots \phi^{(n)}_{i_n} | \rho | \phi^{(1)}_{j_1} \ldots \phi^{(n)}_{j_n} \rangle = G_1(i_1, j_1)G_2(i_2, j_2) \ldots G_n(i_n, j_n)
\]  

(2.14)

where \( D_0 = D_n = 1, G_k(i_k, j_k) \in \mathbb{C}^{D_k \times D_k} \) and \( i_k, j_k \in \{1, \ldots, d_k\} \). Alternatively, an MPO representation may be given in terms of operator bases \( F^{(k)}_{i_k} \):

\[
\langle F^{(1)}_{i_1} \otimes \ldots \otimes F^{(n)}_{i_n} | \rho \rangle = G_1(i_1)G_2(i_2) \ldots G_n(i_n)
\]

(2.15)

where \( D_0 = D_n = 1, G_k(i_k) \in \mathbb{C}^{D_{k-1} \times D_k} \) and \( i_k \in \{1, \ldots, d_k^2\} \). If the operator basis \( F^{(k)}_{i_k} = |\phi^{(k)}_{i_k} \rangle \langle \phi^{(k)}_{i_k}| \) is used, Eq. (2.15) turns into Eq. (2.14). As before, we denote the maximal local and bond dimensions by \( d = \max_k d_k \) and \( D = \max_k D_k \). The number of parameters of an MPO representation is at most \( nd^2D^2 \) and it is an efficient representation if \( d = O(\text{poly}(n)) \) and \( D = O(\text{poly}(n)) \). The operator Schmidt ranks of \( \rho \) provide lower bounds to the bond dimensions of any MPO representation of \( \rho \) [13]:

\[
\text{OSR}(1 \ldots k : k + 1 \ldots n)_\rho \leq D_k
\]

(2.16)

This becomes clear if we rewrite Eq. (2.14) as follows:

\[
\rho = \sum_{b_0 \ldots b_n} H_1(b_0, b_1) \otimes H_2(b_1, b_2) \otimes \ldots \otimes H_n(b_{n-1}, b_n)
\]

(2.17)

where \( H_k(b_{k-1}, b_k) \in \mathbb{C}^{d_k \times d_k} \) and \( [H_k(b_{k-1}, b_k)]_{i_k,j_k} = [G_k(i_k, j_k)]_{b_{k-1},b_k} \). The sum runs over \( b_k \in \{1, \ldots, D_k\} \). It can also be shown that a representation with equality in Eq. (2.16) always exists (see, for example, [14]). If the linear operator \( \rho \) represented by an MPO is a quantum state, it is desirable to ensure that \( \rho \) is positive semi-definite. However, deciding whether a given MPO represents a positive semi-definite operator is an NP-hard problem in the number of parameters of the representation [19], i.e. a numerical solution in polynomial (in \( n \)) time may not be obtained. As an alternative, one can use a PMPS (locally purified MPS) representation of the mixed state. A PMPS representation represents a positive semidefinite linear operator by definition. PMPS representations are also called evidently positive representations and they are introduced in Appendix A.5.

Suppose that a quantum state \( \rho \in \mathcal{D}(\mathcal{H}_1 \ldots n) \) was prepared via quantum operations \( W_k : \mathcal{B}(\mathcal{H}_{k-1}) \rightarrow \mathcal{B}(\mathcal{H}_{k-1,k}) \), i.e.

\[
\rho = W_n W_{n-1} \ldots W_3 W_2 (\sigma)
\]

(2.18)
where $\sigma \in D(H_1)$. Clearly, this is an efficient representation of the quantum state $\rho$ as it is described by at most $n d^n$ parameters. It is known that such a representation can be efficiently, i.e. with at most $\text{poly}(n)$ computational time, converted into an MPO representation or a PMPS representation [18, 19]. Appendix A.5 provides the details of the conversion and of the PMPS (locally purified MPS) representation. The state recovery and reconstruction techniques presented in Section 6 provide a representation of the reconstructed state which is similar to Eq. (2.18). Lemma A.4 in the appendix provides PMPS and MPO representations of the recovered state for techniques based on the Petz recovery map and an MPO representation of the reconstructed state for state reconstruction results.

2.3. The Petz recovery map

The (quantum) relative entropy, for two quantum states $\rho, \sigma \in D(A)$, was defined by Umegaki [28] as

$$D(\rho \parallel \sigma) = \text{Tr}(\rho \log(\rho) - \rho \log(\sigma)),$$

(2.19)

if $\text{supp}\, \rho \subseteq \text{supp}\, \sigma$, and is set equal to $+\infty$ otherwise. For a bipartite quantum state $\rho = \rho_{AB} \in D(AB)$, the mutual information between the subsystems $A$ and $B$ is defined in terms of the von Neumann entropies of $\rho_{AB}$ and its reduced states $\rho_{A} = \text{Tr}_B \rho_{AB}$ and $\rho_{B} = \text{Tr}_A \rho_{AB}$:

$$I(A : B)_{\rho} = S(A)_{\rho} + S(B)_{\rho} - S(AB)_{\rho}.$$

(2.20)

It can also be expressed in terms of the relative entropy as follows:

$$I(A : B)_{\rho} = D(\rho \parallel \rho_A \otimes \rho_B).$$

(2.21)

The quantum conditional mutual information (QCMI) of a tripartite quantum state $\rho \in D(ABC)$ is given by

$$I(A : C|B) := I(A : BC) - I(A : B) = I(AB : C) - I(B : C),$$

(2.22)

and is expressed in terms of the von Neumann entropy as follows:

$$I(A : C|B) = S(AB) + S(BC) - S(ABC) - S(B).$$

(2.23)

As mentioned in the Introduction, a fundamental property of the quantum relative entropy is its monotonicity under quantum operations. This is given by the data processing inequality (DPI): for quantum states $\rho, \sigma \in D(A)$ and a quantum operation $N$ acting on $D(A)$,

$$D(\rho \parallel \sigma) \geq D(N(\rho) \parallel N(\sigma)).$$

(2.24)
For the choice \( \rho = \rho_{ABC}, \sigma = \rho_{AB} \otimes \rho_{C} \) and \( N = \text{Tr}_A \otimes \text{id}_{BC} \), the DPI (2.24) implies that the QCMI of a tripartite state \( \rho_{ABC} \) is always non-negative. Using the definition (2.23) of the QCMI, we further infer that

\[
S(ABC) + S(B) \leq S(AB) + S(BC),
\]

which is the well-known strong subadditivity (SSA) property of the von Neumann entropy.

A necessary and sufficient condition for equality in the DPI (2.24) was derived in Hayden et al. [3] and Petz [2] and is stated in the following theorem.

**Theorem 2.1 (Petz recovery map)** Let \( \rho, \sigma \in D(A) \) be quantum states and \( N: A \rightarrow A' \) be a quantum operation. The equality

\[
D(N(\rho)||N(\sigma)) = D(\rho||\sigma)
\]

holds if and only if there is a linear CPTP map \( R: A' \rightarrow A \) which satisfies

\[
\rho = RN(\rho), \quad \text{and} \quad \sigma = RN(\sigma).
\]

If the above condition is satisfied, the so-called Petz recovery map \( R_{P,\sigma}^N \) satisfies the two equations. On the support of \( N(\rho) \) and for \( \omega \in \mathcal{B}(A') \), this map is given by

\[
R_{P,\sigma}^N(\omega) = \sigma^{1/2} N^* \left( N(\sigma)^{-1/2} \omega N(\sigma)^{-1/2} \right) \sigma^{1/2}
\]

where \( N^* \) is the Hermitian adjoint of \( N \) in the Hilbert–Schmidt inner product.

Further, Hayden et al. derived the following necessary and sufficient condition on the structure of tripartite states satisfying equality in the SSA (2.25) (see Theorem 6 in [3]).

**Theorem 2.2** Let \( \rho \equiv \rho_{ABC} \in D(ABC) \) be a tripartite quantum state. The equality \( I(A : BC)_\rho = I(A : B)_\rho \) (which is equivalent to equality in the SSA (2.25)) holds if and only if there is a decomposition of \( \mathcal{H}_B \) into \( \mathcal{H}_{B_{Lj}} \) and \( \mathcal{H}_{B_{Rj}} \) as

\[
\mathcal{H}_B = \bigoplus_j \mathcal{H}_{B_{Lj}} \otimes \mathcal{H}_{B_{Rj}},
\]

such that \( \rho \) can be written as

\[
\rho_{ABC} = \bigoplus_j p_j \rho_{AB_{Lj}} \otimes \rho_{B_{Rj}C},
\]

for a probability distribution \( \{p_j\} \), and sets of quantum states \( \{\rho_{AB_{Lj}}\} \) and \( \{\rho_{B_{Rj}C}\} \).
3. Reconstruction of bipartite states subjected to local operations

Let \( \rho_{XY} \in D(XY) \) be a bipartite state and let \( \tau_{X'Y'} \in D(X'Y') \) be a state obtained from \( \rho_{XY} \) by the action of local operations:

\[
\tau_{X'Y'} = (N_X \otimes N_Y) \rho_{XY},
\]

where \( N_X : X \rightarrow X' \) and \( N_Y : Y \rightarrow Y' \) denote quantum operations (or more generally, linear maps). We are interested in the conditions under which the original state \( \rho_{XY} \) can be reconstructed from \( \tau_{X'Y'} \) with local maps, i.e.

\[
\rho_{XY} = (R_{X'} \otimes R_{Y'}) (\tau_{X'Y'}). \tag{3.1}
\]

Our reconstruction scheme is particularly useful for states \( \rho \) with low operator Schmidt rank because then a reconstruction of \( \rho \) can be achieved with fewer measurements than required for standard quantum state tomography, as discussed in Remarks 6.3 and 6.14 (see also Section 5.1). The operator Schmidt rank of \( \rho \) is equal to the rank of the matrix \( M_\rho \) (Eqs. (2.7) and (2.9); \( M_\rho \) has size \( d_X^2 \times d_Y^2 \)). Hence, in Section 3.1 we first consider the more general problem of reconstruction of low-rank matrices (which are not necessarily states). Section 3.2 discusses the stability of our matrix reconstruction technique and Section 3.3 shows how it can be used to reconstruct a quantum state.

3.1. Reconstruction of low-rank matrices

Suppose that we want to obtain a matrix \( M \in \mathbb{C}^{m \times n} \) but we only know the entries of the matrix products \( LM \) and \( MR \) where \( L \) and \( R \) are \( r \times m \) and \( n \times s \) complex matrices. We refer to \( LM \), \( MR \) and \( LMR \) as the marginals of the matrix \( M \). Proposition 3.1 states that \( M \) can indeed be obtained from \( LM \) and \( MR \) if the condition \( \text{rk}(LMR) = \text{rk}(M) \) holds. This rank condition implies \( r, s \geq \text{rk}(M) \). If the rank of \( M \) is much smaller than its maximal value, \( \min\{m, n\} \), this provides a way to obtain \( M \) from \( LM \) and \( MR \) which, taken together, have much fewer entries than \( M \). If the matrices \( L \) and \( R \) are restricted to submatrices of permutation matrices, the matrix products \( LM \) and \( MR \) comprise selected rows and columns of \( M \). In this case, Proposition 3.1 provides a reconstruction of a low-rank matrix \( M \) from few rows and columns (cf. [4]).

**Proposition 3.1** Let \( L \in \mathbb{C}^{r \times m} \), \( M \in \mathbb{C}^{m \times n} \) and \( R \in \mathbb{C}^{n \times s} \) be matrices. Then

\[
\text{rk}(LMR) = \text{rk}(M) \iff \exists X \in \mathbb{C}^{s \times r} : M = MRXLM. \tag{3.2}
\]

If the condition is satisfied, \( M = MRXLM \) holds for any matrix \( X \) with \( XCX = C \), \( C = LMR \). The Moore–Penrose pseudoinverse \( X = (LMR)^+ \) has the required property \( XCX = C \).

Furthermore, \( \text{rk}(LM) = \text{rk}(M) \) implies \( \text{rk}(LMR) = \text{rk}(MR) \).
Proof “⇒” of Eq. (3.2): Assume that $\text{rk}(LMR) = \text{rk}(M)$ holds. The property $CXC = C$, $C = LMR$ implies that $LMRXu = u$ holds for all $u \in \text{im}(LMR)$. Let $q = \text{rk}(M) = \text{rk}(LMR)$. Let $u_1, \ldots, u_q$ be a basis of $\text{im}(LMR)$ and set $v_i = Xu_i$, $w_i = MR v_i$. The $v_i$ are linearly independent because $LMR v_i = LMXu_i = u_i$. The $w_i$ are linearly independent because $Lw_i = LMR v_i = u_i$. The $w_i$ is a linearly independent sequence of length $q = \text{rk}(M)$ and they satisfy $w_i \in \text{im}(M)$, i.e. they are a basis of $\text{im}(M)$. Now observe

$$MR XLw_i = MR Xu_i = MR v_i = w_i.$$ (3.3)

As a consequence, $MR XL$ maps any vector from $\text{im}(M)$ to itself. Accordingly, $(MR XL)M = M$ holds.

$\text{rk}(LM) = \text{rk}(M)$ implies $\text{rk}(LMR) = \text{rk}(MR)$: The equality $\text{rk}(LM) = \text{rk}(M)$ implies $M = M(LM)^+LM$ (use the “⇒” direction of Eq. (3.2) for $R = 1$). As a consequence, $MR = M(LM)^+LMR$ and $\text{rk}(MR) \leq \text{rk}(LMR)$ hold. The converse inequality $\text{rk}(LMR) \leq \text{rk}(MR)$ always holds and we arrive at $\text{rk}(LMR) = \text{rk}(MR)$.

“⇐” of Eq. (3.2): Assume that $M = MR XLM$ holds for some matrix $X$. The equality $M = MR XLM$ implies $\text{rk}(M) \leq \text{rk}(MR)$ and $\text{rk}(M) \leq \text{rk}(LM)$. The converse inequalities $\text{rk}(MR) \leq \text{rk}(M)$ and $\text{rk}(ML) \leq \text{rk}(M)$ always hold. As a consequence, we have $\text{rk}(LM) = \text{rk}(M)$ and $\text{rk}(MR) = \text{rk}(M)$. Above, we saw that the former equality implies $\text{rk}(LMR) = \text{rk}(MR)$ which, together with the latter equality $\text{rk}(MR) = \text{rk}(M)$, proves the theorem.

Remark 3.2 A violation of the rank condition $\text{rk}(LMR) = \text{rk}(M)$ does not in general imply that there is no method to obtain $M$ from $LM$ and $MR$. As a trivial example, consider $L = \mathbb{1}$ and $R = 0$. Then, the rank condition is violated for all $M \neq 0$, but $M$ is obtained trivially from $LM = M$. □

Remark 3.3 (Related work) Proposition 3.1 states that $M$ can be obtained from $LM$ and $MR$ if $\text{rk}(LMR) = \text{rk}(M)$ holds. Special cases of Proposition 3.1 have appeared before in several places. If $r = s = \text{rk}(M)$ and $L$ and $R$ select exactly $r = \text{rk}(M)$ rows and columns of $M$, the decomposition $M = MR(LMR)^{-1}LM$ is known as skeleton decomposition of $M$ [4]. Decompositions of the form $M = MR XLM$ where $L$ and $R$ select rows and columns of $M$ are known as pseudoskeleton/CUR decomposition of $M$ and it has been recognized that the truncated Moore–Penrose pseudoinverse $X = (LMR)^+$ may provide a good approximation if $r = s < \text{rk}(M)$ and suitable rows, columns and threshold $\tau$ are chosen [4]; we come back to the case of approximately low rank in Section 3.2. The case $r = s = \text{rk}(M)$, $X = (LMR)^+$ is contained in the results on tensor decompositions by Caiafa and Cichocki [22]. This matrix decomposition with $X = (LMR)^+$, restricted $L$ and $R$ but general $r, s \geq \text{rk}(M)$ forms the basis of MPO reconstruction [1] which is discussed in Section 6. □
3.2. Stability of the reconstruction under perturbation

Suppose that we have a matrix $S$ which satisfies the rank condition
\[ \text{rk}(LSR) = \text{rk}(S) \] (3.4)
for given matrices $L$ and $R$. We want to reconstruct the perturbed matrix
\[ M = S + E, \quad \epsilon = \|E\|. \] (3.5)
and $\epsilon = \|E\|$ is the operator norm of the perturbation. In Theorem 3.4 we provide a reconstruction $\tilde{M}_\tau$ and show that it is close to $M$ if the operator norm $\epsilon$ of the perturbation $E$ is small. A bound on the distance in operator norm between the reconstruction $\tilde{M}_\tau$ and $M$ is provided by
\[ \|\tilde{M}_\tau - M\| \leq \|\tilde{M}_\tau - S\| + \|S - M\| \leq \|\tilde{M}_\tau - S\| + \epsilon \] (3.6)
and Theorem 3.4 provides a bound on $\|\tilde{M}_\tau - S\|$.

Recall that given a matrix $M$, we define $M^+_\tau = (M^\tau)^+$ and $M^\tau$ is given by $M$ with singular values smaller or equal to $\tau$ replaced by zero.

Theorem 3.4 Let $M = S + E$, $\text{rk}(S) = \text{rk}(LSR)$, $\eta = \|L\|\|S\|\|R\|$, $\gamma = \sigma_{\min}(LSR)/\eta$, $\epsilon = \|E\|/\|S\|$. Let $\gamma > 2\epsilon$ and $\epsilon \leq \tau < \gamma - \epsilon$. Then
\[ \|\tilde{M}_\tau - S\| \leq \frac{\|S\|}{\gamma - \epsilon} \left( \frac{4\epsilon}{\gamma} + 2\epsilon + \epsilon^2 \right) \leq \frac{7\epsilon\|S\|}{\gamma(\gamma - \epsilon)}, \quad \tilde{M}_\tau = MR(LMR)^+_{\tau\gamma}LM. \]

Proof We prove the proposition for $\|L\| = \|S\| = \|R\| = 1$ (without loss of generality as explained in Appendix A.2). We have $M = S + E$, $\epsilon = \|E\|$, $\gamma = \sigma_{\min}(LSR)$ and $LMR = LSR + LER$ with $\|LER\| \leq \epsilon$. We insert $S = SR(LSR)^+LS$ and use Lemma 3.5 (provided at the end of this subsection):
\[ \|MR(LMR)^+_{\tau\gamma}LM - S\| \leq \|(LMR)^+_{\tau\gamma} - (LSR)^+\| + 2\|(LMR)^+_{\tau\gamma}\|\epsilon + \|(LMR)^+_{\tau\gamma}\|\epsilon^2 \leq \frac{4\epsilon}{\gamma(\gamma - \epsilon)} + \frac{2\epsilon}{\gamma - \epsilon} + \frac{\epsilon^2}{\gamma - \epsilon} \leq \frac{7\epsilon}{\gamma(\gamma - \epsilon)}. \]

Note that by premise, we have $\epsilon < \gamma < 1$. As a consequence, $1 \leq \frac{1}{\gamma}$ and $\epsilon \leq 1$ (which were used in the last equation) hold. This proves the theorem.

For the interpretation of the theorem, it is convenient to use the case with $\|L\| = \|S\| = \|R\| = 1$ and $\eta = 1$. Theorem 3.4 shows that the reconstruction $\tilde{M}_\tau$ reconstructs the low-rank component $S$ of $M = S + E$ up to a small error if the smallest singular value $\gamma$ of the low-rank component $LSR$ is much larger than the norm $\epsilon$ of the noise component. In addition, the
threshold $\tau$ must be chosen larger than the noise norm $\epsilon$ but smaller than $\gamma - \epsilon$. Appendix A.1 discusses examples which show that the bound from Theorem 3.4 is optimal up to constants and that the reconstruction error can diverge as $\epsilon$ approaches zero if small singular values in $LMR$ are not truncated.

Choosing a suitable threshold $\tau$ is equivalent to estimating the rank of the low rank contribution $S$. If the rank and support of $S$ are known, the measurements $L$ and $R$ can be chosen such that $LSR$ becomes invertible. For this special case, an upper bound on the reconstruction error has been given by Caiafa and Cichocki [22]. Their bound also includes constants which depend on $LSR$ and may diverge as $\gamma$ approaches zero. In Appendix A.3, we generalize their approach to our more general setting and obtain a bound which is similar to Theorem 3.4.

The following Lemma was used in the proof of Theorem 3.4:

**Lemma 3.5** Let $A, B, F \in \mathbb{C}^{m \times n}$. Let $\gamma$ equal $\sigma_{\text{min}}(A)$, $B = A + F$ with $\|F\| \leq \epsilon$. Let $\gamma > 2\epsilon$ and choose $\tau$ such that $\epsilon \leq \tau < \gamma - \epsilon$. Then $\sigma_{\text{min}}(B_\tau) \geq \gamma - \epsilon > \tau > \epsilon \geq \sigma_{r+1}(B) \geq \ldots \geq 0$.

**Proof** The singular values of $B = A + F$ satisfy (see e.g. [29])

$$|\sigma_i(A + F) - \sigma_i(A)| \leq \sigma_1(F) \leq \epsilon$$

and therefore, with $r = \text{rk}(A)$, we obtain

$$\sigma_1(B) \geq \ldots \geq \sigma_r(B) \geq \gamma - \epsilon > \tau > \epsilon \geq \sigma_{r+1}(B) \geq \ldots \geq 0.$$  

This shows already everything except the inequality in Eq. (3.10). To show the latter one, we use $\|X^+ - A^+\| \leq 2\|X^+\|\|A^+\|\|X - A\|$ [30, 29]. Inserting $\|B_\tau - A\| \leq \|B_\tau - B\| + \|B - A\|$ shows the desired inequality.

3.3. Reconstruction of bipartite states

Let $\rho \in \mathcal{D}(XY)$ be a bipartite quantum state and let $\tau \in \mathcal{D}(X'Y')$ be a state obtained from it by the action of local operations:

$$\tau = (N_X \otimes N_Y)\rho,$$

where $N_X : X \to X'$ and $N_Y : Y \to Y'$ denote quantum operations. We are interested in the conditions under which the original state $\rho$ can be reconstructed from $\tau$ with local quantum operations, i.e. $\rho = (R_X' \otimes R_Y')\tau$ with $R_X' : X' \to X$ and $R_Y' : Y' \to Y$. This question can be answered with the matrix decomposition from Section 3.1 without using the positivity properties of $\rho$, $N_{X,Y}$ and $R_{X,Y}$. The result is provided by the following Theorem:
Theorem 3.6 Let $\rho \in \mathcal{B}(XY)$ be a linear operator and let $N_X \in \mathcal{B}(\mathcal{B}(X);\mathcal{B}(X'))$, $N_Y \in \mathcal{B}(\mathcal{B}(Y);\mathcal{B}(Y'))$ be linear maps. Let $\tau = N_X \otimes N_Y(\rho)$. The original operator $\rho$ can be reconstructed from $\tau$ with local linear maps $R_{X'}$, $R_{Y'}$, i.e.

$$\rho = (R_{X'} \otimes R_{Y'})(\tau)$$

if and only if the following equality holds:

$$\text{OSR}(X' : Y')_{\tau} = \text{OSR}(X : Y)_{\rho}. \tag{3.13}$$

If the condition is satisfied, the following linear maps $R_{X'} \in \mathcal{B}(\mathcal{B}(X');\mathcal{B}(X))$ and $R_{Y'} \in \mathcal{B}(\mathcal{B}(Y');\mathcal{B}(Y))$ satisfy Eq. (3.12):

$$R_{X'} = R^M_{M_{ij} N_{ij}^T N_X}, \quad R_{Y'} = R^M_{M_{ij} N_{ij}^T N_Y}, \quad R^M_{L_{ij}N} = L(N L)^+. \tag{3.14}$$

The operators $(\text{id} \otimes N_Y)(\rho)$ and $(N_X \otimes \text{id})(\rho)$ are sufficient to construct the two $R^M_{L_{ij}N}$ maps if $N_X$ and $N_Y$ are known. The superscript $M$ indicates that the reconstruction map is based on matrix reconstruction. If the condition is satisfied, the following equation also holds for $R_{X'}$ and $R_{Y'}$ from Eq. (3.14):

$$\rho = (R_{X'} \otimes \text{id})(\rho) = (\text{id} \otimes R_{Y'})(\rho). \tag{3.15}$$

Remark 3.7 If the rank condition (3.13) is satisfied and $\sigma^{(XY)} = (\text{id} \otimes N_Y)(\rho)$ and $\sigma^{(X'Y')} = (N_X \otimes \text{id})(\rho)$ are given, one can obtain $\rho$ by computing the maps $R_{X'}$ and $R_{Y'}$ (Eq. (3.14)), $\tau = N_X \otimes \text{id}(\sigma^{(XY)})$ and $\rho = (R_{X'} \otimes R_{Y'})(\tau)$. More directly, one can also compute only the map $R_{X'}$ followed by computing $\rho = (R_{X'} \otimes \text{id})(\sigma^{(X'Y')})$ (Eq. (3.15)). The two options correspond to using either Eq. (3.16b) or Eq. (3.16c) to obtain $M_{ij}$. \hfill \Box

Proof (of Theorem 3.6) The operator $\tau$ is given by $\tau = N_X \otimes N_Y(\rho)$, therefore $\text{OSR}(X' : Y')_{\tau} \leq \text{OSR}(X : Y)_{\rho}$ always holds (Corollary 3.9). Let Eq. (3.12) hold. Again by Corollary 3.9, the converse inequality $\text{OSR}(X : Y)_{\rho} \leq \text{OSR}(X' : Y')_{\tau}$ also holds. As a consequence, the two operator Schmidt ranks must be equal.

Let the rank condition (3.13) hold. Lemma 3.8 and $\tau = N_X \otimes N_Y(\rho)$ We obtain:

$$[R_{X'} \otimes R_{Y'}(\tau)]_{ij} = [R_{X'} M_{ij} R_{Y'}^T]_{ij} \tag{3.16a}$$

$$= [M_{ij} N_{ij}^T (N_X M_{ij} N_Y^T)^*(N_X M_{ij} N_Y^T)^* N_X M_{ij}]_{ij} \tag{3.16b}$$

$$= [M_{ij} N_{ij}^T (N_X M_{ij} N_Y^T)^* N_X M_{ij}]_{ij} \tag{3.16c}$$

$$= [M_{ij}]_{ij}. \tag{3.16d}$$

In Eq. (3.16a) and Eq. (3.16b), we used Lemma 3.8 and inserted the maps from Eq. (3.14). In Eq. (3.16c), we used the property $A^*AA^* = A^*$ of the Moore–Penrose pseudoinverse. In Eq. (3.16d), we applied the matrix reconstruction result from Proposition 3.1. The therefor
needed rank condition \( \text{rk}(N_X M \rho N_Y^\dagger) = \text{rk}(M \rho) \) is equivalent to the rank condition (3.13) because of \( \tau = N_X \otimes N_Y(\rho) \) and Lemma 3.8. This shows that Eq. (3.12) holds if Eq. (3.13) is assumed and the maps from Eq. (3.14) are inserted. Eq. (3.15) can be shown by omitting \( R_y \) (or \( R_{x^c} \)) from left hand side of Eq. (3.16a). This finishes the proof of the theorem.

The remainder of the section provides the ingredients used in the preceding proof. It also provides a data processing inequality (DPI) for the operator Schmidt rank which is used below.

**Lemma 3.8** Let \( \rho \in \mathcal{B}(X Y) \) be a linear operator and let \( N_X : \mathcal{B}(X) \to \mathcal{B}(X') \), \( N_Y : \mathcal{B}(Y) \to \mathcal{B}(Y') \), be linear maps. Set \( \tau = (N_X \otimes N_Y)(\rho) \). Then

\[
[M_{\tau}]_{ij} = [N_X M_{\rho} N_Y^\dagger]_{ij}. \tag{3.17}
\]

**Proof** First, note that

\[
[(F_k^{(X)} F_j^{(X)}) \otimes (F_i^{(Y)} F_l^{(Y)})](\rho) = (F_k^{(X)} \otimes F_i^{(Y)})(F_k^{(X)} \otimes F_l^{(Y)})(\rho)
\]

\[
= (F_k^{(X)} \otimes F_i^{(Y)})[\rho]_{kl}. \tag{3.18a}
\]

The proof involves several basic steps:

\[
[M_{\tau}]_{ij} = [\tau]_{ij} = [N_X \otimes N_Y(\rho)]_{ij} \tag{3.19a}
\]

\[
= \langle F_k^{(X)} \otimes F_l^{(Y)}, N_X \otimes N_Y(\rho) \rangle \tag{3.19b}
\]

\[
= \sum_{kl} \langle F_i^{(X)} \otimes F_j^{(Y)}, (N_X \otimes N_Y)(F_k^{(X)} F_l^{(Y)})(\rho) \rangle \tag{3.19c}
\]

\[
= \sum_{kl} \langle F_i^{(X)}, N_X(F_k^{(X)}) \rangle \langle F_j^{(Y)}, N_Y(F_l^{(Y)}) \rangle [\rho]_{kl} \tag{3.19d}
\]

\[
= \sum_{kl} [N_X]_{ik} [N_Y]_{jl} [M_{\rho}]_{kl} \tag{3.19e}
\]

\[
= [N_X M_{\rho} N_Y^\dagger]_{ij}. \tag{3.19f}
\]

In the introduction, we saw that the operator Schmidt rank is given by \( \text{OSR}(X : Y)_\rho = \text{rk}(M_\rho) \) where \( M_\rho \in \mathcal{B}(\mathcal{B}(Y); \mathcal{B}(X)) \) is a linear map. As corollary from Lemma 3.8, we obtain the monotonicity of the operator Schmidt rank under local maps, i.e. a data processing inequality:

**Corollary 3.9** Let \( \rho, N_X, N_Y \) and \( \tau = N_X \otimes N_Y(\rho) \) as in Lemma 3.8. Then,

\[
\text{OSR}(X' : Y'_\tau) \leq \text{OSR}(X : Y)_\rho \tag{3.20}
\]

and

\[
\text{OSR}(X' : Y'_\tau) \leq \min\{\text{rk}(N_X), \text{rk}(N_Y)\}. \tag{3.21}
\]

**Proof** Use the property \( \text{OSR}(X : Y)_\rho = \text{rk}(M_\rho) \) (Eq. (2.9)), the identity \( \text{rk}(M_{\tau}) = \text{rk}(N_X M_\rho N_Y^\dagger) \) (Lemma 3.8) and the rank inequality \( \text{rk}(AB) \leq \min\{\text{rk}(A), \text{rk}(B)\} \) for arbitrary matrices or linear maps \( A \) and \( B \).
4. Petz recovery of bipartite states subjected to local quantum operations

In the previous section, we considered a linear operator \( \rho \in \mathcal{B}(XY) \) subjected to local linear maps \( \mathcal{N}_X \in \mathcal{B}(\mathcal{B}(X); \mathcal{B}(X')) \) and \( \mathcal{N}_Y \in \mathcal{B}(\mathcal{B}(Y); \mathcal{B}(Y')) \),

\[
\tau = (\mathcal{N}_X \otimes \mathcal{N}_Y)(\rho). \tag{4.1}
\]

In Section 3.3 we presented a condition under which \( \rho \) can be reconstructed from \( \tau \) via local linear maps. Here, we discuss the same question for a bipartite quantum state \( \rho \equiv \rho_{XY} \in \mathcal{D}(XY) \) and quantum operations \( \mathcal{N}_X \) and \( \mathcal{N}_Y \). The answer is obtained by restricting Theorem 2.1 to the bipartite setting, i.e. by inserting \( \rho = \rho_{XY}, \sigma = \rho_X \otimes \rho_Y \) and \( \mathcal{N} = \mathcal{N}_X \otimes \mathcal{N}_Y \) [10]:

**Corollary 4.1 (Bipartite Petz recovery map [10])** Let \( \rho \in \mathcal{B}(XY) \) a quantum state and \( \mathcal{N}_X : X \rightarrow X', \mathcal{N}_Y : Y \rightarrow Y' \) quantum operations. Set \( \tau = (\mathcal{N}_X \otimes \mathcal{N}_Y)(\rho) \). The equality

\[
I(X' : Y')_\tau = I(X : Y)_\rho \tag{4.2}
\]

holds if and only if there are quantum operations \( \mathcal{R}_{X'} : X' \rightarrow X \) and \( \mathcal{R}_{Y'} : Y' \rightarrow Y \) which satisfy

\[
\rho = (\mathcal{R}_{X'} \otimes \mathcal{R}_{Y'})(\tau). \tag{4.3}
\]

If the condition is satisfied, the two Petz recovery maps \( \mathcal{R}_{X'} = \mathcal{R}_{\rho_{XY}, \mathcal{N}_X}^{P} \) and \( \mathcal{R}_{Y'} = \mathcal{R}_{\rho_{XY}, \mathcal{N}_Y}^{P} \) satisfy the equation.

In the next section, we explore the relation between bipartite state reconstruction (Theorem 3.6) and bipartite Petz recovery (Corollary 4.1).

5. Comparison of Petz recovery and state reconstruction

In this section, we compare Petz recovery with state reconstruction for a bipartite quantum state \( \rho \in \mathcal{D}(XY) \) subject to local quantum operations \( \mathcal{N}_X : X \rightarrow X' \) and \( \mathcal{N}_Y : Y \rightarrow Y' \):

\[
\tau = (\mathcal{N}_X \otimes \mathcal{N}_Y)(\rho). \tag{5.1}
\]

The reconstruction is to be achieved via local linear maps:

\[
\rho = (\mathcal{R}_{X'} \otimes \mathcal{R}_{Y'})(\tau). \tag{5.2}
\]

State reconstruction and the Petz recovery map both provide maps \( \mathcal{R}_{X'} \) and \( \mathcal{R}_{Y'} \) under the assumption of different conditions on \( \rho \) and \( \tau \) (Theorem 3.6 and Corollary 4.1). There is the following evident relation between state reconstruction and Petz recovery:
Method | $\mathcal{R}_{X'}$ depends on | $\mathcal{R}_{Y'}$ depends on
--- | --- | ---
Petz recovery | $\rho_X, N_X$ | $\rho_Y, N_Y$
Reconstruction | $(\text{id} \otimes N_Y)(\rho), N_X$ | $(N_X \otimes \text{id})(\rho), N_Y$

Table 1: Necessary input data for a recovery or reconstruction of a quantum state $\rho \in \mathcal{D}(XY)$ from $\tau = (N_X \otimes N_Y)(\rho)$. In both cases, $\rho$ is obtained from $\rho = (\mathcal{R}_{X'} \otimes \mathcal{R}_{Y'})(\tau)$.

**Theorem 5.1** Let $\rho \in \mathcal{D}(XY)$ a quantum state and let $N_X : X \to X'$, $N_Y : Y \to Y'$ quantum operations. Let $\tau = N_X \otimes N_Y(\rho)$. Then

$$I(X' : Y')_\tau = I(X : Y)_\rho \Rightarrow \text{OSR}(X' : Y')_\tau = \text{OSR}(X : Y)_\rho.$$  \hspace{1cm} (5.3)

The converse implication does not hold. The premise of the implication (5.3) is equivalent to Petz recovery being possible, i.e. there are CPTP maps $\mathcal{R}_{X'}$ and $\mathcal{R}_{Y'}$ which recover $\rho$ from $\tau$ (Corollary 4.1):

$$\rho = \mathcal{R}_{X'} \otimes \mathcal{R}_{Y'}(\tau).$$  \hspace{1cm} (5.4)

CPTP maps which recover $\rho$ in this way are given by the Petz recovery maps $\mathcal{R}_{X'} = \mathcal{R}_{X', N_X}^\rho$ and $\mathcal{R}_{Y'} = \mathcal{R}_{Y', N_Y}^\rho$ (see Theorem 2.1). The conclusion of the implication (5.3) is equivalent to state reconstruction being possible, i.e. there are linear maps $\mathcal{R}_{X'}$ and $\mathcal{R}_{Y'}$ which reconstruct $\rho$ from $\tau$ (Theorem 3.6):

$$\rho = \mathcal{R}_{X'} \otimes \mathcal{R}_{Y'}(\tau).$$  \hspace{1cm} (5.5)

Linear maps which recover $\rho$ in this way are given by the reconstruction maps $\mathcal{R}_{X'} = \mathcal{R}_{X', N_X}^\rho$ and $\mathcal{R}_{Y'} = \mathcal{R}_{Y', N_Y}^\rho$ (see Theorem 3.6).

**Proof** The premise of Eq. (5.3) implies that the CPTP maps from Eq. (5.4) exist (Corollary 4.1). These CPTP maps are linear maps which satisfy Eq. (5.5), which in turn implies that the conclusion of Eq. (5.3) holds (Theorem 3.6).

A counterexample for the converse implication of Eq. (5.3) will be provided in Section 5.1. □

**Remark 5.2** Suppose that the conclusion of Eq. (5.3) holds while its premise does not hold. If both linear maps $\mathcal{R}_{X'}$ and $\mathcal{R}_{Y'}$ were CPTP, the equality $I(X' : Y')_\rho \leq I(X : Y)_\tau$ would be implied by Eq. (5.5) and $I(X' : Y')_\rho = I(X : Y)_\tau$ would follow (since the converse inequality always holds because of $\tau = (N_X \otimes N_Y)(\rho)$. This would contradict our assumption; i.e. at least one of $\mathcal{R}_{X'}$ and $\mathcal{R}_{Y'}$ is not CPTP. For example, the reconstruction maps for the $W$ state on four qubits (Section 5.1) are non-positive. □

Theorem 5.1 implies that any state which admits Petz recovery also admits state reconstruction. Table 1 compares the quantities on which the state reconstruction maps and the Petz recovery maps depend: Both methods require knowledge of the quantum operations $N_X$ and $N_Y$. The
marginal states $\rho_X$ and $\rho_Y$ are sufficient for computing the Petz recovery maps. To compute the state reconstruction maps, the states $(\text{id} \otimes N_Y)(\rho) \in \mathcal{D}(XY')$ and $(N_X \otimes \text{id})(\rho) \in \mathcal{D}(X'Y)$ are needed. The marginals $\rho_X, \rho_Y, \tau_{X'} \equiv N_X(\rho_X)$ and $\tau_{Y'} \equiv N_Y(\rho_Y)$ (which are also needed for the recovery maps) can be inferred from the states $(\text{id} \otimes N_Y)(\rho)$ and $(N_X \otimes \text{id})(\rho)$. However, in addition, these states contain correlations between the systems $X$ and $Y'$ and between $X'$ and $Y$, respectively. This means that state reconstruction requires more input data for a reconstruction of $\rho$ than state recovery. On the other hand, Theorem 5.1 and the examples in the following subsection show that state reconstruction is successful for a strictly larger set of states than Petz recovery.

5.1. Comparison for four-partite systems

After the comparison of state reconstruction and Petz recovery for bipartite quantum systems, we apply this result to the more specific case of quantum systems which comprise four subsystems. Specifically, we consider four systems $A, B, C,$ and $D$. We insert the partial trace $\text{Tr}_A: AB \rightarrow B$ for $N_X: X \rightarrow X'$ and the partial trace $\text{Tr}_D: CD \rightarrow C$ for $N_Y: Y \rightarrow Y'$ in Theorem 5.1. Accordingly, reconstruction of $\rho \in \mathcal{D}(ABCD)$ from its reduced state $\rho_{BC} = \text{Tr}_{AD}(\rho)$ is achieved with maps $\mathcal{R}_B \in \mathcal{B}(\mathcal{B}(B); \mathcal{B}(AB))$ and $\mathcal{R}_C \in \mathcal{B}(\mathcal{B}(C); \mathcal{B}(CD))$ as

$$\rho = \mathcal{R}_B \otimes \mathcal{R}_C(\rho_{BC}).$$

(5.6)

A straightforward application of Theorem 5.1 provides the implication

$$I(B : C)_\rho = I(AB : CD)_\rho \quad \Rightarrow \quad \text{OSR}(B : C)_\rho = \text{OSR}(AB : CD)_\rho.$$  

(5.7)

Since the W state on four qubits satisfies the conclusion of the last Equation but not its premise it constitutes a counterexample for the converse implication. In addition, it provides a counterexample for the converse implication of Eq. (5.3) in Theorem 5.1. On $n$ qubits, the W state is given by

$$W_n = |w_n\rangle\langle w_n|, \quad |w_n\rangle = \frac{1}{\sqrt{n}} \left[ |10\ldots0\rangle + |010\ldots0\rangle + \cdots + |0\ldots01\rangle \right]$$

(5.8)

and the operator Schmidt rank and mutual information values of the four-qubit W state $W_{ABCD}$ are provided in Table 4 on Page 24.

Above, we presented one possible application of bipartite Petz recovery (Corollary 4.1) to a quadripartite system. It turns out that Petz recovery can be applied to a quadripartite system in three different ways. The first row of Table 2 corresponds to the application of state reconstruction and Petz recovery to a quadripartite system as presented above. Rows two and three of Table 2 present two different ways to apply Petz recovery to a quadripartite system. In total, we have one possible application of state reconstruction and three possible applications of Petz recovery and for each application, there is a condition for successful reconstruction/recovery. These conditions read as follows:
Table 2: Possible applications of state reconstruction and Petz recovery (Theorem 5.1) to a quadripartite system. Corresponding conditions for successful recovery are stated in Lemma 5.3.

| $X \rightarrow X'$ | $Y \rightarrow Y'$ | $N_X$ | $N_Y$ | Condition |
|---------------------|---------------------|--------|--------|-----------|
| $AB \rightarrow B$  | $CD \rightarrow C$  | Tr$_A$ | Tr$_D$ | (C1), (C2) |
| $AB \rightarrow B$  | $CD \rightarrow CD$ | Tr$_A$ | id     | (C3)      |
| $ABC \rightarrow BC$| $D \rightarrow D$   | Tr$_A$ | id     | (C4)      |

**Lemma 5.3** Given a quantum state $\rho \in \mathcal{D}(ABCD)$, consider the equations

\[
\begin{align*}
\text{OSR}(B : C)_\rho &= \text{OSR}(AB : CD)_\rho \quad \text{(C1)} \\
I(B : C)_\rho &= I(AB : CD)_\rho \quad \text{(C2)} \\
I(A : B)_\rho &= I(A : BCD)_\rho \quad \text{(C3)} \\
I(A : BC)_\rho &= I(A : BCD)_\rho \quad \text{(C4)}
\end{align*}
\]

The following implications hold, but the converse implications do not:

\[
\begin{align*}
(C2) \Rightarrow (C1), \\
(C2) \Rightarrow (C3), \\
(C3) \Rightarrow (C4).
\end{align*}
\]

**(Proof)** The implication $(C2) \Rightarrow (C1)$ follows from Theorem 5.1 with the substitutions given in the first row of Table 2.

Eq. $(C2) \Rightarrow$ Eq. $(C3)$: The inequality $I(B : C)_\rho \leq I(B : CD)_\rho \leq I(AB : CD)_\rho$ always holds, therefore $I(B : C)_\rho = I(AB : CD)_\rho$ implies $I(B : CD)_\rho = I(AB : CD)_\rho$. The latter can be written with the conditional mutual information as $I(A : CD|B) = 0$ (Eq. (2.22)). The CMI in turn is also equal to $I(A : CD|B) = I(A : BCD) - I(A : B)_\rho$, which shows the desired equality $I(A : B) = I(A : BCD)$.

Eq. $(C3) \Rightarrow$ Eq. $(C4)$: The inequality $I(A : B)_\rho \leq I(A : BC)_\rho \leq I(A : BCD)_\rho$ always holds, therefore $I(A : B)_\rho = I(A : BCD)_\rho$ implies $I(A : BC)_\rho = I(A : BCD)_\rho$.

Table 4 contains states which show that the converse implications do not hold. The states are constructed from the $n$-qubit W state from Eq. (5.8) and the GHZ and classical GHZ states on $n$ qubits:

\[
\begin{align*}
\text{GHZ}_n &= |\text{GHZ}_n\rangle \langle \text{GHZ}_n|, \quad |\text{GHZ}_n\rangle = \frac{1}{\sqrt{2}} \left[ |0\ldots 0\rangle + |1\ldots 1\rangle \right] \\
c\text{GHZ}_n &= \frac{1}{2} \left[ |0\ldots 0\rangle |0\ldots 0\rangle + |1\ldots 1\rangle |1\ldots 1\rangle \right]
\end{align*}
\]

The values of the operator Schmidt rank and mutual information given in Table 4 show that the converse implications do not hold.
Figure 1: State reconstruction vs. Petz recovery for a quadripartite system. The figure shows the relations between conditions (C1)–(C4) from Lemma 5.3 and several states whose position indicates which of the conditions are satisfied by each state. e.g. \(W_{ABCD}\) satisfies (C1) but it does not satisfy (C2)–(C4). \(\rho_A\) and \(\rho_D\) denote arbitrary states while the W, GHZ and classical GHZ states are defined in Eqs. (5.8) and (5.10).

| Method          | Cond. | Input         | Domain/Range                   | Depends on   |
|-----------------|-------|---------------|--------------------------------|--------------|
| Reconstruction  | (C1)  | \(\rho_{BC}\) | \(B \rightarrow AB, C \rightarrow CD\) | \(\rho_{ABC}, \rho_{BCD}\) |
| Recovery        | (C2)  | \(\rho_{BC}\) | \(B \rightarrow AB, C \rightarrow CD\) | \(\rho_{AB}, \rho_{CD}\) |
| Recovery        | (C3)  | \(\rho_{BCD}\) | \(B \rightarrow AB\)             | \(\rho_{AB}\) |
| Recovery        | (C4)  | \(\rho_{BCD}\) | \(BC \rightarrow ABC\)           | \(\rho_{ABC}\) |

Table 3: Properties of the four reconstruction/recovery settings considered in Fig. 1. The last three columns Input, Domain/Range and Depends on refer to the reconstruction map(s). See Table 2 for further details.

The relations between Eqs. (C1) to (C4) from Lemma 5.3 are illustrated in Fig. 1. The figure also shows which of the conditions are satisfied by the example states from Table 4. For example, the W state \(W_{ABCD}\) on four qubits does not satisfy (C2)–(C4). We can understand that \(W_{ABCD}\) cannot satisfy (C4) by considering the following known result: If (C4) holds, then Theorem 2.2 tells us that the reduced state \(\rho_{AD}\) must be a separable state. However, the reduced state \(\text{Tr}_{BC}(W_{ABCD})\) has a non-positive semidefinite partial transpose and therefore is inseparable, i.e. entangled [31, 32]: The entanglement in the reduced state on \(AD\) mandates that Eq. (C4) is not satisfied.

Figure 1 illustrates that reconstruction and the different applications of Petz recovery work for different subsets of all quadripartite states but one should not forget that they also require different reduced states of \(\rho\) in order to recover \(\rho \in \mathcal{D}(ABCD)\). Table 3 shows the necessary reduced states for each case. In all four cases, the full state \(\rho \in \mathcal{D}(ABCD)\) can be reconstructed from marginal states on only two or three of the systems. Each scheme enables quantum state tomography with incomplete information (i.e. the necessary marginals) if the corresponding condition is assumed to hold. Each scheme also relies on the fact that correlations as measured by the operator Schmidt rank or the mutual information are less than maximal; this restriction...
Table 4: Log-operator Schmidt rank $\kappa(X : Y) = \log_2(\text{OSR}(X : Y))$ and mutual information $I(X : Y)$ of the states shown in Fig. 1. Equations (C1) to (C4) are defined in Lemma 5.3.

| $\kappa(B : C)$ | $\kappa(AB : CD)$ | $I(B : C)$ | $I(AB : CD)$ | $I(A : B)$ | $I(A : BC)$ | $I(A : BCD)$ | $C1$ | $C2$ | $C3$ | $C4$ |
|-----------------|-------------------|-------------|--------------|------------|-------------|--------------|------|------|------|------|
| $\rho_{A \otimes W_{BCD}}$ | 2 2 | $\approx 0.92$ | $\approx 1.84$ | 0 0 | 0 | ✓ ✓ ✓ ✓ | ✓ − ✓ ✓ |
| $W_{ABC} \otimes \rho_D$ | 2 2 | $\approx 0.92$ | $\approx 1.84$ | $\approx 0.92$ | $\approx 1.84$ | ✓ − − ✓ | ✓ − − − |
| $W_{ABCD}$ | 2 2 | $\approx 0.62$ | 2 | $\approx 0.62$ | 1 | $\approx 1.62$ | ✓ − − − |
| $\rho_{A \otimes \text{GHZ}_{BCD}}$ | 1 2 | 1 2 | 0 0 | 0 | − − ✓ ✓ | − − − − |
| $\text{GHZ}_{ABC} \otimes \rho_D$ | 1 2 | 1 2 | 1 2 | 2 | − − − ✓ | − − − − |
| $\text{GHZ}_{ABCD}$ | 1 2 | 1 2 | 1 1 | 2 | − − − − | − − − − |

Table 3 shows that the state $\rho \in D(ABCD)$ can be obtained from $\rho_{ABC}$ and $\rho_{BCD}$ with reconstruction under (C1) but recovery under (C2) requires only the marginal states on $AB$, $BC$ and $CD$. This prompts the question whether smaller marginal states are sufficient to reconstruct a state under the reconstruction condition (C1). For example, one could hope to obtain $\rho$ from $\rho_{AB}$ and $\rho_{BCD}$ but the following two states $\sigma_+$ and $\sigma_-$ show that this is not possible:

$$\sigma_\pm = \frac{1}{32} \left( 2 I \otimes 1 \otimes 1 + (1 \pm Z) \otimes Z \otimes Z \right) \otimes 1, \quad 1 = \left( \begin{array}{c} 1 \\ 0 \\ \end{array} \right), \quad Z = \left( \begin{array}{c} 0 \\ 1 \\ \end{array} \right).$$

(5.11)

The states $\sigma_+$ and $\sigma_-$ have the same marginals on $AB$ and $BCD$ but they do satisfy (C1). As a consequence, $\rho_{AB}$ and $\rho_{BCD}$ are not sufficient to obtain $\rho$ under (C1) and it is now apparent that $\rho_{ABC}$ and $\rho_{BCD}$ are necessary to reconstruct a state under that condition.

6. Efficient reconstruction of states on spin chains via recursively defined measurements

Under suitable conditions, the state of a linear spin chain with $n$ spins can be reconstructed from marginal states of few neighbouring spins with the Petz recovery map [18] or with state reconstruction [1]. In Section 6.1, we explore the relation between Petz recovery and state reconstruction in that setting. In Section 6.2, we generalize both techniques to use long-range measurements instead of or in addition to short-ranging correlations found in marginal states of few neighbouring spins. This allows for the efficient recovery/reconstruction of a larger set of states, as is explained in the following.

1 The exact values of the numerical constants are $2 \log(3) - \frac{4}{3} \approx 1.84, 4 - \frac{3}{2} \log(3) \approx 1.62, \log(3) - \frac{2}{3} \approx 0.92$ and $3 - \frac{1}{2} \log(3) \approx 0.62$.

2 The reduced states $\sigma_{\pm,AB} = \text{Tr}_{CD}(\sigma_\pm) = \frac{1}{4} I \otimes 1$ and $\sigma_{\pm,BCD} = \text{Tr}_{A}(\sigma_\pm) = \frac{1}{4} (2 I \otimes 1 + Z \otimes Z) \otimes 1$ do not depend on the sign. The values of the operator Schmidt rank are $\text{OSR}(AB : CD)_{\sigma_\pm} = \text{OSR}(B : C)_{\sigma_\pm} = 2$.

3 This holds true if $\rho$ is to be reconstructed from marginal states of $\rho$ which include at most $\rho_{ABC}$ and $\rho_{BCD}$.
Motivation for long-range measurements. Consider the following quantum states on $n$ qubits:

$$\text{GHZ}_{\alpha, n} = |\text{GHZ}_{\alpha, n}\rangle \langle \text{GHZ}_{\alpha, n}|, \quad |\text{GHZ}_{\alpha, n}\rangle = \frac{1}{\sqrt{2}} \left( |0 \ldots 0\rangle + e^{i\alpha} |1 \ldots 1\rangle \right)$$  \hspace{1cm} (6.1a)

$$\text{cGHZ}_n = \frac{1}{2} \left[ |0 \ldots 0\rangle \langle 0 \ldots 0| + |1 \ldots 1\rangle \langle 1 \ldots 1| \right]$$  \hspace{1cm} (6.1b)

All states from the set $S_n = \{\text{cGHZ}_n\} \cup \{\text{GHZ}_{\alpha, n} : \alpha \in \mathbb{R}\}$ have the same reduced state $\text{cGHZ}_k$ on $k < n$ qubits. No recovery or reconstruction method which receives only local reduced states as input can distinguish between the states from the set $S_n$ and this is also the reason why no method could recover or reconstruct the four-qubit state $|\text{GHZ}_{0, 4}\rangle = |\text{GHZ}_4\rangle$ in Section 5.1.

Note that the pure states $|\text{GHZ}_{\alpha, n}\rangle$ can be represented as an MPS with bond dimension two (because they are the superposition of two pure product states) and that all states from the set $S_n$ can be represented as an MPO with bond dimension at most four (because they are the sum of at most four tensor product operators) [14].

We call an MPS representation efficient if its bond dimension is at most $D = O(\text{poly}(n))$ and we call a tomography scheme efficient if expectation values of at most $\text{poly}(n)$ simple observables are needed; a possible definition of a simple observable is provided in Remark 6.3. Standard quantum state tomography is not efficient because it requires $\sim \exp(n)$ expectation values. In contrast, it has been shown that any pure state which admits an efficient MPS representation can be determined efficiently from observables with a simple structure.\(^4\) The tomography scheme from Ref. [15] is efficient for the states $|\text{GHZ}_{\alpha, n}\rangle$ but recovery/reconstruction methods based on local reduced states must fail for these states. In Section 6.2, we extend both Petz recovery and state reconstruction in a way which allows the long-range measurements from Ref. [15] to be used and thus the states $|\text{GHZ}_{\alpha, n}\rangle$ to be reconstructed successfully. What is more, we show that there are mixed states which cannot be reconstructed from local reduced states but can be reconstructed from long-range measurements (Remark 6.16). This shows that Petz recovery and state reconstruction with long-range measurements can reconstruct more states than prior techniques (recovery/reconstruction from local reduced states and the tomography scheme from Ref. [15]). Furthermore, state reconstruction can reconstruct any MPO of bond dimension $D$ from $\sim nD$ expectation values of global tensor product observables, as has been shown in related prior work [20]. We build upon that to show that successful, efficient Petz recovery with long-range measurements implies that efficient state reconstruction with long-range measurements is also possible (Theorem 6.17).

**Prior work: MPO reconstruction [1].** Many physically interesting quantum states $\rho \in \mathcal{D}(\mathcal{H}_{1 \ldots n})$ can be represented efficiently, i.e. with $	ext{poly}(n)$ parameters, via an MPO representation [1]. However, standard quantum state tomography requires $\sim \exp(n)$ different expectation

\(^4\) This is shown by the tomography scheme based on unitary operations introduced in Ref. [15]. We discuss it in more detail in Remark 6.8.
values in $\rho$ to reconstruct $\rho$, even if $\rho$ admits such an efficient MPO representation. As an improvement over that, it has been shown\(^5\) that almost all states with an MPO representation of bond dimension $D$ can be reconstructed from their reduced states on $\sim \log(D)$ neighbouring spins if a suitable reconstruction scheme is used [1]. We refer to this reconstruction scheme as \textit{MPO reconstruction} and we rederive it below in Theorem 6.2 as a consequence of our result on bipartite state reconstruction (Theorem 3.6).

\textbf{Prior work: Cross approximation of tensor trains [20].} Our generalization of state reconstruction to long-range measurements in Theorem 6.12 can be used to construct an MPO representation of the quantum state (Remark 6.14). An MPO representation of a quantum state $\rho \in \mathcal{D}(\mathcal{H}_1\cdots\mathcal{H}_n)$ is exactly the same as a tensor train representation of $\rho$ if the operator $\rho$ is regarded as vector from the tensor product vectorspace $\mathcal{B}(\mathcal{H}_1) \otimes \cdots \otimes \mathcal{B}(\mathcal{H}_n)$. Ref. [20] provides a means to reconstruct a tensor of low tensor train rank (i.e. an MPO of low bond dimension) from few entries. This procedure is called \textit{tensor train cross approximation}. When applied to quantum states, tensor train cross approximation allows for the reconstruction of a quantum state from the expectation values of few tensor product observables. Theorem 6.12 is more general because it admits more general measurements; e.g. it also permits the measurements introduced in Ref. [15] (cf. Remarks 6.8 and 6.18).

\textbf{Prior work: Markov entropy decomposition [18].} The strong subadditivity (SSA) property of the von Neumann entropy of a tripartite state $\rho \equiv \rho_{ABC}$ (cf. (2.25) of Section 2.3) can be expressed in terms of the conditional entropy $S(A|B)_\rho = S(AB)_\rho - S(B)_\rho$:

$$S(A|BC)_\rho \leq S(A|B)_\rho$$ (6.2)

If we choose arbitrary subsets $M_k \subset \{1\ldots k\}$, the entropy $S(\rho) = S(1\ldots n)_\rho$ can be rewritten and upper-bounded as follows:

$$S(1\ldots n)_\rho = S(12)_\rho + \sum_{k=2}^{n-1} S(k+1|1\ldots k)_\rho$$ \hspace{1cm} (6.3)

$$\leq S(12)_\rho + \sum_{k=2}^{n-1} S(k+1|M_k)_\rho =: S_M(\rho).$$ \hspace{1cm} (6.4)

In the second step, we applied Eq. (6.2) $n-2$ times. The sets $M_k$ are called \textit{Markov shields} and the upper bound $S_M(\rho)$ is called the \textit{Markov entropy} [18]. In the following, we consider the particular choice $M_k = \{k\}$. In that case, the conditional entropies $S(k+1|M_k)_\rho = S(k+1|k)_\rho$

---

\(^5\) Assume that all spins have the same dimension $d = d_k$, $k \in \{1, \ldots, n\}$. Lemma 1 in the supplementary material of [1] states that $\rho$ can be reconstructed from reduced states on $l+r+1$ neighbouring spins if a certain condition is satisfied. This condition can be satisfied only if both $D^2 \leq d^2 l$ and $D^2 \leq d^2 r$ hold. However, if these two inequalities are satisfied, the given conditions almost always hold for MPO matrices with random entries.
depend only on the reduced state $\rho_{k,k+1}$. As a consequence, the Markov entropy $S_M(\rho)$ is an upper bound on $S(\rho)$ which depends only on the nearest-neighbour reduced states $\rho_{k,k+1}$ ($k \in \{1, \ldots, n-1\}$). For a nearest-neighbour Hamiltonian $H = \sum_{k=1}^{n-1} h_{k,k+1}$ the energy $E = \text{Tr}(\rho H) = \sum_{k=1}^{n-1} \text{Tr}(\rho_{k,k+1} h_{k,k+1})$ is determined by the same reduced states $\rho_{k,k+1}$. Therefore, lower bounds to the free energy $F = E - TS$ of a thermal state at temperature $T$ can be found with a variational algorithm which only uses the reduced states $\rho_{k,k+1}$ \cite{18}.

Equation (6.4) was obtained by applying Eq. (6.2) $n - 2$ times for $A = \{k + 1\}$, $B = \{k\}$ and $C = \{1, \ldots, k - 1\}$ ($k \in \{2, \ldots, n - 1\}$). These inequalities are equivalent to the following inequalities (because Eq. (6.2) is equivalent to $I(B : C) \leq I(AB : C)$):

$$I(1 \ldots k - 1 : k)_{\rho} \leq I(1 \ldots k - 1 : k, k + 1)_{\rho}, \quad (k \in \{2, \ldots, n - 1\}).$$

(6.5)

If equality holds in Eq. (6.5) or, equivalently, in Eq. (6.4), the global state $\rho$ can be obtained from the reduced states $\rho_{k,k+1}$ ($k \in \{1, \ldots, n - 1\}$) via Petz recovery maps (see supplementary material and main text of \cite{18}). We state this known result in Theorem 6.1 and show that these conditions imply that MPO reconstruction (as stated in Theorem 6.2) is possible (Theorem 6.4).

6.1. Reconstruction of states from local marginal states

The state of the spin chain is $\rho \in \mathcal{D}(\mathcal{H}_{1 \ldots n})$ where $\mathcal{H}_{1 \ldots n} = \mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_n$ is the tensor product of the single-spin Hilbert spaces. For each $k \in \{1, \ldots, n\}$, we partition the spins on the chain into two parts:

$$X_k = \{1 \ldots k\} \quad \text{and} \quad \{k + 1 \ldots n\}.$$  

(6.6)

The marginal states $\rho_{X_k} = \text{Tr}_{k+1 \ldots n}(\rho)$, for $k \in \{2, \ldots, n\}$, can be defined recursively via

$$\rho_{X_k} = (\text{id}_{X_k} \otimes \text{Tr}_{k+1})(\rho_{X_{k+1}}), \quad (k \in \{2, \ldots, n - 1\}), \quad \rho_{X_n} = \rho. \quad (6.7)$$

Each partial trace $\text{Tr}_{k+1}$ is a local CPTP map. If the partial trace $\text{Tr}_{k+1}$ does not decrease the mutual information between $\{1 \ldots k - 1\}$ and $\{k, k + 1\}$ for all $k \in \{2 \ldots n - 1\}$, then the $n$-spin state $\rho$ can be recovered from marginal states of two neighbouring spins \cite{18}:

**Theorem 6.1** Let $\rho \in \mathcal{D}(\mathcal{H}_{1 \ldots n})$ be a quantum state. If the conditions

$$I(X_{k-1} : k)_{\rho} = I(X_{k-1} : k, k + 1)_{\rho}, \quad (k \in \{2, \ldots, n - 1\}),$$

(6.8)

are satisfied, then the marginal states $\rho_{X_k} = \text{Tr}_{k+1 \ldots n}(\rho)$ are given by

$$\rho_{X_{k+1}} = (\text{id}_{X_{k+1}} \otimes R_k)(\rho_{X_k}), \quad \rho = \rho_{X_n} = R_{n-1}R_{n-2}\ldots R_3 R_2(\rho_{12}), \quad (6.9)$$

where $k \in \{2, \ldots, n - 1\}$. The recovery maps $R_k : \mathcal{B}(\mathcal{H}_k) \to \mathcal{B}(\mathcal{H}_{k,k+1})$ are given by Petz recovery maps $R_k = R_k^\mathcal{P}_{\rho_{k,k+1}, \text{Tr}_{k+1}}$ (Theorem 2.1). In the above, $\rho_{ij}$ denotes the reduced state of $\rho$ on sites $i$ and $j$. 

27
Proof For $k \in \{2, \ldots, n - 1\}$, apply Corollary 4.1 with $X = X' = X_{k-1}$, $N_X = \text{id}_X$, $Y = \mathcal{H}_{k,k+1}$, $Y' = \mathcal{H}_k$ and $N_Y(\cdot) = \text{Tr}_{k+1}(\cdot)$.

In a similar way, if the partial traces do not decrease certain operator Schmidt ranks, their actions can be reverted with state reconstruction $[1, 20]$:

**Theorem 6.2** Let $\rho \in \mathcal{B}(\mathcal{H}_{1..n})$ be a linear operator. If the conditions

\[
\text{OSR}(k-1 : k)_{\rho} = \text{OSR}(X_{k-1} : k, k+1)_{\rho} \quad (k \in \{3, \ldots, n-1\})
\]

are satisfied, the marginal states $\rho_{X_k} = \text{Tr}_{k+1..n}(\rho)$ are given by

\[
\rho_{X_{k+1}} = (\text{id} \otimes \mathcal{R}_k)(\rho_{X_k}), \quad \rho = \mathcal{R}_{n-1} \mathcal{R}_{n-2} \ldots \mathcal{R}_4 \mathcal{R}_3(\rho_{123})
\]

where $k \in \{3, \ldots, n-1\}$ and $\mathcal{R}_k \in \mathcal{B}(\mathcal{H}_k; \mathcal{B}(\mathcal{H}_{k+1}))$. The maps are given by $\mathcal{R}_k = \mathcal{R}_M^{\mathcal{L}_k, \text{Tr}_{k+1}}$ (Theorem 3.6), $\mathcal{L}_k = \mathcal{M}_k^T$ with $\sigma_k = \rho_{k-1,k,k+1}$ and $\mathcal{M}_k \in \mathcal{B}(\mathcal{B}(\mathcal{H}_{k+1}); \mathcal{B}(\mathcal{H}_{k-1}))$ (Eq. (2.7)).

**Proof** For $k \in \{3, \ldots, n-1\}$, apply Theorem 3.6 with $X = X_{k-1}$, $X' = \{k\}$, $N_X(\cdot) = \text{Tr}_{X_{k-2}}(\cdot)$, $Y = \mathcal{H}_{k,k+1}$, $Y' = \mathcal{H}_k$, $N_Y(\cdot) = \text{Tr}_{k+1}(\cdot)$. Recall that $\sigma_k = \rho_{k-1,k,k+1} = (\text{Tr}_{X_{k-2}} \otimes \text{id})(\rho_{X_{k+1}})$ implies $\mathcal{M}_{\mathcal{O}_k} = (\text{Tr}_{X_{k-2}}) \mathcal{M}_{\rho_{X_{k+1}}}$ where $\mathcal{M}_{\mathcal{O}_k} \in \mathcal{B}(\mathcal{B}(\mathcal{H}_{k+1}); \mathcal{B}(\mathcal{H}_{k-1}))$ and $\mathcal{M}_{\rho_{X_{k+1}}} \in \mathcal{B}(\mathcal{B}(\mathcal{H}_{k+1}); \mathcal{B}(\mathcal{X}_{k-1}))$ (Lemma 3.8). Therefore, the reconstruction map is given by $\mathcal{R}_k = \mathcal{R}_M^{\mathcal{L}_k, \text{Tr}_{k+1}}$ with $\mathcal{L}_k = \mathcal{M}_{\rho_{X_{k+1}}}^T[\text{Tr}_{X_{k-2}}] = M_{\mathcal{O}_k}^T$.

The result from Theorem 6.2 has been obtained previously in [1] under the name reconstruction of quantum states or MPO reconstruction. For a discussion of further related work [20], see Remark 6.19.

**Remark 6.3 (Efficient recovery/reconstruction)** We call a recovery or reconstruction method to obtain $\rho \in \mathcal{D}(\mathcal{H}_{1..n})$ efficient if it satisfies the following conditions. The method provides an efficient representation of $\rho$ (cf. Section 2.2). This representation of $\rho$ can be constructed from suitable input data in at most $\text{poly}(n)$ computational time. As a consequence, the size of the input data may be at most $\text{poly}(n)$ (i.e. at most $\text{poly}(n)$ complex numbers). The necessary input data may be obtained from at most $\text{poly}(n)$ different tensor product expectation values, i.e. expectation values of the form $\text{Tr}[A_1 \otimes \cdots \otimes A_n \otimes A']N(\rho \otimes \rho')$ where $A_k \in \mathcal{B}(\mathcal{H}_k)$, $A' \in \mathcal{B}(\mathcal{H}_{Y'})$, $\rho' \in \mathcal{D}(\mathcal{H}_{Y'})$ and $Y'$ is an ancilla system of dimension $d_{Y'} = O(\text{poly}(n))$. The quantum operation $N$ is constructed from at most $\text{poly}(n)$ quantum operations whose input and output dimension is at most poly($n$). This severely restricts the available measurements because the number of two-qubit gates required to implement e.g. an arbitrary $n$-qubit unitary is exponential in $n$ [33].

---

$^6$ We introduce the ancilla system to capture the precise definition of the measurements in Section 6.2.
Standard quantum state tomography is not efficient because it fails to satisfy any of these criteria. For example, in quantum state tomography \( \sim \exp(n) \) expectation values are required in order to determine \( \rho \).

Clearly, Theorem 6.1 and Theorem 6.2 satisfy all of these criteria because efficient representations are provided and the necessary input data consists only of two- and three-spin marginals of \( \rho \). Lemma A.4 also provides efficient MPO and PMPS representations for Theorem 6.1 and an efficient MPO representation for Theorem 6.2.

Note that the operator Schmidt rank condition (6.10) is different from the mutual information condition (6.8) in that it contains \( \{k-1\} \) instead of \( X_{k-1} \) on the very left. If the partial trace \( \text{Tr}_X_{k-2} \) which maps \( X_{k-1} \) onto \( \{k-1\} \) was left out, the state \( \rho_{X_{k+1}} \) would be needed to construct \( \mathcal{R}_k \). Construction of \( \mathcal{R}_{n-1} \) would need \( \rho_{X_n} = \rho \) and the reconstruction would be neither efficient nor useful. Despite this difference, we show that the premise of state recovery (Theorem 6.1) implies the premise of state reconstruction (Theorem 6.2):

**Theorem 6.4** Let \( \rho \in \mathcal{D}(\mathcal{H}_{1...n}) \) a quantum state. The conditions

\[
I(X_{k-1} : k)_\rho = I(X_{k-1} : k, k+1)_\rho \quad (k \in \{2, \ldots, n-1\}),
\]

imply

\[
\text{OSR}(k-1 : k)_\rho = \text{OSR}(X_{k-1} : k, k+1)_\rho \quad (k \in \{3, \ldots, n-1\}).
\]

In other words, if the state \( \rho \) can be recovered with Petz recovery from the marginals \( \rho_{k,k+1} \) (\( k \in \{1, \ldots, n-1\} \)), then it can also be reconstructed with state reconstruction from the marginals \( \rho_{k-1,k,k+1} \) (\( k \in \{2, \ldots, n-1\} \)).

**Proof** Equation (6.12) implies \( (k \in \{2, \ldots, n-1\} \), use Eq. (2.22))

\[
0 = I(X_{k-1} : k, k+1) - I(X_{k-1} : k) = I(X_{k-1} : k+1|k)
\]

\[
= I(X_{k-1}, k : k+1) - I(k : k+1).
\]

\[
= I(X_k : k+1) - I(k : k+1).
\]

We shift the index of the last equation by one and obtain, for \( k \in \{3, \ldots, n-1\} \),

\[
I(k-1 : k) = I(X_{k-1} : k) = I(X_{k-1} : k, k+1).
\]

This mutual information equality implies the corresponding operator Schmidt rank equality (Theorem 5.1). ■

**Remark 6.5** For \( n = 4 \), Theorem 6.4 reduces to “(C2) implies Eq. (C1)” from Lemma 5.3 if one takes into account that “\( I(B : C) = I(AB : CD) \)” (C2) is equivalent to “\( I(A : B) = I(A : BC) \) and \( I(AB : C) = I(AB : CD) \)” (see the following Lemma 6.6). ■
Lemma 6.6 $I(B : C) = I(AB : CD)$ holds if and only if $I(A : B) = I(A : BC)$ and $I(AB : C) = I(AB : CD)$.

Proof “$\Rightarrow$”: Let $I(B : C) = I(AB : CD)$ hold. We have $I(B : C) \leq I(AB : C) \leq I(AB : CD) = I(B : C)$, which implies that $I(B : C) = I(AB : C) = I(AB : CD)$. As a consequence,

$$0 = I(AB : C) - I(B : C) = I(A : C|B) = I(A : BC) - I(A : B)$$

(6.16)

also holds (Eq. (2.22)). This already shows the proposed conclusion.

“$\Leftarrow$”: Let $I(A : B) = I(A : BC)$ and $I(AB : C) = I(AB : CD)$ hold. The former equality implies Eq. (6.16) and this shows that $I(B : C) = I(AB : C) = I(AB : CD)$ holds. □

6.2. Long-ranged measurements

In this subsection, we generalize recovery and reconstruction to use certain long-range measurements as input and show that successful recovery implies that successful reconstruction is also possible.

6.2.1. Recovery from long-ranged measurements

Recovery and reconstruction of a spin chain state from few-body marginals required that correlations (as measured by the mutual information or the operator Schmidt rank) do not decrease under the following partial traces (Fig. 2):

$$\rho_{X_k} = (\text{id}_{X_k} \otimes \text{Tr}_{k+1})(\rho_{X_{k+1}}) \quad (k \in \{2, \ldots, n-1\}), \quad \rho_{X_n} = \rho.$$  \hspace{1cm} (6.17)

In order to incorporate long-range measurements, we introduce ancillary systems $Y'_k$ ($k \in \{0, \ldots, n\}$, $d_{Y'_0} = d_{Y'_n} = 1$), quantum operations $\mathcal{T}_k : \mathcal{B}(k, Y'_k) \rightarrow \mathcal{B}(Y'_{k-1})$ and define $\rho_k \in \mathcal{D}(X_k, Y'_k)$ via (Fig. 2)

$$\rho_{k-1} = (\text{id}_{X_{k-1}} \otimes \mathcal{T}_k)(\rho_k) \quad (k \in \{2, \ldots, n\}), \quad \rho_n = \rho.$$  \hspace{1cm} (6.18)

The relation between long-range measurements and the maps $\mathcal{T}_k$ is explained Remark 6.7. If the mutual information $I(X_{k-1} : k, Y'_k)_{\rho_k}$ does not decrease when $\mathcal{T}_k$ is applied, then Theorem 6.9 provides a reconstruction of $\rho$ from the states $\sigma_k = \text{Tr}_{X_{k-1}}(\rho_k) \in \mathcal{D}(k, Y'_k)$ (details are specified in the theorem). Before we state the theorem, we explain that measurements on $\sigma_k$ correspond to recursively defined long-range measurements on $\rho$ and we observe that suitable ancilla systems $Y'_k$ and operations $\mathcal{T}_k$ can be determined for any pure MPS.

Remark 6.7 (Recursively defined long-range measurements) In Theorem 6.9 below, the state $\rho$ is reconstructed from the states $\sigma_k = \text{Tr}_{X_{k-1}}(\rho_k) \in \mathcal{D}(k, Y'_k)$ ($k \in \{1 \ldots n\}$). The states $\sigma_k$ can be reconstructed from the expectation values $\text{Tr}(F_i|\sigma_k)$ of a set observables $\{F_i \in \mathcal{B}(k, Y'_k)\}_i$ which is complete, i.e. spans the full vector space $\mathcal{B}(k, Y'_k)$.

7 The observables may be given e.g. by the elements of a POVM.
Remark 6.8 (Example: Pure matrix product states) Suppose that we fix \( l \geq 1 \) and set \( \mathcal{H}_{Y'_k} = \mathcal{H}_{k+1...k+l} \). In this case, \( \rho_k \in \mathcal{D}(X_k, Y'_k) = \mathcal{D}(X_{k+l}) \) (cf. Eq. (6.18)). We define \( \mathcal{T}_k(\cdot) = \text{Tr}_{k+l}(U_k \cdot U^*_k) \) where \( U_k \in \mathcal{B}(\mathcal{H}_{k...k+l}) \) are unitary operators. Suppose further that the unitaries have the property that they transform \( \rho_k \) into

\[
(1 \otimes U_k)\rho_k (1 \otimes U^*_k) = \rho'_k \otimes |\phi_k\rangle \langle \phi_k| \tag{6.20}
\]

where \( \rho'_k \in \mathcal{D}(\mathcal{H}_{1...k+l}) \) and \( |\phi_k\rangle \in \mathcal{H}_{k+l} \) are states. Then, \( \rho_{k-1} = \mathcal{T}(\rho_k) = \rho'_k \) and the action of \( \mathcal{T}_k \) on \( \rho_k \) can be reversed with \( \mathcal{M}_k(\cdot) = U^*_k(\cdot \otimes |\phi_k\rangle \langle \phi_k|)U_k \), i.e. \( \mathcal{M}_k(\rho_{k-1}) = \rho_k \). As a consequence, the mutual information \( I(X_{k-1} : k, Y'_k)_{\rho_k} \) does not decrease if \( \mathcal{T}_k \) is applied (Corollary 4.1) and we can apply Theorem 6.9 to reconstruct \( \rho \). If \( \rho \) is a pure state which has an MPS representation of bond dimension \( D \), then unitaries which satisfy Eq. (6.20) exist if \( l = \lceil \log_d(D) \rceil \) where \( d = \max_k d_k \) is the maximal dimension of a single spin [15]. In this case, Theorem 6.9 provides an efficient reconstruction if \( D = O(\text{poly}(n)) \) (cf. Remark 6.11). As a consequence, any state which can be reconstructed with the pure-state reconstruction scheme based on unitary operations from Ref. [15] can also be reconstructed with Theorem 6.9 if the same unitaries are used. \( \square \)

**Theorem 6.9** Let \( \rho \in \mathcal{D}(\mathcal{H}_{1...n}) \) a quantum state. Let \( Y'_k \) \((k \in \{1, \ldots, n\})\) be ancilla systems with \( \text{dim}(Y'_n) = 1 \) and choose quantum operations \( \mathcal{T}_k : \mathcal{B}(k, Y'_k) \rightarrow \mathcal{B}(Y'_{k-1}) \) \((k \in \{2, \ldots, n\})\). Define
Remark 6.11

\[ \rho_k \in \mathcal{D}(X_k, Y'_k) \quad (k \in \{1, \ldots, n\}) \text{ recursively via} \]
\[ \rho_{k-1} = (\text{id} \otimes \mathcal{T}_k)(\rho_k) \quad (k \in \{2, \ldots, n\}), \quad \rho_n = \rho. \quad (6.21) \]

If the conditions
\[ I(X_{k-1} : Y'_{k-1})_{\rho_{k-1}} = I(X_{k-1} : k, Y'_k)_{\rho_k}, \quad k \in \{2, \ldots, n\}, \]
are satisfied, then
\[ \rho_k = (\text{id} \otimes \mathcal{R}_k)(\rho_{k-1}) \quad (k \in \{2, \ldots, n\}), \quad \rho = \rho_n = \mathcal{R}_n \mathcal{R}_{n-1} \ldots \mathcal{R}_2(\rho_1) \quad (6.23) \]
where the recovery maps $\mathcal{R}_k : \mathcal{B}(Y_{k-1}') \to \mathcal{B}(k, Y'_k)$ are given by Petz recovery maps $\mathcal{R}_k = \mathcal{R}^{P}_{\sigma_k, \mathcal{T}_k}$ (Theorem 2.1) with $\sigma_k = \text{Tr}_{X_{k-1}}(\rho_k) \in \mathcal{D}(k, Y'_k)$.

**Proof** For $k \in \{2, \ldots, n\}$, apply Corollary 4.1 with $X = X' = X_{k-1}$, $\mathcal{N}_X = \text{id}$, $Y = \mathcal{H}_k, Y'_k$, $Y' = Y'_{k-1}$ and $\mathcal{N}_Y = \mathcal{T}_k$.

**Remark 6.10** Theorem 6.9 provides Theorem 6.1 by restricting to the special case $\mathcal{H}_{Y_k'} = \mathcal{H}_{k+1}$ ($k \in \{1, \ldots, n-1\}$), $\mathcal{T}_k(\cdot) = \text{Tr}_{k+1}(\cdot)$ ($k \in \{2, \ldots, n-1\}$), $\mathcal{T}_n = \text{id}$ and using Eq. (6.23) only for $k \in \{2, \ldots, n-1\}$.

**Remark 6.11** Denote by $d_{Y'} = \max_k \dim(Y'_{k})$ the maximal dimension of any ancillary system. If $d_{Y'} = O(\text{poly}(n))$, the recovery scheme from Theorem 6.9 is efficient (it satisfies all conditions from Remark 6.3). Lemma A.4 provides efficient PMPS and MPO representations of $\rho$.

### 6.2.2 Reconstruction from long-ranged measurements

State reconstruction can be generalized similarly but it requires that additional ancillary systems $X'_k$ and linear maps $\mathcal{U}_k$ are introduced (Fig. 3):

**Theorem 6.12** Let $\rho \in \mathcal{B}(\mathcal{H}_{1 \ldots n})$ be a linear operator. Let $Y'_k \quad (k \in \{0, \ldots, n\})$ and $X'_k \quad (k \in \{0, \ldots, n-1\})$ ancilla systems with $\dim(Y'_0) = \dim(Y'_n) = \dim(X'_0) = 1$. Choose linear maps $\mathcal{T}_k \in \mathcal{B}(\mathcal{B}(k, Y'_1); \mathcal{B}(Y'_{k-1})) \quad (k \in \{2, \ldots, n\})$ and $\mathcal{U}_k \in \mathcal{B}(\mathcal{B}(X_{k-1}); \mathcal{B}(X'_{k-1})) \quad (k \in \{1, \ldots, n\}, \quad \mathcal{U}_1 = 1)$. As before (Eq. (6.21)), we define $\rho_k \in \mathcal{B}(X_k, Y'_k)$ via

\[ \rho_{k-1} = (\text{id} \otimes \mathcal{T}_k)(\rho_k) \quad (k \in \{2, \ldots, n\}), \quad \rho_n = \rho. \quad (6.24) \]

In addition, we define (Fig. 3)

\[ \sigma_k = (\mathcal{U}_k \otimes \text{id})(\rho_k) \quad \in \mathcal{B}(X'_{k-1}, k, Y'_k) \quad (k \in \{1, \ldots, n\}), \quad (6.25) \]
\[ \tau_k = (\text{id} \otimes \mathcal{T}_k)(\sigma_k) \quad \in \mathcal{B}(X'_{k-1}, Y'_{k-1}) \quad (k \in \{2, \ldots, n\}). \quad (6.26) \]
Remark 6.13 Expectation values which grows polynomially instead of exponentially with \( n \).

Efficient reconstruction implies that a given state can be reconstructed from a number of representation of state of dimension can only be achieved if the to-be-reconstructed state is not a completely general quantum state of \( n \) systems. In the following, we show that the condition for efficient reconstruction
in particular implies that the operator Schmidt ranks of the state are restricted to growing polynomially (instead of exponentially) with \( n \).

For \( k = \lceil \frac{n}{2} \rceil \), the maximal value of the operator Schmidt rank \( \text{OSR}(X_k : k + 1 \ldots n)_\rho \) is
\[
(\min\{d_1 \ldots d_k, d_{k+1} \ldots d_n\})^2 = O(\exp(n)) \quad \text{and it is assumed e.g. for maximally entangled pure states.}
\]
Suppose that \( \rho \) can be reconstructed efficiently. The equality \( \rho = (\text{id}_{X_k} \otimes R_n \ldots R_{k+1})(\rho_k) \) (Eq. (6.28)) implies \( \text{OSR}(X_k : k + 1 \ldots n)_\rho \leq \text{rk}(R_{k+1}) \) (Corollary 3.9). The rank of \( R_{k+1} \) is, in turn, upper bounded by \( \text{rk}(R_{k+1}) \leq d^2 \), \( O(\text{poly}(n)) \). I.e. the operator Schmidt rank of \( \rho \) is at most \( \text{OSR}(X_k : k + 1 \ldots n)_\rho = O(\text{poly}(n)) \). In conclusion, any state which admits an efficient reconstruction with Theorem 6.12 has a small operator Schmidt rank in the sense that it does not grow exponentially but only polynomially with the number of spins \( n \).

**Remark 6.15 (Recursively defined long-range measurements)** In Theorem 6.12, \( \rho \) is reconstructed from \( \sigma_k = \mathcal{B}(X_{k-1}', k, Y_k') \) \((k \in \{1 \ldots n\})\), noting that \( \sigma_1 = \rho_1 \). As above (Remark 6.7), measurements on \( \sigma_k \) correspond to recursively defined long-range measurements on \( \rho \) (Fig. 3):
\[
\text{Tr}(F \sigma_k) = \text{Tr}(F(U_k \otimes (T_n \ldots T_{k+1}))(\rho)) = \text{Tr}(G \rho), \quad F \in \mathcal{B}(X_{k-1}', k, Y_k'), \quad G = [(U_k' \otimes (T_{k+1}^* \ldots T_n^*))](F^*)^T \in \mathcal{B}({\mathcal{H}}_{1 \ldots n}).
\]

If the superoperators involved are quantum operations and the operator \( F \) is Hermitian, \( G \) is Hermitian as well and there is a correspondence between observables on \( \sigma_k \) and recursively defined long-ranged observables on \( \rho \). Otherwise, the correspondence holds only in an abstract sense between operators \( F \in \mathcal{B}(X_{k-1}', k, Y_k') \) and \( G \in \mathcal{B}({\mathcal{H}}_{1 \ldots n}). \)

**Remark 6.16 (Mixed state which requires long-range measurements)** Remark 6.8 showed that any pure MPS can be recovered with Theorem 6.9 if the unitary operations from [15] are used. Below, we show that recovery with Theorem 6.9 implies that reconstruction with Theorem 6.12 is also possible (see Theorem 6.17). The following simple mixed state shows that Theorems 6.9 and 6.12 can recover more than pure matrix product states and more than recovery or reconstruction from local reduced states (Theorems 6.1 and 6.2): The state
\[
\rho = \frac{1}{2} \left[ |0 \ldots 0\rangle \langle 0 \ldots 0| + |100 \ldots 01\rangle \langle 100 \ldots 01| \right] \in \mathcal{D}({\mathcal{H}}_{1 \ldots n})
\]

does not admit recovery or reconstruction from local reduced states because it turns into a product state if the first or last site is traced out; this unavoidably reduces the mutual information from non-zero to zero and the operator Schmidt rank from larger than one to one. The state admits recovery or reconstruction via Theorem 6.9 and Theorem 6.12 if the following definitions are used: Assuming uniform local dimensions \( d = d_k \ (k \in \{1, \ldots, n\}) \), set \( \mathcal{H}_{k'} = \mathcal{H}_{k+1}, \ T_k = \text{Tr}_{k+1} \text{SWAP}_{k,k+1} \ (k \in \{2, \ldots, n-1\}) \), \( T_n = \mathds{1}_n \), \( X_{k-1}' = \mathcal{H}_{k-1}, \ U_k = \text{Tr}_{1 \ldots k-2} \text{SWAP}_{1,k-1} \ (k \in \{2, \ldots, n\}) \) and \( \mathcal{U}_1 = 1 \).


\[
\text{SWAP}_{kl}(\cdot) = S_{kl} \cdot S_{kl}, \quad S_{kl} = \sum_{i=1}^{d_k} \sum_{j=1}^{d_l} |ij\rangle \langle ji| \quad \text{where} \ |i\rangle \text{ and} \ |j\rangle \text{ are orthonormal bases of} \ \mathcal{H}_k \text{ and} \ \mathcal{H}_l.
\]

\[8\] The swap gate is given by \( \text{SWAP}_{kl}(\cdot) = S_{kl} \cdot S_{kl}, \quad S_{kl} = \sum_{i=1}^{d_k} \sum_{j=1}^{d_l} |ij\rangle \langle ji| \quad \text{where} \ |i\rangle \text{ and} \ |j\rangle \text{ are orthonormal bases of} \ \mathcal{H}_k \text{ and} \ \mathcal{H}_l.\]
Theorem 6.9 and Theorem 6.12 are given by \( \rho_n = \rho \),
\[
\rho_k = \frac{1}{2} \left[ |0\ldots0踪0\ldots0踪| + |100\ldots0踪100\ldots0踪| \right] \in \mathcal{D}(\mathcal{H}_{1..k+1})
\] (6.31a)

where \( k \in \{1..n-1\} \). The states \( \sigma_k \) and \( \tau_k \) used in Theorem 6.12 are given by \( \sigma_1 = \rho_1 \), \( \sigma_n = \rho_1 \in \mathcal{D}(\mathcal{H}_{n-1,n}) \),
\[
\sigma_k = \frac{1}{2} \left[ |000踪000踪| + |101踪101踪| \right] \in \mathcal{D}(\mathcal{H}_{k-1,k,k+1}) (k \in \{2..n-1\}) \),
\] (6.31b)
\[
\tau_k = \frac{1}{2} \left[ |00踪00踪| + |11踪11踪| \right] \in \mathcal{D}(\mathcal{H}_{k-1,k}) (k \in \{2..n\}) \). \] (6.31c)

These states show that the conditions from Theorem 6.9 and Theorem 6.12 are satisfied. \( \square \)

6.2.3. Recovery vs. reconstruction for long-ranged measurements

In this section, we show that the conditions for state recovery (Theorem 6.9) imply that state reconstruction (Theorem 6.12) is also possible. The premise of Theorem 6.9 implies the premise of Theorem 6.12 for \( U_k = \text{id} \). However, Theorem 6.12 does not provide a useful reconstruction with \( U_k = \text{id} \) because the necessary input \( \sigma_n \) for the construction of \( R_n \) would be \( \sigma_n = \rho \). In Theorem 6.4 we used the symmetry of the conditional mutual information to work around this but this is no longer possible because \( T_k \) was introduced. Note that Eq. (6.22) implies the same equality for operator Schmidt ranks and that Eq. (6.23) provides MPO representations of the \( \rho_k \) (Lemma A.4). It is well-known that maps \( U_k \) suitable for Theorem 6.12 can be obtained directly from the matrices of the MPO representation after the matrices have been transformed into a suitable orthogonal (mixed-canonical) form ([14, 25]: see also Remark A.1 in the appendix). The maps \( U_k \) obtained in this way are given by partial isometries on the vector space of linear operators. Such a map is not guaranteed to be completely positive or trace preserving, i.e. it does not represent a quantum operation and it may not allow for an efficient implementation in a given quantum experiment. An alternative construction has been put forward in [20]:

9 Here, maps \( U_k \) are provided whose matrix representation is given by a submatrix of a permutation matrix in a product basis of \( \mathcal{B}(X_k) = \mathcal{B}(\mathcal{H}_1) \otimes \cdots \otimes \mathcal{B}(\mathcal{H}_k) \). We use this result to prove that efficient recovery implies efficient reconstruction in Theorem 6.17. Remark 6.18 discusses advantages and disadvantages of the two different choices for \( U_k \) mentioned in this paragraph.

**Theorem 6.17** Let the premise of Theorem 6.9 hold. Set \( \dim(Y'_0) = 1 \) and \( X'_k = Y'_k \). There are linear maps \( U_k \in \mathcal{B}(\mathcal{B}(X_{k-1})); \mathcal{B}(X'_{k-1})) (k \in \{2,\ldots,n\}) \) such that
\[
\text{OSR}(X'_{k-1}:Y'_{k-1})_{U_k} = \text{OSR}(X_{k-1}:Y'_k)_{\rho_k} (k \in \{2,\ldots,n\})
\] (6.32)

9 We provide a formal description of the corresponding part of their work in Lemmata A.2 and A.3.
holds where \( \tau_k = (\mathcal{U}_k \otimes \text{id}_{i_{k-1}})(\rho_{k-1}) \) is the same operator as in (6.26). Choose operator bases \( F_{i_k}^{(k)}, F_{j_k}^{(X_j')}, \) and \( F_{i_k}^{(Y_i')} \). There is an efficient algorithm to choose suitable maps \( \mathcal{U}_k \) and they can be chosen such that their matrix representation (Eq. (2.6)) is a submatrix of a permutation matrix. In this case, the resulting reconstruction is efficient (in the sense of Remark 6.3) if recovery is efficient and if the operator bases are chosen such that they contain only Hermitian operators.

**Proof** Lemma A.4 provides an MPO representation of the states \( \rho_k \) from Eq. (6.23). It is well-known that maps \( \mathcal{U}_k \) can be chosen recursively such that \( \text{OSR}(X_{k-1}': Y_{k-1}')_{t_k} = \text{OSR}(X_{k-1} : Y_{k-1}')_{t_k} \) holds if an MPO representation of \( \rho_{k-1} \) is given [14, 25]. As \( \text{OSR}(X_{k-1} : Y_{k-1}')_{t_k} = \text{OSR}(X_{k-1} : k, Y_{k-1}')_{t_k} \) is implied by (6.22) (Theorem 5.1), it is clear that Eq. (6.32) holds as well.

It was also recognized that the maps \( \mathcal{U}_k \) can be chosen such that their matrix representation is a submatrix of a permutation matrix [20]. We provide a self-contained description of the corresponding procedure in Lemma A.2.

Let \( \mathcal{U}_k \) the matrix representation of \( \mathcal{U}_k \) and suppose that \( \mathcal{U}_k \) is a submatrix of a permutation matrix. Denote by \( f_k(j) = \{(i_1, \ldots, i_{k-1}) : (i_k, j) \} \) the set of columns with a non-zero entry in a given row of \( \mathcal{U}_k \) (where \( i_j \in \{1 \ldots d_j^2 \} \) and \( j \in \{1 \ldots d_{X_{k-1}'} \} \)). The matrix elements of \( \sigma_k = (\mathcal{U}_k \otimes \text{id})(\rho_k) \) are given by (insert an identity map (2.5) into Eq. (6.25))

\[
\begin{align*}
\langle F_{j_k}^{(X_{k-1}')} & \otimes F_{i_k}^{(k)} \otimes F_{i_1}^{(Y_1')}, \sigma_k \rangle \\
= & \sum_{i_1 \ldots i_{k-1}} [\mathcal{U}_k]_{j_1 \ldots i_{k-1}} \langle F_{i_1}^{(1)} \otimes \cdots \otimes F_{i_{k-1}}^{(k-1)} \otimes F_{i_k}^{(k)} \otimes F_{i_1}^{(Y_1')}, \rho_k \rangle \\
= & \begin{cases} 
0, & \text{if } f_k(j) = \emptyset, \\
\langle F_{i_1}^{(1)} \otimes \cdots \otimes F_{i_k}^{(k)} \otimes F_{i_1}^{(Y_1')}, \rho_k \rangle, & \text{with } \{(i_1, \ldots, i_{k-1})\} = f_k(j).
\end{cases}
\end{align*}
\]

(6.33)

(6.34)

Here, we used that \( |f_k(j)| \leq 1 \) because \( \mathcal{U}_k \) is a submatrix of a permutation matrix. The last equation shows that \( \sigma_k \), which needs to be known for reconstruction of \( \rho_k \), can be determined from at most \( (d_{X_{k-1}'}d_{Y_{k-1}'})^2 \) tensor product expectation values in \( \rho_k \). The structure of the given expectation values is permitted for efficient reconstruction (Remark 6.3) and the number of expectation values is at most \( \text{poly}(n) \) if recovery is efficient. Furthermore, efficient recovery implies that the MPO representation of the \( \rho_k \) as well as the procedures to determine \( \mathcal{U}_k \) and \( f_k(j) \) are efficient as well ([20]; for details see Lemmata A.2 and A.3). This finishes the proof of the theorem.

**Remark 6.18** The singular values of \( M_{\sigma_k} \in \mathcal{B}(\mathcal{B}(k, Y_{k}'); \mathcal{B}(X_{k-1}')) \) equal those of \( M_{\rho_k} \in \mathcal{B}(\mathcal{B}(k, Y_{k}'); \mathcal{B}(X_{k-1})) \) if the maps \( \mathcal{U}_k \) are suitable partial isometries on the vector space of linear operators (cf. Remark A.1). For reconstruction stability (Theorem 3.4), this is the optimal case (if the maps \( \mathcal{T}_k \) are predefined). If the maps \( \mathcal{U}_k \) are restricted to submatrices of permutation matrices, the singular values of \( M_{\sigma_k} \) are smaller than or equal to those of \( M_{\rho_k} \) (because \( \mathcal{U}_k \) has unit operator norm). If the smallest non-zero singular value decreases, then
stability of the reconstruction is reduced (Theorem 3.4; cf. [20, 34]). In the worst case, the smallest non-zero singular value decreases by a factor exponential in $n$ because of the recursive construction of the $U_k$ [34]. However, empirical results show that this worst-case behaviour is usually not observed in practice [1, 20, 34, 35].

If the maps $T_k$ are not predefined, the singular values of $M_{\rho_k}$ equal those of $M_{\rho} \in \mathcal{B}(\mathcal{H}_k; \mathcal{B}(X_{k-1}))$ if the maps $U_k$ and $T_k$ are suitable partial isometries on the vector space of linear operators (Remark A.1). In this case, Theorem 6.17 allows reconstruction of an arbitrary MPO (or matrix product state/tensor train) with optimal reconstruction stability. However, it remains an open question whether this can be fully exploited e.g. in the reconstruction of quantum states as the necessary measurements may not allow for an efficient implementation if the maps $U_k$ and $T_k$ are general partial isometries on the vector space of linear operators.

The situation is different if the state $\rho$ is a pure matrix product state. Here, partial isometries which act on the Hilbert spaces themselves can be obtained ([15], cf. Remarks 6.8 and A.1). These partial isometries can be implemented via unitary control of the quantum system and they have the property that they preserve the singular values of $M_{\rho}$. This also shows that the tomography scheme for pure matrix product states based on local unitary operations and proposed in [15] provides maps $T_k$ and $U_k$ for state recovery and reconstruction with optimal stability.

Remark 6.19 (Related work) Note that nowhere in the proof of Theorems 3.6 and 6.12 did we use the fact that $\rho$ is a linear operator on $\mathcal{H}_{1\ldots n}$. The theorems apply equally well to arbitrary vectors $|\psi\rangle \in \mathcal{H}_{1\ldots n}$ on tensor product vector spaces $\mathcal{H}_{1\ldots n} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$. The components of $\langle i_1 \ldots i_n |\psi\rangle$ of $|\psi\rangle$ in a product basis define a tensor $t \in \mathbb{C}^{d_1 \times \cdots \times d_n}$ (i.e. an array with $n$ indices).

A result similar to Theorem 6.12 has been obtained before in the context of tensor train representations [20, 35, 34]. Their result is formulated for a tensor with $n$ indices, i.e. replace $\mathcal{B}(\mathcal{H}_k)$ by $\mathcal{H}_k$, $\mathcal{B}(Y_k')$ by $Y_k'$, $\text{OSR}(X_k' - 1 : Y_k' - 1)$ by $\text{rk}(M_{\tau_k})$, etc. They restrict to $\dim(X_k') = \dim(Y_k') = \text{rk}(M_{\tau_k})$. In this case, the pseudoinverse in the reconstruction maps $R_k$ (defined in Theorem 6.12) is just the regular inverse (cf. Remark 3.3). They also restrict $U_k$ and $T_k$ to submatrices of permutation matrices in a fixed product basis. In addition, they provide an algorithm which attempts to determine suitable maps $U_k$ and $T_k$ incrementally and efficiently. Similar work has been carried out for the Tucker and hierarchical Tucker tensor representations [5, 22, 21] and the relation between this work and the matrix reconstruction from Section 3.1 will be explored elsewhere [36].

Acknowledgements

We acknowledge discussions with Oliver Marty. Work in Ulm was supported by an Alexander von Humboldt Professorship, the ERC Synergy grant BioQ, the EU projects QUCHIP and
A. Appendix

A.1. Optimality of the stability bound

The following examples show that the bound from Theorem 3.4 is optimal up to constants and that the reconstruction error can diverge as \( \epsilon \) approaches zero if small singular values in \( \text{LMR} \) are not truncated.

Theorem 3.4 provides an upper bound on the reconstruction error of a reconstructible matrix \( S \) (the signal) which is perturbed by some error matrix \( E \). The following example shows that the upper bound from the theorem is optimal up to constants:

\[
L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad S = \begin{pmatrix} \Delta & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad E = \eta \epsilon \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\]

The eigenvalues of \( S \) are \( \pm \sqrt{1 + \Delta^2} \) so \( \eta = \|S\| = \sqrt{1 + \Delta^2} \) approaches unity as \( \Delta \to 0 \). For simplicity, we might assume that using \( \eta \approx 1 \) is sufficient for the discussion of this example but we keep \( \eta > 1 \). We have \( \|E\|/\eta = \epsilon \). Suppose that we choose \( \epsilon \) and \( \tau \) such that \( 0 < \epsilon < \tau < \frac{1}{3\sqrt{2}} \). Set \( c = \tau + 2\epsilon \), i.e. \( 0 < c < \frac{1}{\sqrt{2}} \), and set \( \Delta = c/\sqrt{1 - c^2} \). Then \( \Delta < 1 \) and \( \eta < \sqrt{2} \). In addition, we obtain

\[
(\tau + 2\epsilon)\eta = c\sqrt{1 + \Delta^2} = c\sqrt{1 - \frac{c^2}{1 - c^2}} = \Delta.
\]

The eigenvalues of \( \text{LSR} \) are \( \pm \Delta \). This provides us \( \gamma = \Delta/\eta = \tau + 2\epsilon \). Therefore, the condition \( \epsilon \leq \tau < \gamma - \epsilon \) is automatically satisfied and, as a consequence, \( 2\epsilon < \gamma \) holds as well. LER can change the eigenvalues of \( \text{LSR} \) at most by \( \epsilon \) (cf. proof of Lemma 3.5), so no truncation occurs.

In this case, the reconstruction error has exactly the scaling from the theorem:

\[
\|\text{MR}(\text{LMR})^\dagger\text{LM} - S\| = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \eta \epsilon & 0 \\ 0 & 0 & -\eta \epsilon \end{pmatrix} = \max \left\{ \eta \epsilon, \frac{\eta \epsilon}{\Delta^2} \right\} \geq \frac{\eta \epsilon}{\Delta^2} \geq \frac{\epsilon}{9\sqrt{2}(\gamma - \epsilon)^2}.
\]

Note that the conditions from above imply \( \Delta^2/\eta = \eta \gamma^2 < \sqrt{2}\gamma^2 \) and that \( \gamma = \tau + 2\epsilon \geq 3\epsilon \). The latter implies \( -\epsilon \geq -\gamma/3 \) and \( 3(\gamma - \epsilon) \geq 2\gamma \geq \gamma \). Combining the relations provides the bound \( \Delta^2/\eta < \sqrt{2}\gamma^2 \leq 9\sqrt{2}(\gamma - \epsilon)^2 \) used above.
One may ask whether thresholds $\tau$ outside the interval permitted by the theorem reconstruct $M$ successfully. In this example, a threshold which is large enough to produce a different reconstruction will replace at least one of the two singular values of the reconstruction by zero. As the two singular values of $S$ are equal, the reconstruction error will be at least $\|S\|$ in this case. i.e. larger thresholds do not provide a successful reconstruction in the sense that the error in operator norm is significantly smaller than $\|S\|$. In this example, neither smaller nor larger thresholds (than the ones permitted by Theorem 3.4) provide an improved reconstruction: Smaller thresholds do not change the reconstruction at all because thresholding does not reduce the rank of $LMR$ in this example. However, the following example shows that thresholding is in general necessary to obtain an error which satisfies the bound from Theorem 3.4. We keep $L$ and $R$ from above and choose

$$S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad E = \epsilon \begin{pmatrix} 0 & 0 & 0 \\ 0 & e^2 & 1-e^2 \\ 0 & 1-e^2 & e^2 \end{pmatrix}. $$

We have $\eta = \|S\| = 1$, $\gamma = 1$ and the eigenvalues of $E/\epsilon$ are 1 and $-1 + 2e^2$ such that $\|E\| = \epsilon$; we choose $\epsilon < \frac{1}{2}$ such that choosing a $\tau$ from $\epsilon \leq \tau < 1 - \epsilon$ is permitted. The eigenvalues of $LMR$ are 1 and $e^3$. We obtain (using $\epsilon \leq \frac{1}{2}$)

$$\|MR(LMR)^+LM - S\| = \left\| \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1-e^2}{\epsilon} & e(1-e^2) \\ 0 & e\epsilon(1-e^2) & \epsilon \end{pmatrix} \right\| \geq \frac{(1-\epsilon)^2}{\epsilon} \geq \frac{1-2e^2}{\epsilon} \geq \frac{1}{2\epsilon}. $$

Without truncating small singular values, the error diverges as $\epsilon \to 0$, i.e. it does not satisfy the bound from Theorem 3.4. Here, the effect of $E$ is completely erased by truncation:

$$(LMR)_{\tau}^2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \|MR(LMR)^+LM - S\| = 0. $$

A.2. The stability bound for matrices with non-unit operator norm

In this section, we provide an argument which extends the proof of Theorem 3.4 from matrices $S$ with unit operator norm to matrices $S$ with arbitrary operator norm. Suppose that the matrix $M$ is the sum of a signal $S$ and an noise contribution $E$, $M = S + E$. The signal satisfies $\text{rk}(S) = \text{rk}(LSR)$, but we only know the strength $\|E\|$ of the noise. Suppose that for $\|S\| = \|L\| = \|R\| = 1$, we obtain some error bound of the form

$$\|MR(LMR)^+LM - S\| \leq f(\epsilon, \gamma, \tau), \quad \epsilon = \|E\|, \quad \gamma = \sigma_{\min}(LMR). \quad (A.1)$$

We can obtain an error bound for $M' = S' + E'$ where $S'$, $L'$, and $R'$ have arbitrary norms as follows: Set $M = M'/\|S'\|$, $S = S'/\|S'\|$, $E = E'/\|E'\|$, $L = L'/\|L'\|$, $R = R'/\|R'\|$. With these
definitions, we have
\[
\|MR(LMR)_t^+LM - S\| = \frac{\|M'R'(L'M'R')_t^+L'M' - S'\|}{\|S'\|},
\] (A.2)
where \(\tau' = \|L\|\|R\|\|S\|\tau\). Therefore, the bound from the last but one equation implies
\[
\|M'R'(L'M'R')_t^+L'M' - S'\| \leq \|S'\|f(\epsilon, \gamma, \tau),
\] (A.3)
\[
\epsilon = \frac{\|E'\|}{\|S'\|}, \quad \gamma = \frac{\sigma_{\min}(L'M'R')}{\|L'\|\|R'\|\|S'\|}, \quad \tau = \frac{\tau'}{\|L'\|\|R'\|\|S'\|}.
\] (A.4)
In proofs, we assume \(\|S\| = \|L\| = \|R\| = 1\) and we use \(\epsilon = \|E\|\).

A.3. Alternative proof of the stability bound

In this section we obtain a bound similar to the one from Theorem 3.4 using the ansatz by Caiafa and Cichocki [22].

As above, we use \(M = S + E, \|S\| = \|L\| = \|R\| = 1\) and \(\text{rk}(S) = \text{rk}(LSR)\).

Note that \(\text{rk}(S) = \text{rk}(LSR)\) implies \(\text{rk}(S) = \text{rk}(LS) = \text{rk}(SR) = \text{rk}(LSR)\). We use the matrix reconstruction Proposition 3.1 several times, sometimes with \(L\) or \(R\) replaced by the identity matrix. The proposition e.g. provides \(S = S(LS)^+LS\). Using that identity, we obtain the following two equalities:
\[
S = S(LS)^+L(M - E)
\] (A.5)
\[
M = S + E = S(LS)^+LM + (1 - S(LS)^+L)E
\] (A.6)

In the same way, we obtain:
\[
S = (M - E)R(SR)^+S
\] (A.7)
\[
M = S + E = MR(SR)^+S + E(1 - R(SR)^+S)
\] (A.8)

We will also use
\[
LMR(LMR)_t^+LMR = (LMR)_t = [(LMR)_t - LMR] + LSR + LER.
\] (A.9)

We decompose into three parts:
\[
MR(LMR)_t^+LM = (A) + (B) + (C)
\] (A.10)
We insert Eq. (A.6) for $M$ at the beginning of the expression and Eq. (A.8) for $M$ on the end of the expression. In the following equations, spaces separate factors which come from different equations. In part (A) below, we insert Eq. (A.9).

\[(A) = S(LS)^+ LM R(LMR)_{\tau}^+ L MR(SR)^+ S \]
\[= S(LS)^+ [(LRM)_{\tau} - LMR] (SR)^+ S \]
\[+ S(LS)^+ LSR (SR)^+ S \]
\[+ S(LS)^+ LER (SR)^+ S \quad (A.11a)\]

\[(B) = S(LS)^+ LM R(LMR)_{\tau}^+ L E(1 - R(SR)^+ S) \]
\[+ (1 - S(LS)^+ L)E R(LMR)_{\tau}^+ L MR(SR)^+ S \quad (A.11b)\]

\[(C) = (1 - S(LS)^+ L)E R(LMR)_{\tau}^+ L E(1 - R(SR)^+ S) \quad (A.11c)\]

The expression in Eq. (A.11a) is equal to $S$. We use the relation $\|AA_{\tau}\| = \|A_{\tau}A_{\tau}\| \leq 1$ and obtain the following bound:

\[
\|MR(LMR)_{\tau}^+ LM - S\| \\
\leq \|LS\|^+ \|SR\|^+ \left(\|LMR\|_{\tau} - LMR\| + \epsilon\right) \\
+ \|LS\|^+ \|1 - R(SR)^+ S\| \epsilon \\
+ \|SR\|^+ \|1 - S(LS)^+ L\| \epsilon \\
+ \|1 - S(LS)^+ L\| \epsilon^2 \|1 - R(SR)^+ S\| \|LMR\|_{\tau}^+ \epsilon^2 \\
\quad (A.12)
\]

This bound has been given by Caiafa and Cichocki [22] for the case that $L$ and $R$ have exactly $r = \text{rk}(S)$ rows and columns (such that the matrix $LSR$ is invertible). They proceed by defining constants $a$, $b$ and $c$ which are independent of the threshold $\tau$ and of noise strength $\epsilon = \|E\|$ and obtain a bound of the form $a\tau + b\epsilon + c\epsilon^2/\tau$.

We continue by analyzing how all terms in the last equation depend on $L$, $R$ and $S$. This will provide a bound similar to that of Theorem 3.4.

Because $LSR$, $LS$ and $SR$ have all rank $r = \text{rk}(S)$, the relation $\sigma_{\text{min}}(LSR) = \sigma_r(LSR)$ holds for these three matrices. We obtain

\[
\gamma = \sigma_{\text{min}}(LSR) = \sigma_r(LSR) \leq \sigma_r(LS)\sigma_1(R) \leq \sigma_r(LS) = \sigma_{\text{min}}(LS), \quad (A.13)
\]

where the first inequality is provided by Ref. [37] (Theorem 3.3.16, page 178). This provides

\[
\|S(LS)^+ L\| \leq \|LS\|^+ = \frac{1}{\sigma_{\text{min}}(LS)} \leq \frac{1}{\gamma} \quad (A.14)
\]

and the same bound applies to $\|R(SR)^+ S\|$. Note that $\gamma \leq \|S\| = 1$. Further, for any square matrix $A$, we have

\[
\|I - A\| \leq \max\{\sigma_1(1), \sigma_1(A) - 1\} = \max\{1, \|A\| - 1\}. \quad (A.15)
\]
Using $1 \leq 1/\gamma$, we obtain

$$\|I - S(LS)^+L \| \leq \frac{1}{\gamma}, \quad \|I - R(SR)^+S\| \leq \frac{1}{\gamma}$$  \hfill (A.16)

and

$$\|MR(LMR)^+LM - S\| \leq \frac{\|LMR\| - LMR\| + 3\epsilon + \|(LMR)^+\|\epsilon^2}{\gamma^2}.$$ \hfill (A.17)

The inequality holds for arbitrary values of $\gamma$, $\tau$ and $\epsilon$.

Now, we assume $\epsilon \leq \tau < \gamma - \epsilon$ and use bounds from Lemma 3.5. This provides

$$\|MR(LMR)^+LM - S\| \leq \frac{4\epsilon}{\gamma^2} + \frac{\epsilon^2}{\gamma^2(\gamma - \epsilon)} \leq \frac{5\epsilon}{\gamma^2}.$$ \hfill (A.18)

Without the assumption $\epsilon \leq \tau < \gamma - \epsilon$, we obtain

$$\|MR(LMR)^+LM - S\| \leq \frac{\tau + 3\epsilon}{\gamma^2} + \frac{\epsilon^2}{\gamma^2\tau}.$$ \hfill (A.19)

This is again the bound $a\tau + b\epsilon + c\epsilon^2/\tau$, but with $a = 1/\gamma^2$, $b = 3/\gamma^2$ and $c = 1/\gamma^2$. If we select $\epsilon = \tau$ in this bound, we obtain exactly Eq. (A.18). As $(LMR)^+$ is the same operator for all $\tau \in [\epsilon, \gamma - \epsilon)$, the bound holds not only for $\epsilon = \tau$ but for all $\tau \in [\epsilon, \gamma - \epsilon)$: We obtain Eq. (A.18) again.

### A.4. Known results on matrix product representations

This section reviews known results on matrix product state/tensor train representations used in Section 6.2.3. It also provides full formal details for the results which were used.

Given a tensor $t \in \mathbb{C}^{d_1 \times \ldots \times d_n}$, a matrix product representation of the tensor is given by

$$[t]_{i_1, \ldots, i_n} = G_1(i_1) \ldots G_m(i_m)H_{m+1}(i_{m+1})G_{n+1}(i_n) \quad (i_k \in \{1, \ldots, d_k\})$$ \hfill (A.20)

where $D_0 = D_n = 1$, $H_m \in \mathbb{C}^{D_m \times D_n}$, $G_k(i_k) \in \mathbb{C}^{D_{k-1} \times D_k}$, $i_k \in \{1, \ldots, d_k\}$ and $m \in \{1, \ldots, n-1\}$. For simplicity, $H_m = 1_{D_n}$ may be used. The $G_k$ are called the cores of the representation while the matrices $G_k(i_k)$ give the representation its name. The left and right unfoldings of the cores are given by

$$G_k^L \in \mathbb{C}^{D_{k-1} \times d_k \times D_k} \quad \quad \quad G_k^R \in \mathbb{C}^{D_{k-1} \times d_k \times D_k}$$ \hfill (A.21)

---

10 This notation is partially inspired by [38] but notation in the tensor train literature does not seem to be uniform.
and they have the same entries as $G_k$, e.g. $[G_k^L]_{b_{k-1}i_kb_k} = [G_k(i_k)]_{b_{k-1}b_k}$. The left and right interface matrices are given by

$G_{\leq k} = (G_{\leq k-1} \otimes 1_{d_k}) G_k^L \in \mathbb{C}^{d_1 \ldots d_k \times D_k}$ \hspace{1cm} (k \in \{1 \ldots m\}), \hspace{1cm} (A.22)

$G_{> k} = (1_{d_k+1} \otimes G_{> k+1})(G_k^R)^\top \in \mathbb{C}^{d_k+1 \ldots d_n \times D_k}$ \hspace{1cm} (k \in \{m \ldots n-1\}) \hspace{1cm} (A.23)

where $G_{\leq 0} = 1$ and $G_{> n} = 1$. The unfolding $t_k$ is the $d_1 \ldots d_k \times d_{k+1} \ldots d_n$ matrix with the same entries as $t$ and it can be written as

$t_m = G_{\leq m} H_m (G_{> m})^\top$ \hspace{1cm} (A.24)

It is well-known that a singular value decomposition of the unfolding $t_k$ can be obtained efficiently [14, 25]:

**Remark A.1** Fix $m \in \{1, \ldots, n-1\}$, let $t$ have a matrix product representation as in Eq. (A.20) with positive-semidefinite diagonal $H_m$ and assume orthogonal cores, i.e.

$$(G_k^L)^\top G_k^L = 1_{D_k} \quad (k \leq m)$$

$$(G_{\leq k})^\top G_{\leq k} = 1_{D_k} \quad (k \leq m)$$

$$(G_k^R)^\top G_k^R = 1_{D_k} \quad (k > m)$$

$$(G_{> k})^\top G_{> k} = 1_{D_k} \quad (k \geq m)$$ \hspace{1cm} (A.25a)

$$(G_{\leq k})^\top G_{\leq k} = 1_{D_k} \quad (k \leq m)$$

$$(G_{> k})^\top G_{> k} = 1_{D_k} \quad (k \geq m)$$ \hspace{1cm} (A.25b)

An arbitrary matrix product representation can be efficiently converted into such an orthogonal representation [25, 14]. Then

$t_m = G_{\leq m} H_m (G_{> m})^\top$ \hspace{1cm} (A.26)

is a singular value decomposition of the unfolding matrix $t_m$. Let $U_{\leq m} = (G_{\leq m})^\top$ and $V_{> m} = (G_{> m})^\top$. Then

$U_{\leq m} t_m (V_{> m})^\top = H_m$ \hspace{1cm} (A.27)

has the same singular values as $t_m$. \hfill \Box

The following Lemma provides an efficient, incremental construction of matrices $U_{\leq k}$ and $V_{> k}$ such that the matrix $U_{\leq m} t_m (V_{> m})^\top$ has the same rank as $t_m$. More general matrices are permitted than in Eq. (A.27) and the rank is preserved (Eq. (A.31c)) but the singular values of $U_{\leq m} t_m (V_{> m})^\top$ can differ from those of $t_m$. The proof of Lemmata A.2 and A.3 has been sketched in [20]. In the premise of the following Lemma, it is possible to choose $U_k$ and $V_k$ as submatrices of permutation matrices (the case considered in [20]), but the actual proof is independent of this choice.

---

11 See Ref. [25]; this is called a mixed-canonical representation in Ref. [14].
**Lemma A.2** Assume a matrix product representation of \( t \) as in Eq. (A.20) with \( H_m = 1 \). In the following, \( m \in \{1 \ldots n - 1 \} \) is fixed, \( j \in \{1 \ldots m \} \) and \( k \in \{m \ldots n - 1 \} \). Choose matrices \( U_j \in \mathbb{C}^{D_j \times D_j} \) and \( V_{k+1} \in \mathbb{C}^{D_k \times D_k} \). Set \( U_{\leq 0} = 1, V_{\geq n} = 1 \) and

\[
U_{\leq j} = U_j (U_{\leq j-1} \otimes 1_{d_j}) \in \mathbb{C}^{D_j \times d_{j-1} \times D_j},
\]

\[
V_{\geq k} = V_{k+1} (1_{d_{k+1}} \otimes V_{\geq k+1}) \in \mathbb{C}^{D_k \times d_{k+1} \times D_k}.
\] (A.28a)

In addition, set

\[
\tilde{U}_j = [(U_{\leq j-1} G_{\leq j-1}) \otimes 1_{d_j}] G_{k}^{L} \in \mathbb{C}^{D_j \times d_j \times D_j},
\]

\[
\tilde{V}_k = [1_{d_{k+1}} \otimes (V_{\geq k+1} G_{\geq k+1})] (G_{k+1}^{R})^{\top} \in \mathbb{C}^{d_{k+1} \times D_k \times D_k}.
\] (A.29a)

If the rank equalities

\[
\mathrm{rk}(U_j \tilde{U}_j) = \mathrm{rk}(\tilde{U}_j)
\] (A.30a)

\[
\mathrm{rk}(V_k \tilde{V}_k) = \mathrm{rk}(\tilde{V}_k)
\] (A.30b)

hold, then the following rank equalities hold as well:

\[
\mathrm{rk}(U_{\leq j} G_{\leq j}) = \mathrm{rk}(G_{\leq j})
\] (A.31a)

\[
\mathrm{rk}(V_{\geq k} G_{\geq k}) = \mathrm{rk}(G_{\geq k})
\] (A.31b)

\[
\mathrm{rk}(U_{\leq m} t_m (V_{\geq m})^{\top}) = \mathrm{rk}(t_m)
\] (A.31c)

**Proof** Note that

\[
U_{\leq j} G_{\leq j} = U_j \tilde{U}_j, \quad V_{\geq k} G_{\geq k} = V_k \tilde{V}_k.
\] (A.32)

The matrices \( \tilde{U}_j \) and \( \tilde{V}_k \) can be computed efficiently by using Eq. (A.32) to compute \( U_{\leq j} G_{\leq j} \) and \( V_{\geq k} G_{\geq k} \). If we refer to Proposition 3.1 in the remainder of the proof, we use the fact that \( \mathrm{rk}(LM) = \mathrm{rk}(M) \) implies \( \mathrm{rk}(LMR) = \mathrm{rk}(MR) \) for three matrices \( L, M \) and \( R \).

For \( j = 1 \), we have \( U_{\leq j} = U_j \) and \( \tilde{U}_j = G_j^L = G_{\leq j} \), i.e. (A.30a) implies (A.31a) for \( j = 1 \). Suppose that Eq. (A.31a) holds for some \( j - 1 \in \{1 \ldots m - 1 \} \), i.e.

\[
\mathrm{rk}(U_{\leq j-1} G_{\leq j-1}) = \mathrm{rk}(G_{\leq j-1})
\] (A.33)

which implies

\[
\mathrm{rk}((U_{\leq j-1} G_{\leq j-1}) \otimes 1_{d_j}) = \mathrm{rk}(G_{\leq j-1} \otimes 1_{d_j}).
\] (A.34)

This in turn implies (Proposition 3.1)

\[
\mathrm{rk}(\tilde{U}_j) = \mathrm{rk}([(U_{\leq j-1} G_{\leq j-1}) \otimes 1_{d_j}] G_j^L) = \mathrm{rk}([G_{\leq j-1} \otimes 1_{d_j}] G_j^L) = \mathrm{rk}(G_{\leq j}).
\] (A.35)
Then
\[ \text{rk}(U_{<j} G_{<j}) = \text{rk}(U_j \tilde{U}_j) = \text{rk}(\tilde{U}_j) = \text{rk}(G_{<j}) \]  
\tag{A.36}

where we used in turn Eq. (A.32), Eq. (A.30a) and the last but one equation. This shows that Eq. (A.31a) holds for \( j \in \{1 \ldots m\} \).

The proof of Eq. (A.31b) proceeds in the same way: For \( k = n - 1 \), we have \( V_{>k} = V_{k+1} \) and \( \tilde{V}_k = (G_{k+1}^R)^t = G_{>k} \), i.e. (A.30b) implies (A.31b) for \( k = n - 1 \). Suppose that (A.31b) holds for some \( k + 1 \in \{m + 1 \ldots n\} \), i.e.
\[ \text{rk}(\tilde{V}_k) = \text{rk}([I_{d_{k+1}} \otimes (V_{>k+1} G_{>k+1})](G_{k+1}^R)^t) = \text{rk}([I_{d_{k+1}} \otimes (G_{>k+1})](G_{k+1}^R)^t) \]

This implies
\[ \text{rk}(V_{>k} G_{>k}) = \text{rk}(V_k \tilde{V}_k) = \text{rk}(\tilde{V}_k) = \text{rk}(G_{>k}). \]  
\tag{A.37}

We have used, in turn, (A.32), (A.30b) and the last but one equation. This implies that (A.30b) holds for \( k \in \{m \ldots n - 1\} \).

The unfolding can be written as \( t_m = G_{\leq m}(G_{>m})^t \) (Eq. (A.24)). Applying Eqs. (A.31a) and (A.31b) and Proposition 3.1 provides
\[ \text{rk}(t_m) = \text{rk}(G_{\leq m}(G_{>m})^t) = \text{rk}(U_{\leq m} G_{\leq m}(G_{>m})^t) = \text{rk}(U_{\leq m} G_{\leq m}(G_{>m})^t (V_{>m})^t) \]
\[ = \text{rk}(U_{\leq m} t_m (V_{>m})^t). \]  
\tag{A.38}

This shows that Eq. (A.31c) holds and finishes the proof.

In the last Lemma, it was possible to choose \( U_j \) as submatrices of permutation matrices. The following Lemma shows that this implies that the \( U_{<j} \) are submatrices of permutation matrices as well and that the position of the non-zero entries of \( U_{<j} \) can be computed efficiently.

**Lemma A.3**  *In the following, let \( j \in \{1 \ldots m\} \). Choose matrices \( U_j \in \mathbb{C}^{D_j \times D_{j-1} d_j} \) which are submatrices of permutation matrices. Set \( U_{\leq 0} = 1 \) and set*
\[ U_{<j} = U_j(U_{\leq j-1} \otimes 1_{d_j}) \in \mathbb{C}^{D_j \times d_{j-1} \ldots d_1}. \tag{A.39} \]

*In the following, we also assume \( i_j \in \{1 \ldots d_j\} \) and \( b_j \in \{1 \ldots D_j\} \). Denote the set of unit entries in a given row of \( U_j \) and \( U_{<j} \) by*
\[ f_j(b_j) = \{(b_{j-1}, i_j) : [U_j]_{b_j, (b_{j-1}, i_j)} = 1\}, \tag{A.40a} \]
\[ f_{<j}(b_j) = \{(i_1, \ldots, i_j) : [U_{<j}]_{b_j, (i_1 \ldots i_j)} = 1\}. \tag{A.40b} \]
The latter set is given by

\[ f_{\leq j}(b_j) = \{(i_1, \ldots, i_j) : (b_{j-1}, i_j) \in f_j(b_j) \text{ and } (i_1, \ldots, i_{j-1}) \in f_{\leq j-1}(b_{j-1})\}, \tag{A.41} \]

and \( U_{\leq j} \) is a submatrix of a permutation matrix, i.e. we have \(|f_{\leq j}(b_j)| \leq 1\) and the entries of \( U_{\leq j} \) are given by

\[ [U_{\leq j}]_{b_j(i_1 \ldots i_j)} = \begin{cases} 1, & \text{if } (i_1, \ldots, i_j) \in f_{\leq j}(b_j), \\ 0, & \text{otherwise.} \end{cases} \tag{A.42} \]

**Proof** As \( U_j \) is a submatrix of a permutation matrix, we have \(|f_j(b_j)| \leq 1\) and the entries of \( U_j \) are given by

\[ [U_j]_{b_j(b_{j-1}, i_j)} = \begin{cases} 1, & \text{if } (b_{j-1}, i_j) \in f_j(b_j), \\ 0, & \text{otherwise.} \end{cases} \tag{A.43} \]

The entries of \( U_{\leq j} \) are given by (Eqs. (A.39) and (A.43))

\[ [U_{\leq j}]_{b_j(i_1 \ldots i_j)} = \sum_{b_{j-1}} [U_j]_{b_j(b_{j-1}, i_j)} [U_{\leq j-1}]_{b_{j-1},(i_1 \ldots i_{j-1})} = \sum_{b_{j-1} : (b_{j-1}, i_j) \in f_j(b_j)} [U_{\leq j-1}]_{b_{j-1},(i_1 \ldots i_{j-1})} \]

\[ = \begin{cases} [U_{\leq j-1}]_{b_{j-1},(i_1 \ldots i_{j-1})} \text{ where } (b_{j-1}, i_j) \in f_j(b_j) & \text{if } |f_j(b_j)| = 1, \\ 0 & \text{otherwise.} \end{cases} \]

Here, we used that there is at most one element \((b_{j-1}, i_j) \in f_j(b_j)\). Note that \(|f_{\leq j}(b_j)| \leq 1\) and Eq. (A.42) hold for \( j = 0 \) (as \( f_{\leq 0}(1) = \{1\} \)); assume that the two conditions hold for some \( j - 1 \in \{0 \ldots m - 1\} \). In the following, we show that they also hold for \( j \). If we apply the assumption to the last equation, we obtain

\[ [U_{\leq j}]_{b_j(i_1 \ldots i_j)} = \begin{cases} 1, & \text{if } (b_{j-1}, i_j) \in f_j(b_j) \text{ and } (i_1, \ldots, i_{j-1}) \in f_{\leq j-1}(b_{j-1}), \\ 0, & \text{otherwise.} \end{cases} \tag{A.45} \]

The set \( f_{\leq j}(b_j) \) has at most one element because \( f_j(b_j) \) has at most one element and because we assumed that \( f_{\leq j-1}(b_{j-1}) \) (for the single possible value of \( b_{j-1} \)) has at most one element as well. This finishes the proof.

**A.5. Sequence of local quantum operations as PMPS representation**

This section introduces the locally purified matrix product state (PMPS) representation and discusses the known fact that a sequentially prepared mixed quantum state can be represented as a PMPS or as an MPO (Lemma A.4). The PMPS representation \([27, 13]\) provides an alternative to the MPO representation for positive semidefinite operators such as mixed quantum
Suppose that a quantum state \( \rho \) is given in terms of \( n \) ancilla systems of dimensions \( d'_k \) with bases \( \{ |q_{i_k}^{(k)} \rangle \}_{k=1}^{d'_k} \). A PMPS representation of \( \rho \) is given by

\[
\rho = \text{Tr}_{Y_1 \ldots Y_n} (|\Psi \rangle \langle \Psi|)
\]

\[
\{ \phi_{i_1}^{(1)} | q_{i_2}^{(2)} \ldots q_{i_n}^{(n)} | \Psi \rangle \} = G_1(i_1, i'_1) G_2(i_2, i'_2) \ldots G_n(i_n, i'_n) \]

\[
G_k(i_k, i'_k) \in \mathbb{C}^{D_{k-1} \times D_k}
\]

where the overline denotes the complex conjugate. Equation (A.47) shows that given a PMPS representation of bond dimension \( D \), we can directly construct an MPO representation with bond dimension \( D^2 \). However, an MPO representation with bond dimensions smaller than \( D^2 \) can exist. It has been shown that there is a family of quantum states on \( n \) systems which can be represented as an MPO with bond dimension independent of \( n \) but the bond dimension of any PMPS representation of those states increases with \( n \) [13]. This is an advantage of the MPO representation, but on the other hand, deciding whether a given MPO representation represents a positive semidefinite operator is an NP-hard problem, i.e. a solution in polynomial (in \( n \)) time is unlikely [19]. The PMPS representation has the advantage that it always represents a positive semidefinite operator by definition. The relative merits of the MPO and PMPS representations of a mixed quantum state depend on the application.

Suppose that a quantum state \( \rho \in \mathcal{D}(\mathcal{H}_{1 \ldots n}) \) was prepared via quantum operations \( \mathcal{W}_k : \mathcal{B}(\mathcal{H}_{k-1}) \rightarrow \mathcal{B}(\mathcal{H}_{k-1,k}) \), i.e.

\[
\rho = \mathcal{W}_n \mathcal{W}_{n-1} \ldots \mathcal{W}_3 \mathcal{W}_2 (\sigma), \quad \sigma \in \mathcal{D}(\mathcal{H}_1).
\]

Clearly, this is an efficient representation of the quantum state \( \rho \) as it is described by at most \( nd^6 \) parameters. It is known that such a representation can be efficiently – i.e. with at most \( \text{poly}(n) \) computational time – converted into a PMPS representation or a MPO representation [18, 19]. The following Lemma provides the technical details of the conversion:

**Lemma A.4** Let \( Y'_k (k \in \{ 1, \ldots, n \}) \) be systems with \( \text{dim}(Y'_k) = 1 \). Let \( \rho_1 \in \mathcal{B}(1, Y'_1) \) and \( \mathcal{W}_k \in \mathcal{B}(\mathcal{B}(Y'_{k-1}); \mathcal{B}(k, Y'_k)) \). A linear operator \( \rho \in \mathcal{B}(\mathcal{H}_{1 \ldots n}) \) is given by

\[
\rho = \rho_{n-1} \rho_{k} = (\text{id} \otimes \mathcal{W}_k)(\rho_{k-1}) \quad (k \in \{ 2, \ldots, n \}).
\]

Let \( F_{i_1}^{(k)} \) and \( F_{b_k}^{(Y'_k)} \) be orthonormal operator bases of systems \( k \) and \( Y'_k \). An MPO representation of \( \rho \) is given by

\[
[G_1(i_1)]_{b_0, b_1} = \langle F_{i_1}^{(1)} \otimes F_{b_1}^{(Y'_1)}, \rho_1 \rangle,
\]

\[
[G_k(i_k)]_{b_{k-1}, b_k} = \langle F_{i_k}^{(k)} \otimes F_{b_k}^{(Y'_k)}, \mathcal{W}_k(F_{b_{k-1}}^{(Y'_{k-1})}) \rangle, \quad (k \in \{ 2, \ldots, n \}).
\]
The bond dimensions of the representation are given by $D_k = (d_{\Gamma_k})^2$ ($k \in \{1, \ldots, n-1\}$, $D_0 = D_n = 1$). Now, let $\rho_1$ be a quantum state and $W_k$ be CPTP linear maps. In addition, let

$$\rho_1 = \sum_{i=1}^{r_1} |\psi_i\rangle\langle\psi_i|,$$

$$W_k(\cdot) = \sum_{i=1}^{r_1} E_i^{(k)} \cdot (E_i^{(k)})^\ast \quad (k \in \{2, \ldots, n\}) \quad (A.51)$$

be a decomposition of $\rho_1$ into $r_1 = \text{rk}(\rho_1)$ orthogonal vectors and a Kraus decomposition of $W_k$ where $r_k$ equals the Kraus rank of $W_k$ (i.e. $r_1 \leq d_1 d_{\Gamma_1}$ and $r_k \leq d_k d_{\Gamma_k} d_{\Gamma_{k-1}}$). Let $\{|i_k\rangle\}_{i_k}$ and $\{|b_k\rangle\}_{b_k}$ be orthonormal bases of $H_k$ and $\Gamma_k$, respectively. A PMPS representation of $\rho$ is given by

$$[G_1(i_1, i_1')]_{b_0, b_1} = \langle i_1 | \gamma_1^1 \langle b_1 | \psi_{i_1}' \rangle, \quad (A.52a)$$

$$[G_k(i_k, i_k')]_{b_{k-1}, b_k} = \langle i_k | \gamma_k \langle b_k | E_i^{(k)} \rangle \langle b_{k-1} | \gamma_{k-1} \rangle = \langle k \in \{2, \ldots, n\} \rangle. \quad (A.52b)$$

The bond dimensions of the representation are given by $D_k = d_{\Gamma_k}^2$ ($k \in \{1, \ldots, n-1\}$, $D_0 = D_n = 1$) and the ancilla dimensions are given by $r_k$.

**Proof** For the MPO representation, evaluate the operator basis elements of $\rho$ from Eq. (A.49) and compare with the operator basis elements of the representation (Eq. (2.15)). For the PMPS representation, evaluate the matrix entries of $\rho$ from Eq. (A.49) inserting Eq. (A.51) and compare with the matrix entries of the representation (Eq. (A.46)).

**A.6. General linear maps as measurements**

Consider a quantum state $\rho \in D(Y)$ and an arbitrary linear map $N \in \mathcal{B}(\mathcal{B}(Y); \mathcal{B}(X))$. Below, we work with linear maps $N$ which are not necessarily CPTP and therefore do not represent a physical operation on the quantum state. Such a map $N$ is of relevance only if $N(\rho)$ can be obtained from the outcomes of physical measurements on $\rho$. In this section, we show how this can be achieved, allowing the reconstruction scheme from Section 3.3 to be used for quantum state tomography.

Firstly, we construct a set of observables $G_i \in \mathcal{B}(Y)$ ($i \in \{1, \ldots, 2d_X^2\}$) whose expectation values $\text{Tr}(G_i \rho)$ in the state $\rho$ can be used to compute $N(\rho)$. Secondly, we construct a POVM with $2d_X^2 + 1$ elements $E_i \in \mathcal{B}(Y)$ such that the outcome probabilities $\text{Tr}(E_i \rho)$ in the state $\rho$ can also be used to compute $N(\rho)$.

We denote the components of $N(\rho)$ in an operator basis $F_i^{(X)}$ of $X$ by $s_i$:

$$s_i = \langle F_i^{(X)}, N(\rho) \rangle \quad i \in \{1, \ldots, d_X^2\} \quad d_X = \dim(X) \quad (A.53)$$

The key tool is the following property of the map $N^*$:

$$s_i = \langle N^*(F_i^{(X)}), \rho \rangle = \text{Tr}(H_i \rho), \quad H_i = \left( N^*(F_i^{(X)}) \right)^* \quad (A.54)$$
Since $H_i$ may not be Hermitian, we use its Hermitian and skew-Hermitian components:

$$G_i = \frac{1}{2}(H_i + H_i^*) \quad G_{i+d_X^2} = \frac{1}{2i}(H_i - H_i^*) \quad H_i = G_i + \text{i}G_{i+d_X^2} \quad (A.55)$$

Using the observables $G_i$, the components $s_i$ can be expressed as follows:

$$s_i = \langle N^*(F_i^{(X)}), \rho \rangle = \text{Tr}(G_i \rho) + \text{i} \text{Tr}(G_{i+d_X^2} \rho) \quad (A.56)$$

In other words, the expectation values of the $G_i$ provide the real and imaginary parts of $s_i$:

$$\text{Re}(s_i) = \text{Tr}(G_i \rho) \quad \text{Im}(s_i) = \text{Tr}(G_{i+d_X^2} \rho) \quad (A.57)$$

If these expectation values can be measured, we already obtain a way to obtain $N(\rho)$ from physical measurements on $\rho$ even if $N$ is not CPTP. Furthermore, we construct a POVM whose measurement on $\rho$ also allows to determine $N(\rho)$. We choose coefficients $c_i \in \mathbb{R}$ and $c > 0$ such that the following operators become positive semidefinite:

$$G_i + c_i 1 \geq 0 \quad 1 - c \sum_{i=1}^{2d_X^2} (G_i + c_i 1) \geq 0 \quad (A.58)$$

We define $2d_X^2 + 1$ POVM elements by

$$E_0 = 1 - \sum_{i=1}^{2d_X^2} E_i \quad E_i = c(G_i + c_i 1) \quad (A.59)$$

Clearly, the expectation values of the $G_i$ are related to the POVM probabilities by

$$\text{Tr}(G_i \rho) = \frac{1}{c} (\text{Tr}(E_i \rho) - c_i). \quad (A.60)$$

The coefficients $s_i$ of $N(\rho)$ can be obtained from these expectation values using Eq. (A.57). As a consequence, the POVM probabilities of the given POVM allow us to determine $N(\rho)$ even if $N$ is not CPTP.

B. References

[1] T. Baumgratz, D. Gross, M. Cramer, and M. B. Plenio. “Scalable Reconstruction of Density Matrices”. *Phys. Rev. Lett.* 111(2) (2013), p. 020401. doi: 10.1103/physrevlett.111.020401. arXiv: 1207.0358.

[2] D. Petz. “Monotonicity of Quantum Relative Entropy Revisited”. *Rev. Math. Phys.* 15 (2003), pp. 79–91. doi: 10.1142/S0129055X03001576. arXiv: quant-ph/0209053.

[3] P. Hayden, R. Jozsa, D. Petz, and A. Winter. “Structure of States Which Satisfy Strong Subadditivity of Quantum Entropy with Equality”. *Commun. Math. Phys.* 246 (2004), pp. 359–374. doi: 10.1007/s00220-004-1049-z. arXiv: quant-ph/0304007.
[4] S. A. Goreinov, E. E. Tyrtyshnikov, and N. L. Zamarashkin. “A theory of pseudoskeleton approximations”. *Linear Algebra Appl.* **261**(1) (1997), pp. 1–21. doi: [10.1016/S0024-3795(96)00301-1](https://doi.org/10.1016/S0024-3795(96)00301-1).

[5] C. F. Caiafa and A. Cichocki. “Generalizing the column–row matrix decomposition to multi-way arrays”. *Linear Algebra Appl.* **433**(3) (2010), pp. 557–573. doi: [10.1016/j.laa.2010.03.020](https://doi.org/10.1016/j.laa.2010.03.020).

[6] F. Hiai and D. Petz. “The proper formula for relative entropy and its asymptotics in quantum probability”. *Commun. Math. Phys.* **143**(1) (1991), pp. 99–114. doi: [10.1007/BF02100287](https://doi.org/10.1007/BF02100287). https://projecteuclid.org/euclid.cmp/1104248844.

[7] T. Ogawa and H. Nagaoka. “Strong converse and Stein’s lemma in quantum hypothesis testing”. *IEEE Trans. Inf. Theory* **46**(7) (2000), pp. 2428–2433. doi: [10.1109/18.887855](https://doi.org/10.1109/18.887855). arXiv: quant-ph/9906090.

[8] D. Petz. “Sufficient subalgebras and the relative entropy of states of a von Neumann algebra”. *Commun. Math. Phys.* **105**(1) (1986), pp. 123–131. doi: [10.1007/BF01212345](https://doi.org/10.1007/BF01212345). https://projecteuclid.org/euclid.cmp/1104115260.

[9] D. Petz. “Sufficiency of channels over von Neumann algebras”. *Quart. J. Math.* **39**(1) (1988), pp. 97–108. doi: [10.1093/qmath/39.1.97](https://doi.org/10.1093/qmath/39.1.97).

[10] M. Piani, P. Horodecki, and R. Horodecki. “No-Local-Broadcasting Theorem for Multi-partite Quantum Correlations”. *Phys. Rev. Lett.* **100**(9) (2008), p. 090502. doi: [10.1103/PhysRevLett.100.090502](https://doi.org/10.1103/PhysRevLett.100.090502). arXiv: 0707.0848 [quant-ph].

[11] M. A. Nielsen, C. M. Dawson, J. L. Dodd, A. Gilchrist, D. Mortimer, T. J. Osborne, M. J. Bremner, A. W. Harrow, and A. Hines. “Quantum dynamics as a physical resource”. *Phys. Rev. A* **67**(5) (2003), p. 052301. doi: [10.1103/PhysRevA.67.052301](https://doi.org/10.1103/PhysRevA.67.052301). arXiv: quant-ph/0208077.

[12] A. Datta and G. Vidal. “Role of entanglement and correlations in mixed-state quantum computation”. *Phys. Rev. A* **75**(4) (2007), p. 042310. doi: [10.1103/PhysRevA.75.042310](https://doi.org/10.1103/PhysRevA.75.042310). arXiv: quant-ph/0611157.

[13] G. De las Cuevas, N. Schuch, D. Pérez-García, and J. I. Cirac. “Purifications of multipartite states: limitations and constructive methods”. *New J. Phys.* **15**(12) (2013), p. 123021. doi: [10.1088/1367-2630/15/12/123021](https://doi.org/10.1088/1367-2630/15/12/123021). arXiv: 1308.1914 [quant-ph].

[14] U. Schollwöck. “The density-matrix renormalization group in the age of matrix product states”. *Ann. Phys.* **326**(1) (2011), pp. 96–192. doi: [10.1016/j.aop.2010.09.012](https://doi.org/10.1016/j.aop.2010.09.012). arXiv: 1008.3477.

[15] 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”. *Nat. Commun.* **1**(9) (2010), p. 149. doi: [10.1038/ncomms1147](https://doi.org/10.1038/ncomms1147). arXiv: 1101.4366.
[16] T. Baumgratz, A. Nüßeler, M. Cramer, and M. B. Plenio. “A scalable maximum likelihood method for quantum state tomography”. New J. Phys. 15(12) (2013), p. 125004. doi: 10.1088/1367-2630/15/12/125004. arXiv: 1308.2395.

[17] B. P. Lanyon, C. Maier, M. Holzäpfel, T. Baumgratz, C. Hempel, P. Jurcevic, I. Dhand, A. S. Buyskikh, A. J. Daley, M. Cramer, M. B. Plenio, R. Blatt, and C. F. Roos. “Efficient tomography of a quantum many-body system”. Nat. Phys. (2017). doi: 10.1038/nphys4244. arXiv: 1612.08000 [quant-ph].

[18] D. Poulin and M. B. Hastings. “Markov Entropy Decomposition: A Variational Dual for Quantum Belief Propagation”. Phys. Rev. Lett. 106(8) (2011), p. 080403. doi: 10.1103/physrevlett.106.080403. arXiv: 1012.2050.

[19] M. Kliesch, D. Gross, and J. Eisert. “Matrix-Product Operators and States: NP-Hardness and Undecidability”. Phys. Rev. Lett. 113(16) (2014), p. 160503. doi: 10.1103/PhysRevLett.113.160503. arXiv: 1404.4466 [quant-ph].

[20] I. V. Oseledets and E. E. Tyrtyshnikov. “TT-cross approximation for multidimensional arrays”. Linear Algebra Appl. 432(1) (2010), pp. 70–88. doi: 10.1016/j.laa.2009.07.024.

[21] J. Ballani, L. Grasedyck, and M. Kluge. “Black box approximation of tensors in hierarchical Tucker format”. Linear Algebra Appl. 438(2) (2013), pp. 639–657. doi: 10.1016/j.laa.2011.08.010.

[22] C. F. Caiafa and A. Cichocki. “Stable, Robust, and Super Fast Reconstruction of Tensors Using Multi-Way Projections”. IEEE Trans. Signal Process. 63(3) (2015), pp. 780–793. doi: 10.1109/TSP.2014.2385040.

[23] R. Penrose. “A generalized inverse for matrices”. Math. Proc. Cambridge Philos. Soc. 51 (03 1955), pp. 406–413. doi: 10.1017/S0305004100030401.

[24] R. Rado. “Note on generalized inverses of matrices”. Math. Proc. Cambridge Philos. Soc. 52 (03 1956), pp. 600–601. doi: 10.1017/S0305004100031601.

[25] I. V. Oseledets. “Tensor-train decomposition”. SIAM J. Sci. Comput. 33(5) (2011), pp. 2295–2317. doi: 10.1137/090752286.

[26] M. Zwolak and G. Vidal. “Mixed-State Dynamics in One-Dimensional Quantum Lattice Systems: A Time-Dependent Superoperator Renormalization Algorithm”. Phys. Rev. Lett. 93(20) (2004), p. 207205. doi: 10.1103/PhysRevLett.93.207205. arXiv: cond-mat/0406440.

[27] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac. “Matrix Product Density Operators: Simulation of Finite-Temperature and Dissipative Systems”. Phys. Rev. Lett. 93(20) (2004), p. 207204. doi: 10.1103/PhysRevLett.93.207204. arXiv: cond-mat/0406426.

[28] H.Umegaki. “Conditional expectation in an operator algebra, IV (Entropy and information)”. Kodai Math. Sem. Rep. 14(2) (1962), pp. 59–85. doi: 10.2996/kmj/113844604.
[29] G. W. Stewart. “On the Perturbation of Pseudo-Inverses, Projections and Linear Least Squares Problems”. SIAM Rev. 19(4) (1977), pp. 634–662. doi: 10.1137/1019104. http://www.jstor.org/stable/2030248.

[30] P.-Å. Wedin. “Perturbation theory for pseudo-inverses”. BIT Numer. Math. 13(2) (1973), pp. 217–232. doi: 10.1007/BF01933494.

[31] A. Peres. “Separability Criterion for Density Matrices”. Phys. Rev. Lett. 77 (1996), pp. 1413–1415. doi: 10.1103/PhysRevLett.77.1413. arXiv: quant-ph/9604005.

[32] M. Horodecki, P. Horodecki, and R. Horodecki. “Separability of mixed states: necessary and sufficient conditions”. Phys. Lett. A 223 (1996), pp. 1–8. doi: 10.1016/S0375-9601(96)00706-2. arXiv: quant-ph/9605038.

[33] M. A. Nielsen and I. L. Chuang. Quantum Computation and Quantum Information. 9th ed. Cambridge: Cambridge University Press, 2007. doi: 10.1017/CBO9780511976667.

[34] D. V. Savostyanov. “Quasioptimality of maximum-volume cross interpolation of tensors”. Linear Algebra Appl. 458 (2014), pp. 217–244. doi: 10.1016/j.laa.2014.06.006.

[35] D. V. Savostyanov and I. V. Oseledets. “Fast adaptive interpolation of multi-dimensional arrays in tensor train format”. In: The 2011 International Workshop on Multidimensional (nD) Systems. 2011, pp. 1–8. doi: 10.1109/nDS.2011.6076873.

[36] M. Holzäpfel. To be published. PhD thesis. Universität Ulm.

[37] R. A. Horn and C. R. Johnson. Topics in Matrix Analysis. 1st ed. Cambridge: Cambridge University Press, 1991.

[38] D. Kressner, M. Steinlechner, and A. Uschmajew. “Low-Rank Tensor Methods with Subspace Correction for Symmetric Eigenvalue Problems”. SIAM J. Sci. Comput. 36(5) (2014), A2346–A2368. doi: 10.1137/130949919.