IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Purifications of multipartite states: limitations and constructive methods

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2013 New J. Phys. 15 123021

(http://iopscience.iop.org/1367-2630/15/12/123021)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 130.183.90.175
This content was downloaded on 05/02/2014 at 10:54

Please note that terms and conditions apply.

New Journal of Physics
The open access journal for physics

Purifications of multipartite states: limitations and constructive methods

Gemma De las Cuevas[1,4], **Norbert Schuch**[2], **David Pérez-García**[3] **and J Ignacio Cirac**[1]

[1] Max Planck Institute for Quantum Optics, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany
[2] Institut für Quanteninformation, RWTH Aachen, Aachen, Germany
[3] Departamento de Análisis Matemático, Facultad de CC Matemáticas, Universidad Complutense de Madrid, E-28040 Madrid, Spain
E-mail: gemma.delascuevas@mpq.mpg.de

New Journal of Physics **15** (2013) 123021 (25pp)
Received 9 August 2013
Published 10 December 2013
Online at http://www.njp.org/
doi:10.1088/1367-2630/15/12/123021

Abstract. We analyze the description of quantum many-body mixed states using matrix product states and operators. We consider two such descriptions: (i) as a matrix product density operator of bond dimension D; and (ii) as a purification that is written as a matrix product state of bond dimension D'. We show that these descriptions are inequivalent in the sense that D' cannot be upper bounded by D only. Then we provide two constructive methods to obtain (ii) out of (i). The sum of squares (sos) polynomial method scales exponentially in the number of different eigenvalues, and its approximate version is formulated as a semidefinite program, which gives efficient approximate purifications whose D' only depends on D. The eigenbasis method scales quadratically in the number of eigenvalues, and its approximate version is very efficient for rapidly decaying distributions of eigenvalues. Our results imply that a description of mixed states which is both efficient and locally positive semidefinite does not exist, but that good approximations do.

[4] Author to whom any correspondence should be addressed.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

© IOP Publishing Ltd and Deutsche Physikalische Gesellschaft

Contents

1. Introduction **2**
2. The setting **3**
2.1. Definitions . . . 3
2.2. The problem . . . 4
3. Inequivalence of the two forms **5**
3.1. Inequivalence for bipartite classical states . . . 5
3.2. Inequivalence for multipartite classical states . . . 7
4. Purification methods **9**
4.1. sos polynomial method . . . 9
4.2. Eigenbasis method . . . 14
4.3. Comparison of approximation methods . . . 16
5. Conclusions and outlook **20**
Acknowledgments **21**
Appendix A. Semidefinite program **21**
Appendix B. sos polynomials for the exponentially decaying distribution of eigenvalues **23**
References **24**

1. Introduction

Quantum many-body systems appear in a variety of fields in physics, such as condensed matter, quantum chemistry or high-energy physics. Since their Hilbert space description is intractable (as it scales exponentially with the system size), a number of methods have been proposed to describe them efficiently. One of them are tensor networks [1], which have been particularly successful in describing one-dimensional pure states with matrix product states (MPS) [2]. Its canonical form has facilitated the distinction between injective and non-injective MPS [3], (which determines the ground state degeneracy of the parent Hamiltonian, and is linked to many other physical properties, see e.g. [4–6]), which has led to the classification of gapped phases of one-dimensional systems [7]. This mathematical understanding has allowed to characterize global properties of the state, such as topological order or symmetries, in a local way [8, 9]. This is in contrast with one-dimensional *mixed states*, whose description with tensor networks is much more scarce. While the class of matrix product density operators (MPDOs) has been defined [10, 11], a canonical form has not been found. This is despite the fact that mixed states are used to describe quantum many-body systems at finite temperature, systems out of equilibrium, dissipative dynamics or lack of knowledge of the state of the system.

One of the difficulties in defining a canonical form for mixed states is that with MPDOs one cannot verify locally that the global tensor is positive semidefinite. This implies that local truncations of the tensors generally destroy the global positivity, causing instability in numerical algorithms. An alternative is to use the local purification form, where the mixed state is purified and the purification is expressed as an MPS. While its local tensors are positive semidefinite, it is not known if an efficient MPDO form guarantees an efficient local purification. That is, are the two descriptions equivalent, or can the local purification be much more costly than the MPDO form (see figure 1)?

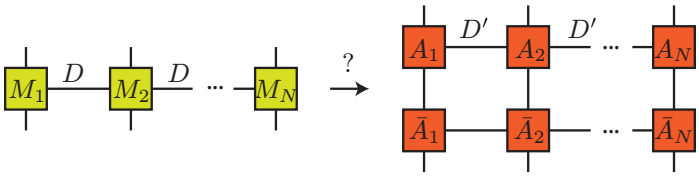


Figure 1. (Left) the MPDO form of a mixed state writes it as a sum of D tensor products. The matrices M_i are generally not positive semidefinite. (Right) the local purification form of ρ, where the purifying state is written as an MPS of bond dimension D'. The question is: can D' be upper bounded by D?

In this paper we address this question and show that these two descriptions are *inequivalent*. Specifically, we provide classical multipartite states whose MPDO form has a fixed cost, but whose local purification form has an unbounded cost. Then, we provide *two constructive purification methods* applicable to all multipartite density matrices, which relate the two forms but also involve the number of (different) eigenvalues. The sum of squares (sos) polynomial method has an exact version which scales exponentially in the number of different eigenvalues. Its approximate version is formulated as a semidefinite program (SDP), which shows an efficient and robust behavior for all the tested distributions of eigenvalues. The eigenbasis method has an exact version which scales multiplicatively with the number of eigenvalues, and its approximate version gives very efficient purifications for rapidly decaying distributions of eigenvalues.

This paper is organized as follows. First, we present the problem in section 2. Then we show the inequivalence of the MPDO form and the local purification form in section 3. In section 4 we present the two purification methods: the sos polynomial method (section 4.1), and the eigenbasis method (section 4.2), both with its main idea, its exact and its approximate version, and in section 4.3 we compare the two approximate methods. Finally we conclude and mention further directions in section 5.

2. The setting

We now present the question that concerns us, which is whether the MPDO form and the local purification form of a mixed state can be related. We will first present some notation and definitions (section 2.1), and then introduce the problem (section 2.2).

2.1. Definitions

Let us first fix some notation. We write $A \succeq 0$ to denote that the matrix A is positive semidefinite (i.e. Hermitian with non-negative eigenvalues), and $A \geqslant 0$ to denote that it is non-negative (i.e. with non-negative entries). We also write $\mathrm{rank}(\rho)$ ($\mathrm{rank}(A)$) to denote the number of non-zero eigenvalues (non-zero singular values) of the matrix ρ (A). The trace norm of a matrix A is defined as $\|A\|_1 = \sum_{i=1}^{r} s_i$, where $r = \mathrm{rank}(A)$ and s_i are its singular values. Given a multipartite state with N subsystems arranged on a one-dimensional line (henceforth called a 1D state), we call a linear bipartition a splitting of the form $1, \ldots, k$ versus $k+1, \ldots, N$ for any k. Finally, given a multipartite state $|\psi\rangle = \sum \psi_{i_1 \ldots i_N} |i_1 \ldots i_N\rangle$ with Schmidt rank D'_k across the bipartition $i_1, \ldots, i_k$ versus the rest, we define the Schmidt rank of $|\psi\rangle$ as $\mathrm{SR}(\psi) = \max_k D'_k$.

We focus on 1D mixed states with N d-level systems and open boundary conditions, described by a density matrix ρ,

$$\rho = \sum_{i_1, j_1, \ldots, i_N, j_N = 1}^{d} \rho_{i_1,\ldots,i_N}^{j_1,\ldots,j_N} |i_1, \ldots, i_N\rangle\langle j_1, \ldots, j_N|. \tag{1}$$

The are two natural ways of describing ρ locally. The first one is the *MPDO form* [10, 11], which is defined as

$$\rho = \sum_{\alpha_1=1}^{D_1} \ldots \sum_{\alpha_{N-1}=1}^{D_{N-1}} M_1^{\alpha_1} \otimes M_2^{\alpha_1,\alpha_2} \otimes \ldots \otimes M_N^{\alpha_{N-1}}, \tag{2}$$

where $M_k^{\alpha_{k-1},\alpha_k}$ are $d \times d$ matrices for $1 < k < N$, and $M_1^{\alpha_1}$ ($M_N^{\alpha_{N-1}}$) is a row (column) vector of size d. Note that the subindex only indicates the subsystem that the matrix is describing. Here D_k (for all k) is the minimal dimension such that (2) holds[5]. The *operator Schmidt rank* of ρ is defined as[6]

$$\mathrm{OSR}(\rho) := \max_k D_k = D. \tag{3}$$

The second form is the *local purification form*, which is obtained by purifying the mixed state ρ (living in system S) into a pure state $|\Psi\rangle$ (living in S, S'), and expressing $|\Psi\rangle$ as an MPS[7]

$$\rho = \mathrm{Tr}_{S'} |\Psi\rangle\langle\Psi| \tag{4}$$

$$|\Psi\rangle = \sum_{\beta_1=1}^{D'_1} \ldots \sum_{\beta_{N-1}=1}^{D'_{N-1}} A_1^{\beta_1} \otimes A_2^{\beta_1,\beta_2} \otimes \ldots \otimes A_N^{\beta_{N-1}}. \tag{5}$$

Here $A_k^{\beta_{k-1},\beta_k}$ are $d \times d_{\mathrm{a}_k}$ matrices for $1 < k < N$, where d_{a_k} is the dimension of the local ancilla, and $A_1^{\beta_1}$ ($A_N^{\beta_{N-1}}$) is a row (column) vector of size $d \times d_{\mathrm{a}_1}$ ($d \times d_{\mathrm{a}_{N-1}}$). Here D'_k is the Schmidt rank of $|\Psi\rangle$ of the bipartition $\mathrm{Tr}_{k+1\ldots N}|\Psi\rangle\langle\Psi|$ versus the rest. We define the *purification rank* of ρ as

$$\mathrm{rank}_{\mathrm{puri}}(\rho) := \max_k D'_k = D'. \tag{6}$$

2.2. The problem

We want to find out if the MPDO and the local purification form are equivalent, or if the latter can be arbitrarily more costly than the former. The advantage of the local purification form is that the local tensors are positive semidefinite, since they are of the form $A_k A_k^\dagger$. In contrast, in the MPDO form this is generally not true, i.e. $M_k^{\alpha_{k-1},\alpha_k} \not\succeq 0$, thus it is locally invisible that

[5] The minimal D_k (for all k) is unique as it does not depend on the order on which the decompositions are made. To see this, note that D_k is the bond dimension of the bipartition $i_1 j_1 \ldots i_k j_k$ versus the rest of the pure state $|\psi\rangle = \sum \rho_{i_1,\ldots,i_N}^{j_1,\ldots,j_N} |i_1, \ldots, i_N\rangle|j_1, \ldots, j_N\rangle$.

[6] The operator Schmidt rank of ρ must not be confused with its tensor rank, which is the minimal r such that $\rho = \sum_{\alpha=1}^{r} M_1^\alpha \otimes M_2^\alpha \otimes \ldots \otimes M_N^\alpha$.

[7] There are infinitely many purifications of a given mixed state, and they are all related by an isometry on the ancillary system. We always consider the purification which minimizes the Schmidt rank of the purifying state along every linear bipartition and which is local (i.e. an ancillary subsystem S'_i is attached to each subsystem S_i).

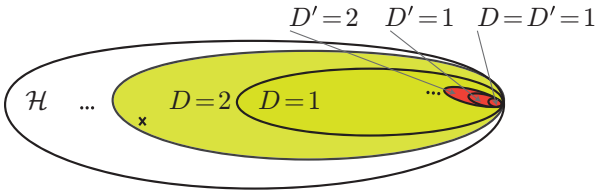


Figure 2. The Hilbert space $\mathcal{H}$ of mixed states. How much larger is the family of states with fixed operator Schmidt rank D compared to that with fixed purification rank D'? Are they comparable in size, or are there states requiring very small D and very large D' (such as the one marked with a cross)?

$\rho \succeq 0$. This is a problem theoretically as well as numerically, as local truncations destroy the global positivity.

On the other hand, the problem of the local purification form is that it is not known how much larger the purification rank D' may be compared to the operator Schmidt rank D (see figure 2). Thus, we want to see if there is a transformation from the MPDO to the local purification form in which the bond dimension increases in a controlled way. That is, we want to find out if D' can be upper bounded by a function of D.

Observe that it is very easy to obtain a bound in the other direction, namely $D \leqslant D'^2$, simply by choosing

$$[M_k]_{i_k,j_k} = \sum_{z_k=1}^{d_{a_k}} [A_k]_{i_k,z_k} \otimes ([A_k]_{j_k,z_k})^\dagger, \tag{7}$$

where i_k, j_k, z_k are physical indices. On the other hand, both ranks can be upper bounded by the physical dimension,

$$D \leqslant d^N, \quad D' \leqslant d^N. \tag{8}$$

3. Inequivalence of the two forms

We now show that the MPDO form and the local purification form are inequivalent. More precisely, we provide a family of multipartite classical states with a constant operator Schmidt rank across every linear bipartition, and an unbounded purification rank (result 2). We will first show the separation for bipartite states (section 3.1), and then for multipartite states (section 3.2).

3.1. Inequivalence for bipartite classical states

We now show the inequivalence of the two forms for bipartite classical states, which follows from recent results in convex analysis shown by Gouveia *et al* [12].

Result 1. *Let ρ denote a bipartite classical mixed state. Let $D = OSR(\rho)$ and $D' = \mathrm{rank}_{\mathrm{puri}}(\rho)$. Then D' cannot be upper bounded by a function f that only depends on D (in particular, f does not depend on ρ), i.e.*

$$\nexists f: \quad D' \leqslant f(D). \tag{9}$$

Before proving the result, we establish a relation between the MPDO and the local purification form of classical states and two other known decompositions. Consider the family of classical bipartite states (i.e. diagonal in the computational basis $\{|x\rangle\}$) of physical dimension t

$$\rho_t = \sum_{x,y=1}^{t} S_t(x, y)|x, y\rangle\langle x, y|, \tag{10}$$

where S_t is a non-negative matrix of size $t \times t$ and $S_t(x, y)$ denotes its x, y component. Now, the MPDO form of ρ_t corresponds to the singular value decomposition of S_t. Hence, the operator Schmidt rank of ρ_t corresponds to the rank of S_t, $\mathrm{rank}(S_t)$. On the other hand, the local purification form of ρ_t corresponds to the positive semidefinite factorization of S_t, in which S_t is expressed as

$$S_t(x, y) = \mathrm{Tr}(E_x F_y) \quad E_x, F_y \succeq 0 \tag{11}$$

for all x, y (the local purification form is obtained by expressing $E_x = A_x A_x^\dagger$ and $F_y = B_y B_y^\dagger$). Thus, the purification rank of ρ_t corresponds to the positive semidefinite rank of S_t, $\mathrm{rank}_{\mathrm{psd}}(S_t)$, which is defined as the minimal r such that there exist matrices $E_x, F_y \succeq 0$ of size $r \times r$ such that (11) holds (see figure 3). The positive semidefinite factorization was very recently introduced in [13] and shown to be related to the quantum communication complexity [13] and the quantum correlation complexity of S_t [14]. In summary, for classical states such as ρ_t, it holds that[8]

$$\begin{aligned} &\mathrm{OSR}(\rho_t) = \mathrm{rank}(S_t), \\ &\mathrm{rank}_{\mathrm{puri}}(\rho_t) = \mathrm{rank}_{\mathrm{psd}}(S_t). \end{aligned} \tag{12}$$

Now we are ready to prove result 1.

Proof. We consider classical bipartite states of the form (10), and we focus on a class of non-negative matrices S_t called slack matrices of polytopes, defined as follows (see e.g. [12]). A convex polytope is defined as the intersection of a finite set of halfspaces $\{h_j(x) \leqslant b_j\}$, or as the convex hull of a set of vertices $\{v_i\}$ [16]. Its slack matrix S is defined so that its (i, j) entry contains the distance from hyperplane j to vertex i, i.e. $S(i, j) = b_j - h_j(v_i)$ (see figure 4).

Now we let S_t be the slack matrix of the two-dimensional (2D) regular polytope with t vertices (and thus also t faces), called the regular t-gon[9]. Gouveia *et al* [12] show that[10]

$$\begin{aligned} &\mathrm{rank}(S_t) = 3 \quad \forall t \\ &\mathrm{rank}_{\mathrm{psd}}(S_t) \sim \log t. \end{aligned} \tag{13}$$

[8] The non-negative factorization [15] of non-negative matrices such as S_t has also been defined, and has traditionally received more attention, in particular in its connections to classical communication complexity. In this case, S_t is expressed as a product of two non-negative matrices X and Y. The non-negative rank of S_t, $\mathrm{rank}_+(S_t)$, is the minimal r such that there exist $X, Y \geqslant 0$ of size $t \times r$ and $r \times t$, respectively, such that $S_t = XY$. Note that this factorization cannot be defined for general quantum states.

[9] If centered at the 0 point, the regular t-gon is defined by $P_t = \mathrm{conv}\{(\cos(2\pi k/t), \sin(2\pi k/t)) : 0 \leqslant k < t\} \subseteq \mathbb{R}^2$ [16].

[10] More precisely, they show that $\Omega(\log t) \leqslant \mathrm{rank}_{\mathrm{psd}}(S_t) \leqslant \log t$. We use the O and Ω notation as customary, see e.g. [17].

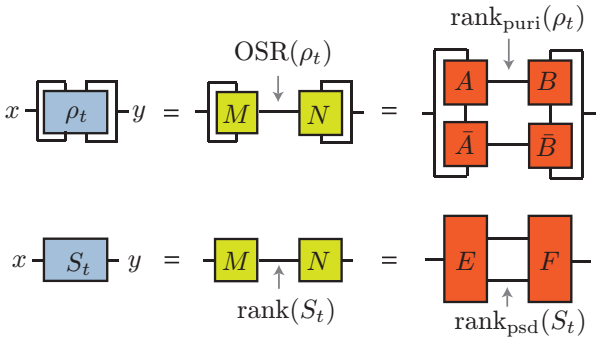
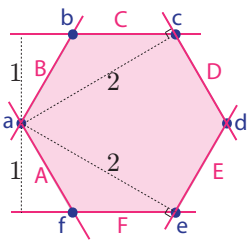


Figure 3. (Left) the classical bipartite state ρ_t (equation (10)) (with a joint ket and bra index for the party in state x and the party in state y), and below its associated non-negative matrix S_t. (Middle) the operator Schmidt decomposition of ρ_t (with dimension $\mathrm{OSR}(\rho_t)$) corresponds to the singular value decomposition of S_t (with dimension $\mathrm{rank}(S_t)$). (Right) the local purification form of ρ_t (with dimension $\mathrm{rank}_{\mathrm{puri}}(\rho_t)$) corresponds to the positive semidefinite factorization of S_t (with dimension $\mathrm{rank}_{\mathrm{psd}}(S_t)$).



$$S_6 = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \end{array} \begin{array}{c} \begin{array}{cccccc} A & B & C & D & E & F \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 2 & 2 & 1 \\ 0 & 1 & 2 & 2 & 1 & 0 \\ 1 & 2 & 2 & 1 & 0 & 0 \\ 2 & 2 & 1 & 0 & 0 & 1 \\ 2 & 1 & 0 & 0 & 1 & 2 \\ 1 & 0 & 0 & 1 & 2 & 2 \end{bmatrix} \end{array}$$

Figure 4. The slack matrix of the regular hexagon S_6 contains in its (i, j) entry the distance from vertex i (here labeled with lowercase letters) to hyperplane j (labeled with uppercase letters). Note that the matrix is circulant.

Using the equivalences (12), this implies that

$$\begin{aligned} &\mathrm{OSR}(\rho_t) = 3 \quad \forall t \\ &\mathrm{rank}_{\mathrm{puri}}(\rho_t) \sim \log t. \end{aligned} \tag{14}$$

That is, the operator Schmidt rank of ρ_t is constant for all t, whereas its purification rank grows unboundedly with t. It follows that there does not exist an upper bound of $\mathrm{rank}_{\mathrm{puri}}(\rho_t)$ that depends only on $\mathrm{OSR}(\rho_t)$. □

3.2. Inequivalence for multipartite classical states

Now we show a more general form of separation, namely we provide classical multipartite states with a constant operator Schmidt rank across every linear bipartition, and an unbounded purification rank.

Result 2. *Let ρ denote a mutipartite classical mixed state. Let $D = OSR(\rho)$ and $D' = \mathrm{rank}_{\mathrm{puri}}(\rho)$. Then D' cannot be upper bounded by a function f that depends only of D (in*

particular, f is independent of ρ), i.e.

$$\nexists f: \quad D' \leqslant f(D). \tag{15}$$

Proof. Consider the family of states ρ_t (equation (10)) with $t = 2^m$ (with natural m). That is, S_t is the slack matrix of the square, octagon, etc. Let the binary representations of the row index x and column index y be $(x_1, \ldots, x_m)$ and $(y_1, \ldots, y_m)$, respectively.

Since the polygon is regular, S_t is a circulant matrix. Hence, it is diagonalized by the Fourier transform F_t, with components $F_t(x, y) = \omega^{xy}$, where $\omega = \exp(\mathrm{i}2\pi/2^m)$. That is,

$$S_t(x, y) = \sum_{\alpha=1}^{3} L^{\alpha}_{x_1,\ldots,x_m} R^{\alpha}_{y_1,\ldots,y_m}, \tag{16}$$

where

$$\begin{aligned} L^{\alpha}_{x_1,\ldots,x_m} &= \exp\left[-2\pi\mathrm{i}\alpha\, 0.x_1 \ldots x_m\right]\lambda_\alpha, \\ R^{\alpha}_{y_1,\ldots,y_m} &= \exp\left[2\pi\mathrm{i}\alpha\, 0.y_1 \ldots y_m\right] \end{aligned} \tag{17}$$

and $\{\lambda_\alpha\}$ are the eigenvalues of S_t. Now we decompose the tensors L and R into smaller tensors,

$$\begin{aligned} L^{\alpha}_{x_1\ldots x_m} &= (M_1)^{\alpha}_{x_1} \ldots (M_m)^{\alpha,\alpha}_{x_m}, \\ R^{\alpha}_{y_1\ldots y_m} &= (M_{m+1})^{\alpha,\alpha}_{y_1} \ldots (M_{2m})^{\alpha}_{y_m}, \end{aligned} \tag{18}$$

where each M_k depends only on one bit (x_k or y_k),

$$\begin{aligned} (M_k)^{x_k}_{\alpha,\alpha} &= \exp\left[-2\pi\mathrm{i}\alpha\, \tfrac{x_k}{2^k}\right] \quad 1 < k < m, \\ (M_m)^{x_m}_{\alpha,\alpha} &= \exp\left[-2\pi\mathrm{i}\alpha\, \tfrac{x_m}{2^m}\right]\lambda_\alpha, \\ (M_k)^{y_{k-m}}_{\alpha,\alpha} &= \exp\left[2\pi\mathrm{i}\alpha\, \tfrac{y_{k-m}}{2^{k-m}}\right] \quad m < k < 2m. \end{aligned} \tag{19}$$

That is, each M_k is a 3×3 diagonal matrix, except for M_1 and M_{2m}, which are a row and a column vector, respectively, defined analogously. This shows that ρ_t has operator Schmidt rank 3 across every linear bipartition.

We know from equation (14) that the purification rank of ρ_t along the x versus y bipartition grows unboundedly like $\log t$. Clearly, a small purification rank across any bipartition would imply a small purification rank across the x versus y bipartition. Thus, the purification rank is unbounded at least across one bipartition (see figure 5). □

This shows that a small operator Schmidt rank across all linear bipartitions does not imply a small purification rank.

Let us make a final remark. Note that a possible purification of ρ_t is

$$\begin{aligned} \rho_t &= \mathrm{tr}_{\mathrm{anc}}|\varphi_t\rangle\langle\varphi_t|, \\ |\varphi_t\rangle &= \sum_{x,y=1}^{t} \sqrt{S_t(x, y)}|x, x, y, y\rangle, \end{aligned} \tag{20}$$

where the second and fourth index refer to the ancillary states. Thus, $\mathrm{rank}_{\mathrm{puri}}(\rho_t) \leqslant \mathrm{SR}(\varphi_t)$. Above we have seen that $\mathrm{rank}_{\mathrm{puri}}(\rho_t) \sim \log t$, thus the Schmidt rank of $|\varphi_t\rangle$ grows with t, and

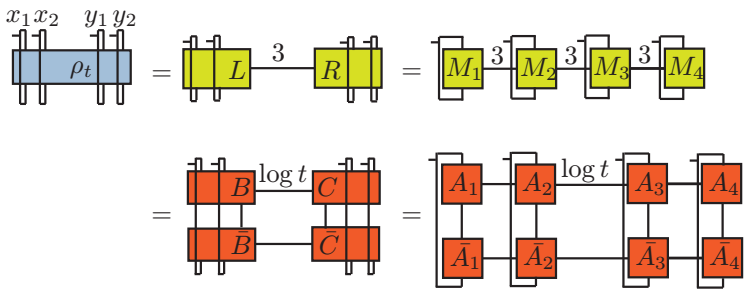


Figure 5. The classical state ρ_t (equation (10)), where S_t is the slack matrix of the regular t-gon, shows the separation of the MPDO form and the local purification form. (Middle) in the bipartite case, ρ_t has operator Schmidt rank 3 for all t, and purification rank $\sim \log t$. (Right) in the multipartite case, for $t = 2^m$, ρ_t has operator Schmidt rank 3 across every linear bipartition, and purification rank $\sim \log t$ at least across one bipartition.

so does its preparation cost[11]. Now consider the state obtained by taking the square of the coefficients of $|\varphi_t\rangle$,

$$|\varphi_t^{\odot 2}\rangle := \sum_{x,y=1}^{t} S_t(x, y)|x, x, y, y\rangle. \tag{21}$$

From result 2 it follows that the Schmidt rank of this state across any linear bipartition is 3. Hence its preparation cost is constant with t [18]. Thus, we see that transforming (21) to (20), i.e. taking the Hadamard square root of the coefficients, may have a high cost in the Schmidt rank, and thus in the preparation cost of a pure state.

4. Purification methods

We will now present two constructive purification methods: the sos polynomial method (section 4.1) and the eigenbasis method (section 4.2). Both are applicable to all multipartite mixed states, and can be used to construct exact and approximate purifications. We will compare both approximation methods and see that they are complementary for various eigenvalue distributions in section 4.3.

4.1. sos polynomial method

We will first present the idea of the sos polynomial method (section 4.1.1), and then explain how to use it to construct exact (section 4.1.2) and approximate purifications (section 4.1.3).

4.1.1. The idea. The idea of the sos polynomial method is the following: given a mixed state ρ, we construct a purifying state as a sum of powers of ρ (up to certain degree), where each power is attached to an ancillary state. If the degree is large enough, there exists a choice of the ancillary states such that this purifying state is an exact purification for ρ. If the degree is not large enough, one can find the ancillary states with an ansatz of sos polynomials or with an SDP.

[11] In the sense of the dimension of the ancilla required to prepare the state in a sequential scheme [18].

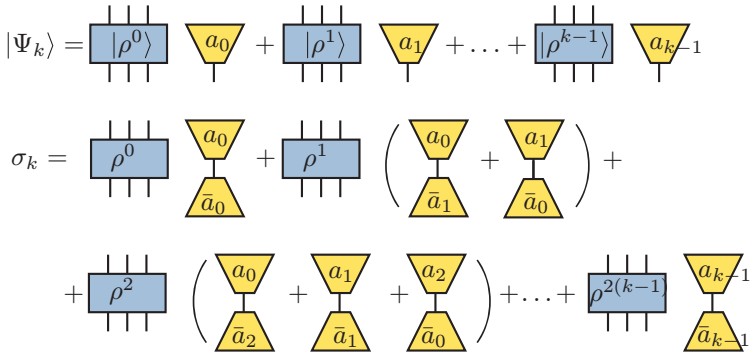


Figure 6. The sos polynomial method constructs a state $|\Psi_k\rangle$ as a sum of powers of ρ, from 0 to $k-1$, each attached to an ancillary state, $|a_l\rangle$. $|\Psi_k\rangle$ is used as purification of σ_k, which has the same eigenvectors as ρ, but its eigenvalues are a sos polynomial of the eigenvalues of ρ. If $k = m$, where m is the number of different non-negative eigenvalues of ρ, σ_k can be an exact purification of ρ. If $k < m$, one can search for the σ_k closest to ρ with an SDP.

Specifically, consider a multipartite density matrix ρ of size $d^N \times d^N$, with spectral decomposition

$$\rho = \sum_{i=1}^{n} \lambda_i |\phi_i\rangle\langle\phi_i| + \sum_{i=n+1}^{d^N} \lambda_i |\phi_i\rangle\langle\phi_i|, \tag{22}$$

where $\lambda_i = 0$ for $i > n$. We construct a purifying state $|\Psi_k\rangle$ as a sum of powers of ρ, from 0 to $k-1$, where each power is attached to an ancillary state (see figure 6),

$$|\Psi_k\rangle = \sum_{l=0}^{k-1} |\rho^l\rangle_{KB} \otimes |a_l\rangle_A, \tag{23}$$

where $|\rho\rangle$ denotes a vectorized matrix ρ (i.e. given $\rho = \sum_{i,j} \rho_{i,j}|i\rangle\langle j|$ we define $|\rho\rangle_{KB} = \sum_{i,j} \rho_{i,j}|i\rangle_K|j\rangle_B$). We use $|\Psi_k\rangle$ as the purifying state of a density matrix σ_k,

$$\sigma_k = \mathrm{Tr}_{BA}|\Psi_k\rangle\langle\Psi_k| = \sum_{i=1}^{d^N} p_k(\lambda_i)\, |\phi_i\rangle\langle\phi_i|. \tag{24}$$

Here p_k is a polynomial that can be written as

$$p_k(\lambda) = (1, \lambda, \ldots, \lambda^{k-1})\, R_k\, (1, \lambda, \ldots, \lambda^{k-1})^T, \tag{25}$$

where R_k is a positive semidefinite matrix which is the Gram matrix of the ancillary states, i.e. its (i, j) component is $R_k(i, j) = \langle a_i|a_j\rangle$. This polynomial is sos, as can be readily seen by writing $R_k = A^T A$,

$$p_k(\lambda) = \sum_{u=1}^{r} y_u(\lambda)^2 \tag{26}$$

$$y_u(\lambda) = A_u(1, \lambda, \ldots, \lambda_{k-1})^T, \tag{27}$$

where A_u denotes the uth row of A, and $r = \mathrm{rank}(R_k)$. Note that, since the polynomial is univariate, the set of sos polynomials is identical to the set of non-negative polynomials (i.e. with the property $p(x) \geqslant 0$ for all x) [19].

Now, by construction, the purification rank of σ_k is at most the Schmidt rank of $|\Psi_k\rangle$ (equation (24)). Denoting $\mathrm{OSR}(\rho) = D$, this is at most $1 + D + D^2 + \cdots + D^{k-1}$ (equation (23)), and thus

$$\mathrm{rank}_{\mathrm{puri}}(\sigma_k) \leqslant \mathrm{SR}(\Psi_k) \leqslant \frac{D^k - 1}{D - 1}. \tag{28}$$

We are interested in making σ_k as close as possible to ρ. We will show that if $k = m$, where m is the number of different non-negative eigenvalues of ρ,[12] there exists ancillary states such that $p_k(\lambda_i) = \lambda_i$ for all i, and thus $\sigma_m = \rho$, i.e. σ_m is an exact purification of ρ (section 4.1.2). If, on the contrary, $k < m$, one can choose $k - 1$ points and construct the p_k that passes through them, or one can find the p_k that minimizes the distance between σ_k and ρ (in trace norm) with an SDP (section 4.1.3).

4.1.2. Exact case. We now show how to build exact purifications with the sos polynomial method.

Result 3. *Let ρ denote a multipartite density matrix with m different non-negative eigenvalues. Let $D = OSR(\rho)$ and $D' = \mathrm{rank}_{\mathrm{puri}}(\rho)$. The sos polynomial method provides an exact purification of ρ with*

$$D' \leqslant \frac{D^m - 1}{D - 1}. \tag{29}$$

Proof. To ease the notation we consider that the different non-negative eigenvalues are the first m (i.e. $\lambda_i \neq \lambda_j$ for $i \neq j$ and $i, j = 1, \ldots, m$). Consider the construction of the purifying state σ_k as explained in section 4.1.1. Define the vector $|v_k^i\rangle$ as

$$|v_k^i\rangle = (1, \lambda_i, \ldots, \lambda_i^{k-1})^{\mathrm{T}} \tag{30}$$

so that the sos polynomial evaluated at λ_i can be written as $p_k(\lambda_i) = \langle v_k^i|R_k|v_k^i\rangle$ (compare with equation (25)). Now we choose $k = m$. The important observation is that the set of vectors $\{|v_m^i\rangle\}_{i=1}^m$ is linearly independent (since only different eigenvalues are considered). Hence, there exists another set $\{|w_m^j\rangle\}_{j=1}^m$ which is biorthogonal to it, i.e. $\langle v_m^i|w_m^j\rangle = \delta_{ij}$ for all i, j. Then we choose R_m as follows:

$$R_m = \sum_{j=1}^{m} \lambda_j |w_m^j\rangle\langle w_m^j|. \tag{31}$$

This satisfies that $p_m(\lambda_i) = \langle v_m^i|R_m|v_m^i\rangle = \lambda_i$ for all i. From equation (24) it follows that $\sigma_m = \rho$ and thus σ_m is an exact purification of ρ. Finally, using equation (28) with $k = m$, the claim of the result follows. □

Note that result 3 depends on the number of different non-negative eigenvalues m because $\sigma_k = \rho$ requires that $p(\lambda_i) = \lambda_i$ for all i, and these are only m independent conditions.

[12] By non-negative eigenvalues, we mean that the 0 should also be counted as an eigenvalue. That is, if ρ is rank deficient, m is the number of different non-zero eigenvalues plus one.

4.1.3. Approximate case: SDP. Building approximate purifications with the sos polynomial method can be done analytically or numerically, as we show next.

Result 4. *Let ρ denote a multipartite density matrix of dimension $d^N \times d^N$, with* $\mathrm{rank}(\rho) = n$. *Let $D = OSR(\rho)$. The sos polynomial method provides a density matrix σ_k with*

$$\mathrm{rank}_{\mathrm{puri}}(\sigma_k) \leqslant \frac{D^k - 1}{D - 1} \tag{32}$$

and

$$\|\rho - \sigma_k\|_1 = \sum_{i=1}^{n} |\lambda_i - p_k(\lambda_i)| + (d^N - n)\, p_k(0), \tag{33}$$

where p_k is a sos polynomial of degree $2(k-1)$. p_k can be constructed by choosing $k-1$ non-negative points through which it must pass. Alternatively, the optimal p_k can be found with an SDP.

Proof. *sos polynomial that passes through $k-1$ points.* We construct the sos polynomial p_k by letting it pass through $k-1$ chosen points $\{\mu_1, \ldots, \mu_{k-1}\}$. We use the Lagrange basis 'squared' to this end

$$p_k(\lambda) = \sum_{j=1}^{k-1} \mu_j l_j(\lambda), \tag{34}$$

$$l_j(\lambda) = \prod_{i=1, i\neq j}^{k-1} \left(\frac{\lambda - \mu_i}{\mu_j - \mu_i}\right)^2, \tag{35}$$

where we have omitted the dependence of l_j on k. This satisfies $p_k(\mu_j) = \mu_j$ for $j = 1, \ldots, k-1$. Note that the degree of p_k is $2(k-1)$ as required. The distance (33) depends on the points $\{\mu_i\}$, thus the difficulty lies in choosing them.

The optimization problem as an SDP. Let $\{\lambda_i\}_{i=1}^{d^N}$ denote the eigenvalues of ρ, with $\lambda_i = 0$ for $n < i \leqslant d^N$. We search for the sos polynomial p_k that minimizes the distance (33), i.e. for the positive semidefinite matrix R_k that minimizes it (see equation (25)). The optimization problem thus reads

$$\begin{aligned} &\min \sum_{i=1}^{d^N} |\lambda_i - \langle v_k^i | R_k | v_k^i \rangle| \\ &\text{s.t. } R_k \succeq 0. \end{aligned} \tag{36}$$

The objective function can be made linear by introducing the slack variables z

$$\begin{aligned} &\min \sum_{i=1}^{d^N} z_i \\ &\text{s.t. } z_i \geqslant \lambda_i - \langle v_k^i | R_k | v_k^i \rangle \qquad i = 1, \ldots, d^N, \\ &\quad z_i \geqslant -\lambda_i + \langle v_k^i | R_k | v_k^i \rangle \quad i = 1, \ldots, d^N, \\ &\quad R_k \succeq 0. \end{aligned} \tag{37}$$

The optimization variables are now $R_k \succeq 0$ and $z \geqslant 0$, and the constraints are linear in them. Thus, this is an SDP optimization problem (see appendix A for the precise formulation). □

In words, the SDP searches for the sos polynomial p_k (of degree $2(k-1)$) whose distance to the eigenvalue distribution is minimal in trace norm. This formulation is consistent with the fact that optimization over sos polynomials can be done with SDPs [19].

Note that the non-trivial condition is that the polynomial be sos (equivalently, non-negative), since otherwise one could take $p(\lambda) = \lambda$. Approximating λ for $0 \leqslant \lambda \leqslant \lambda_{\max}$ and another non-negative function elsewhere has the problem that the function at $\lambda = 0$ is non-analytical.

We remark that there may exist exact solutions for $m/2 < k < m$, since $p_k(\lambda) - \lambda$ has degree $2(k-1)$, and thus can have $2(k-1)$ real roots. However, we only know how to construct the exact solution for the case $k = m$ (with the idea of the proof of result 3).

Remark 1. *Non-orthogonal ancillary states. Observe that the exact solution for $k = m$ involves non-orthgonal ancillary states, since R_m is non-diagonal (equation (31)).*

This is so because the basis $\{|v_m^i\rangle\}_{i=1}^m$ is not orthogonal, and neither is the biorthogonal basis $\{|\omega_m^j\rangle\}_{j=1}^m$. More generally, the solution for $k < m$ also involves non-orthogonal ancillary states, since orthogonal states result in a diagonal R_k, which renders a polynomial with only even powers with non-negative coefficients (i.e. a monotonously increasing polynomial for positive λ). In contrast, a non-diagonal R yields polynomials with possibly odd powers with negative coefficients, thus with various minima, rendering a better approximation of λ in the desired interval.

Remark 2. *Real ancillary vectors. The ancillary vectors $\{|a_j\rangle\}$ can be taken real without loss of generality. To see this, first write the polynomial (equation (25)) as*

$$p_k(\lambda) = \sum_{s,t=0}^{k-1} R_k(s,t)\lambda^{s+t} = \sum_{l=0}^{2(k-1)} c_l \lambda^l, \tag{38}$$

where $c_l = \sum_{s+t=l} R_k(s,t)$.

That is, each coefficient c_l is the sum of the lth antidiagonal of R_k. Thus c_l depends on diagonal elements of R_k, which are non-negative, or on the sum of an element and its transposed $(R_k(s,t) + R_k(t,s))$, which is real. Therefore c_l only depends on the real part of R. If $R \succeq 0$, then $\mathrm{Re}(R)$ is positive semidefinite for all real vectors $|r\rangle$, since $\langle r|\mathrm{Re}(R) + \mathrm{i}\mathrm{Im}(R)|r\rangle \geqslant 0$, which implies $\mathrm{i}\langle r|\mathrm{Im}(R)|r\rangle = 0$. Since we only consider contractions with real vectors (namely $|v_k^i\rangle$), we can restrict R to its real part. Thus, its spectral decomposition reads $R = ODO^{\mathrm{T}}$, where O is orthogonal, and D is diagonal and nonnegative. This readily yields $R = O\sqrt{D}\sqrt{D}O^{\mathrm{T}} = AA^{\mathrm{T}}$, where the ith row of A contains the coefficients of the ancillary state $\langle a_i|$ expressed in the eigenbasis of R, which are real.

While the exact sos polynomial method depends on m (result 3) and the approximate on d^N (result 4),we will see in section 4.3 that in practice there exists good ansätze of sos polynomials which make it independent of both. We will also compare this approximate method with the eigenbasis method, which we present next.

4.2. Eigenbasis method

Now we turn to the eigenbasis method, for which we present the main idea (section 4.2.1), and how to use it to construct exact (section 4.2.2) and approximate purifications (section 4.2.3).

4.2.1. The idea. The idea of the eigenbasis method is to consider the standard purification of ρ, which is the spectral decomposition, and upper bound the Schmidt rank of each eigenstate of ρ. This is done by constructing a basis of the range of ρ where each basis element is the image (under the map ρ) of a product state. Thus, each has Schmidt rank at most D, where D is the operator Schmidt rank of ρ. Expressing each eigenstate in terms of this basis, we see that each can have Schmidt rank at most Dn, where n is the number of non-zero eigenvalues of ρ. Thus the purification rank of ρ is at most Dn^2.

To be more precise, consider the standard purification of ρ, obtained from its spectral decomposition (see equation (22)) as

$$\rho = \mathrm{Tr}_A |\Psi\rangle\langle\Psi| \tag{39}$$

$$|\Psi\rangle = \sum_{i=1}^{n} \sqrt{\lambda_i}\, |\phi_i\rangle_S |i\rangle_A. \tag{40}$$

By construction, the purification rank of ρ is at most the Schmidt rank of $|\Psi\rangle$, which is at most n times the maximum Schmidt rank of $|\phi_i\rangle$,

$$\mathrm{rank}_{\mathrm{puri}}(\rho) \leqslant \mathrm{SR}(\Psi) \leqslant n \max_i \mathrm{SR}(\phi_i). \tag{41}$$

Our goal is to upper bound $\max_i \mathrm{SR}(\phi_i)$ as a function of D. To this end, we build a basis of the range of ρ where each basis element $|\chi_\alpha\rangle$ is the image (under the map ρ) of a certain product state $|p_\alpha\rangle$ (see figure 7),

$$|\chi_\alpha\rangle = \rho\, |p_\alpha\rangle, \tag{42}$$

$$|p_\alpha\rangle = \bigotimes_{i=1}^{N} |p_{\alpha_i}\rangle_i, \tag{43}$$

where $\alpha = (\alpha_1, \ldots, \alpha_N)$, α_i is a label of the ith product state, and the subindex outside the ket denotes the subsystem that the state is describing.

Now, consider the MPDO form of ρ as in equation (2), with operator Schmidt rank D. The Schmidt rank of $|p_\alpha\rangle$ is one, and applying ρ to $|p_\alpha\rangle$ increases the Schmidt rank by at most D. It follows that

$$\mathrm{SR}(\chi_\alpha) \leqslant D. \tag{44}$$

On the other hand, we consider the spectral decomposition of ρ (equation (22)), and use the eigenstates as a basis of the range. In particular, we express $|\chi_\alpha\rangle$ in terms of this basis, with coefficients $f_{i\alpha}$ of the linear combination,

$$|\chi_\alpha\rangle = \sum_{i=1}^{n} |\phi_i\rangle \underbrace{\lambda_i \langle\phi_i|p_\alpha\rangle}_{f_{i\alpha}}. \tag{45}$$

Since the range has dimension n, there are at most n linearly independent $|\chi_\alpha\rangle$. The idea of the exact result (section 4.2.2) is to show that one can invert the relation (45) and express $|\phi_i\rangle$ as a linear combination of (at most n) $|\chi_\alpha\rangle$. The approximate method discards the smallest eigenvalues, and applies the exact result to the truncated density matrix (section 4.2.3).

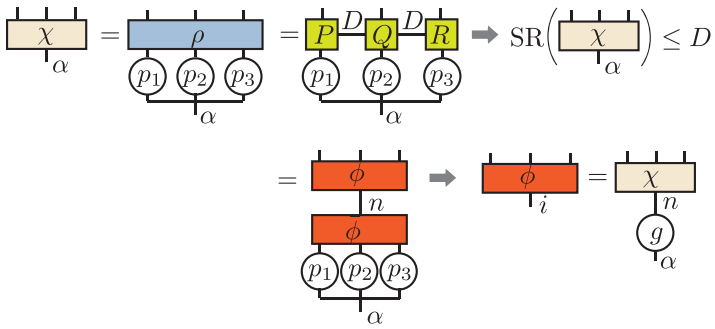


Figure 7. The eigenbasis method constructs a state $|\chi_\alpha\rangle$ as the image (under ρ) of the product state $\otimes_{i=1}^{N}|p_i\rangle_\alpha$, here shown for $N=3$. (Upper line) using the operator Schmidt decomposition of ρ we see that $\mathrm{SR}(\chi_\alpha)\leqslant D$. (Lower line) using the spectral decomposition of ρ, we express each eigenstate $|\phi_i\rangle$ as a linear combination of (at most n) $|\chi_\alpha\rangle$ (equation (47)), and upper bound the Schmidt rank of each eigenstate by Dn.

4.2.2. Exact case. We now show how to use the eigenbasis method to construct an exact purification.

Result 5. *Let ρ denote a multipartite density matrix with* $\mathrm{rank}(\rho)=n$. *Let $D=OSR(\rho)$ and $D'=\mathrm{rank}_{\mathrm{puri}}(\rho)$. Then the eigenbasis method constructs an exact purification of ρ with*

$$D'\leqslant Dn^2. \tag{46}$$

Proof. Let ρ be a multipartite density matrix with spectral decomposition as in (22) (thus with $\mathrm{rank}(\rho)=n$). Building upon section 4.2.1, we only need to show that there exist n product states $\{|p_\alpha\rangle\}_{\alpha=1}^{n}$ such that their images under ρ, $\{|\chi_\alpha\rangle=\rho|p_\alpha\rangle\}_{\alpha=1}^{n}$, form a basis of the range of ρ. Take a product basis $\{|x\rangle\}$ with $x=1,\ldots,d^N$, which spans the whole space. Then $\{\rho|x\rangle\}$ spans the range of ρ. Then, we can select a basis $\rho|x_s\rangle$ of the range of ρ with $s=1,\ldots,n$. We call the latter product states $|p_\alpha\rangle$ with $\alpha=1,\ldots,n$. Then for every $|\psi\rangle$ in the support of ρ there exist $\{c_\alpha\}$ such that $|\psi\rangle=\sum_{\alpha=1}^{n}c_\alpha|p_\alpha\rangle$. Now for every state $|\phi\rangle$ in the range of ρ there exists $|\psi\rangle$ in the support of ρ such that $|\phi\rangle=\rho|\psi\rangle=\sum_{\alpha=1}^{n}c_\alpha\rho|p_\alpha\rangle$, where we have used the linearity of ρ. Thus, $\{\rho|p_\alpha\rangle\}_{\alpha=1}^{n}$ forms a basis of the range of ρ.

This implies that the coefficient matrix $(f_{i\alpha})$ (equation (45)) is full-rank and hence can be inverted. We denote the elements of the inverse by $g_{\alpha j}$, i.e. $\sum_{\alpha=1}^{n}f_{i\alpha}g_{\alpha j}=\delta_{ij}$, and we invert equation (45) to express the eigenvector $|\phi_i\rangle$ as a linear combination of the vectors $|\chi_\alpha\rangle$,

$$|\phi_i\rangle=\sum_{\alpha=1}^{n}g_{\alpha i}|\chi_\alpha\rangle. \tag{47}$$

Since $\mathrm{SR}(\chi_\alpha)\leqslant D$, we obtain

$$\mathrm{SR}(\phi_i)\leqslant Dn, \tag{48}$$

for all i. Finally, using (41), the claim of the result follows. □

4.2.3. Approximate case. We now show how to use the eigenbasis method to build approximate purifications.

Result 6. *Let ρ denote a multipartite density matrix with spectral decomposition as in (22), and let $D = OSR(\rho)$. The eigenbasis method provides a density matrix σ_s with* $\mathrm{rank}(\sigma_s) = s$ *such that*

$$\mathrm{rank}_{\mathrm{puri}}(\sigma_s) \leqslant Ds^2, \tag{49}$$

and its distance to ρ is

$$\|\rho - \sigma_s\|_1 \leqslant 2 \sum_{i=s+1}^{n} \lambda_i. \tag{50}$$

Proof. We construct σ_s with the largest s eigenvalues of ρ, i.e.

$$\sigma_s = \frac{1}{\mathcal{N}} \sum_{i=1}^{s} \lambda_i |\phi_i\rangle\langle\phi_i|, \tag{51}$$

where $\mathcal{N} = \sum_{i=1}^{s} \lambda_i$. By direct calculation one can see that (50) holds with equality. Applying result 5 to σ_s yields equation (49). □

Clearly, this method yields good approximations for rapidly decaying distributions of eigenvalues, for which the distance (50) is small. In the next section we make these statements precise, and compare this method to the sos polynomial method.

4.3. Comparison of approximation methods

We now compare the two approximation methods for a state ρ with $\mathrm{rank}(\rho) = n$ with the following eigenvalue distributions (where the eigenvalues are ordered in non-increasing magnitude).

(i) Uniform distribution, defined as $\lambda_j = 1/n$ for all j.
(ii) Equally spaced distribution, defined as $\lambda_j = j\frac{2}{n(n+1)}$ for $j = 1, \ldots, n$.
(ii) Random distribution, defined as $\lambda_j = \mathcal{N} b_j$, where b_j is a random number in a fixed interval and $\mathcal{N} = 1/\sum_{j=1}^{n} b_j$.
(iv) One fixed eigenvalue and the rest equally spaced, defined as $\lambda_1 = 1/2$ and $\{\lambda_j = j\mathcal{N}\}_{j=2}^{n}$ where $\mathcal{N} = 1/(n(n+1) - 2)$.
(v) Exponentially decaying distribution, defined as $\lambda_j = \mathcal{N}\exp(-bj)$ where $\mathcal{N} = (1 - e)/(e^{-n} - 1)$.

For each distribution, each method provides a matrix σ at distance ϵ from ρ,

$$\|\rho - \sigma\|_1 \leqslant \epsilon, \tag{52}$$

such that

$$\mathrm{rank}_{\mathrm{puri}}(\sigma) \leqslant f(D, \epsilon, n), \tag{53}$$

where $D = \mathrm{OSR}(\rho)$. Our goal is to determine $f(D, \epsilon, n)$ with the sos polynomial method (section 4.3.1) and the eigenbasis method (section 4.3.2).

4.3.1. sos polynomial method. We start with the sos polynomial method, for which we first present analytical and then numerical results obtained with the SDP, both based on result 4.

Firstly, the uniform distribution is the easiest for this method. Using the exact result (result 3) with $m = 1$ we obtain $D' = 1$, since this distribution describes the maximally mixed state. The sos polynomial that constructs this exact purification is $p_1(\lambda) = 1/n$.

Secondly, the sos polynomial that best approximates the line λ in the interval $[0, \lambda_1]$ is a good ansatz for the equally spaced and random distribution. Moreover, these polynomials can be chosen so that their distance to the distribution is independent of n, as we show next.

Result 7. *Consider a distribution of n eigenvalues whose largest eigenvalue $\lambda_1 \sim 1/n$, such as the equally spaced or the random distribution. Then there exists a sos polynomial p_k such that $\sum_{i=1}^{n} |\lambda_i - p_k(\lambda_i)|$ is independent of n.*

Proof. Let p_k denote the sos polynomial of degree $2(k-1)$ that best approximates the straight line λ in the interval $[0, 1]$. This can be rescaled to approximate the straight line in $[0, \lambda_1]$,

$$p'_k(\lambda) = \lambda_1 \; p_k(\lambda/\lambda_1). \tag{54}$$

Now, let $\varepsilon = \max_{\lambda\in[0,1]} |p_k(\lambda) - \lambda|$. Then $\max_{\lambda\in[0,\lambda_1]} |p'_k(\lambda) - \lambda| = \lambda_1\, \varepsilon$, which implies that

$$\sum_{i=1}^{n} |p'_k(\lambda_i) - \lambda_i| \leqslant n\lambda_1\, \varepsilon. \tag{55}$$

If $\lambda_1 \sim 1/n$, this upper bound is independent of n. □

For the distribution with one fixed eigenvalue and the rest equally spaced, it holds that $\lambda_2 \sim 1/n$, hence the sos polynomial $q_k(\lambda) = p_{k-1}(\lambda)(\lambda - \lambda_1)^2$, where p_{k-1} is defined in the proof of result 7, also has a distance independent of n. Denoting the distance $\|\rho - \sigma_k\|_1 = \epsilon = g(k)$, and using result 4, this implies that one can upper bound the purification rank of σ of these three distributions by a function independent of n, namely $\mathrm{rank}_{\mathrm{puri}}(\sigma) \leqslant O(D^{g^{-1}(\epsilon)-1})$.

Finally, for the exponentially decaying distribution, we present an ansatz of sos polynomials whose distance decreases exponentially with k and is independent of n.

Result 8. *Consider the exponentially decaying distribution of eigenvalues $\lambda_j = a\mathrm{e}^{-bj}$ (with $a, b > 0$) for $j = 1, \ldots, n$, and the sos polynomial that passes through the largest $k-2$ eigenvalues and 0, constructed as*

$$p_k(\lambda) = \lambda^2 \sum_{r=1}^{k-2} \frac{1}{\lambda_r} l_r(\lambda), \tag{56}$$

$$l_r(\lambda) = \prod_{j=1, j\neq r}^{k-2} \left(\frac{\lambda - \lambda_j}{\lambda_r - \lambda_j}\right)^2. \tag{57}$$

The distance from this polynomial to the distribution decays exponentially in k, i.e.

$$\sum_{i=1}^{n} |\lambda_i - p_k(\lambda_i)| \leqslant O(\mathrm{e}^{-k}). \tag{58}$$

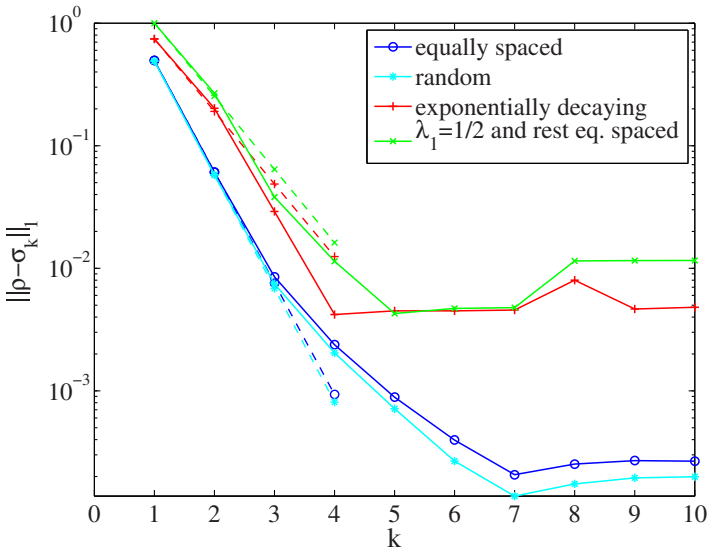


Figure 8. The distance $\|\rho - \sigma_k\|_1$ versus k for the distributions (ii) to (v) with $n = 100$ as found by the SDP (solid lines) and the exponential fits (dashed lines). The figure shows an exponential decrease in the distance for small values of k, and the program does not reduce the distance further due to numerical errors.

Proof. See appendix B. □

The exponential decrease in the distance

$$\|\rho - \sigma_k\|_1 \leqslant A \exp(-Bk) \tag{59}$$

implies that the purification rank of σ scales polynomially in D

$$\mathrm{rank}_{\mathrm{puri}}(\sigma) \leqslant \frac{D^{\ln(A/\epsilon)/B} - 1}{D - 1} = O(D^{\ln(1/\epsilon)-1}). \tag{60}$$

For the exponentially decaying distribution, we provide A and show that $B = b$ in appendix B.

Let us now analyze how well these eigenvalue distributions are approximated with the SDP. For distributions (ii) to (v), the SDP gives an exponential decrease of the distance in k and independent of n (equation (59)) with $A \approx 4$ and $B \approx 2$ for distributions (ii), (iii), $A \approx 3$ and $B \approx 1.3$ for (iv) and for (v) with $b = 1$ we find $A \approx 4$ and $B \approx 1.3$ (see figures 8 and 9). The sos polynomials found by the SDP for the equally spaced distribution are shown in figure 10. We remark that the numerics does not improve beyond $k \approx 4$, as very small numbers (such as powers of small eigenvalues) are numerically treated as 0. We have rescaled the eigenvalues $\lambda_i \to \lambda_i/\lambda_1$ for all i, which allows us to use the SDP until $k \approx 7$ for some distributions.

4.3.2. Eigenbasis method. We now use the eigenbasis method (result 6) to upper bound the purification rank of σ (equation (53)) for the eigenvalue distributions (i)–(v).

(i) Uniform distribution. This is the hardest distribution, since the smallest eigenvalues are as large as possible. We obtain

$$\mathrm{rank}_{\mathrm{puri}}(\sigma) \leqslant Dn^2(1 - \epsilon/2)^2. \tag{61}$$

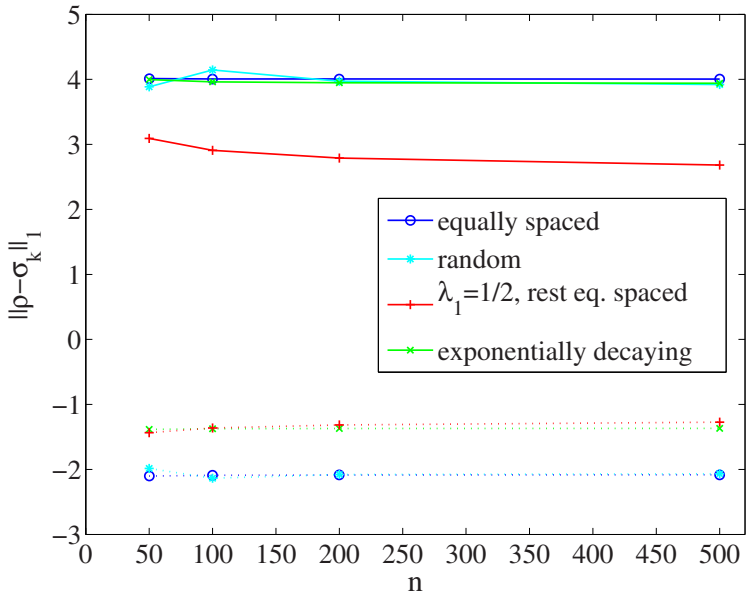


Figure 9. The factors A (solid line) and B (dashed line) (see equation (59)) obtained with the fits of the results of the SDP versus n. This shows that A and B are independent of n.

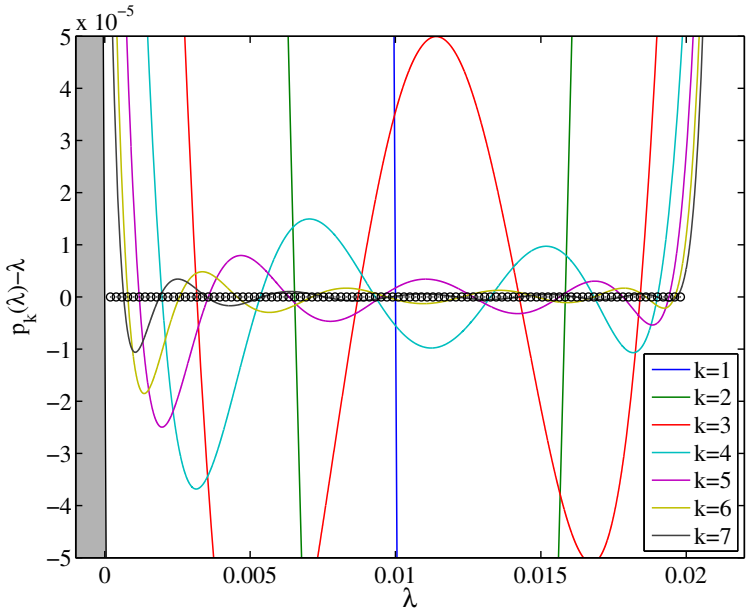


Figure 10. The sos polynomial found by the SDP minus λ, $p_k(\lambda) - \lambda$, for various k, for the equally spaced distribution of eigenvalues with $n = 100$. The black, open circles are the eigenvalues. The larger k is, the smaller the distance from the polynomial to the eigenvalues. The shaded area is that for which $p_k(\lambda) - \lambda \leqslant -\lambda$ i.e. it is unattainable for the sos polynomial.

(ii) Equally spaced distribution. We obtain

$$\begin{aligned}\mathrm{rank}_{\mathrm{puri}}(\sigma) &\leqslant \frac{D}{4}\left(\sqrt{1+4n(n+1)(1-\epsilon/2)}-1\right)^2\\ &\approx Dn^2(1-\epsilon/2).\end{aligned} \tag{62}$$

(iii) Random distribution. The distance depends on the particular random distribution; in the worst case, it is the uniform distribution, hence it is upper bounded by equation (61).

(iv) One fixed eigenvalue and the rest equally spaced. We obtain

$$\begin{aligned}\mathrm{rank}_{\mathrm{puri}}(\sigma) &\leqslant \frac{D}{4}\left(\sqrt{1+4n(n+1)(1-\epsilon)+8\epsilon}-1\right)^2\\ &\approx Dn^2(1-\epsilon).\end{aligned} \tag{63}$$

(v) Exponentially decaying distribution. Assuming that $e^{-n} \ll \epsilon$, the purification rank of σ grows linearly in D and quadratically in $\ln(1/\epsilon)$, i.e.

$$\begin{aligned}\mathrm{rank}_{\mathrm{puri}}(\sigma) &\leqslant \frac{D}{b^2}(\ln(2/\epsilon))^2\\ &= O(D\ln(1/\epsilon)^2).\end{aligned} \tag{64}$$

In summary, for the uniform distribution, the sos polynomial is the best, as it shows exactly that $D' = 1$. For the equally spaced, random, and one large eigenvalue and the rest equally spaced distributions, the sos polynomial method is better as it is independent of n and scales polynomially in D. Finally, for the exponentially decaying distribution, the eigenbasis method is better, as it scales linearly in D (and quadratically in $\ln(1/\epsilon)$). We thus see that the sos polynomial method is very robust, as it yields the same scaling for very different eigenvalue distributions, while the eigenbasis method works well only for rapidly decaying distributions of eigenvalues, in which case it works better than the other.

5. Conclusions and outlook

In this paper we have analyzed the efficiency of representing a mixed state as an MPDO and as a local purification, and we have shown that the latter can be arbitrarily more costly than the former. In particular, we have provided a family of multipartite classical states with a constant operator Schmidt rank D across each linear bipartition and an unbounded purification rank D' (result 2). This shows that, in the exact case, one cannot upper bound D' by a function of D only.

Then we have presented two constructive purifications methods which are applicable to any multipartite density matrix. The exact sos polynomial method implies that $D' \leqslant O(D^{m-1})$, where m is the number of different eigenvalues. Its approximate version consists of finding the sos polynomial which passes through certain points, and the optimal one can be found with an SDP (result 4). For the four tested eigenvalue distributions, this method upper bounds D' by a polynomial function of D which is independent of n, thus showing a robust and efficient behavior.

The exact eigenbasis method implies that $D' \leqslant Dn^2$, where n is the number of eigenvalues. Its approximate version discards the smallest eigenvalues (result 6), and for the exponentially decaying distribution it yields D' scaling linearly in D (and independent of n).

Our inequivalence result (result 1) implies that a single canonical form which is both efficient and locally positive semidefinite cannot exist in the exact case. Note that this is also a numerical limitation, as contracting 2D projected entangled pair states (PEPS) requires determining the 1D density matrices which are at the boundary [20]. In order to have an efficient algorithm for 2D PEPS it is thus necessary to use an efficient description of such 1D mixed states, hence to work with the MPDO form. On the other hand, for the algorithm to be stable, it is necessary that the positivity of the operator can be checked efficiently, which means for example, that it can be checked locally. While in practice one can work with MPDOs of low dimension, one would hope to reexpress such MPDOs in terms of a purification fulfilling the latter requirement, but without increasing significantly the bond dimension. Our results show that this may not be possible. At the same time, the results of section 4.3 show that one can construct efficient approximate purifications of various relevant eigenvalue distributions. One should nonetheless analyze how successive truncations of this approximate purification affect the total error.

Let us now mention some open questions. Concerning the inequivalence of the two forms, we believe that a larger separation could be obtained with states of the form $\rho = I \otimes I - P_1 \otimes Q_1 - P_2 \otimes Q_2$ where P_i, Q_i are Hermitian operators constrained by the fact $\rho \succeq 0$, but otherwise with random entries. Concerning the sos polynomial method, it would be interesting to find sos polynomials whose distance to the equally spaced distribution decreases exponentially with the degree k. It would also be appealing to combine both purification methods in a single one with the best of each, but this requires to split the density matrix into different 'eigenspace sectors' (such as in $\rho = \rho_1 + \rho_2$ where $\rho_{1,2} = \sum_{i \leqslant r, i > r} \lambda_i |\phi_i\rangle\langle\phi_i|$), and it is not clear how the operator Schmidt rank of ρ relates to that of ρ_1 or ρ_2. On a more practical level, it would be worth devising a purification method that works sequentially and does not require to know the spectral decomposition of the density matrix.

Our results also have connections to other research areas. In particular, the results on decomposability of mixed states translate one-to-one to divisibility properties of completely positive (CP) maps via the Choi–Jamiołkowski isomorphism. While the divisibility of CP maps with a fixed dimension of the intermediate space has been studied e.g. in [21, 22], our approach would allow to extend this study to varying middle dimension. Finally, our investigations are also related to communication complexity, since the positive semidefinite rank determines the quantum communication complexity [13] and the quantum correlation complexity [14] of the associated matrix.

Acknowledgments

We thank M C Bañuls, F Pastawski, M Piani for discussions. GDLC and NS acknowledge support from the Alexander von Humboldt foundation. DPG acknowledges support from Spanish grants MTM2011-26912 and QUITEMAD, and European CHIST-ERA project CQC. JIC acknowledges support from the EU Integrated Project SIQS.

Appendix A. Semidefinite program

Here we give an exact formulation of optimization problem of (36) as an SDP. As mentioned in the text, we first make the objective function linear by introducing slack variables $\{z_i \geqslant 0\}$

which upper bound the absolute values, as in equation (37). Then we rewrite the constraints so that the variables (z, R) appear on one side of the inequality and the data λ on the other,

$$\begin{aligned} &\min \sum_{i=1}^{d^N} z_i \\ &\text{s.t. } -z_i - \langle v_k^i | R_k | v_k^i \rangle \leqslant -\lambda_i \quad i = 1, \ldots, d^N, \\ &\quad -z_i + \langle v_k^i | R_k | v_k^i \rangle \leqslant \lambda_i \qquad i = 1, \ldots, d^N, \\ &\quad R_k \succeq 0. \end{aligned} \tag{A.1}$$

It is now straightforward to verify that this is an SDP optimization problem, whose standard formulation is given by

$$\begin{aligned} &\min_X \langle C, X \rangle \\ &\text{s.t. } \langle A_j, X \rangle \leqslant b_j \quad j = 1, \ldots, s \\ &\quad X \succeq 0. \end{aligned} \tag{A.2}$$

Comparing this with (A.1), we identify that variables take the following values in our problem. First,

$$X = \mathrm{diag}(z_1, \ldots, z_{d^N}) \oplus R, \tag{A.3}$$

where $\oplus$ denotes direct sum. Thus X is a positive semidefinite matrix (of size $d^N + k$) because the variables z are non-negative and R is positive semidefinite. Then,

$$C = I_{d^N} \oplus 0_k, \tag{A.4}$$

where I_{d^N} is the identity matrix of size d^N and 0_k the zero matrix of size $k \times k$. The matrix constraints are given by

$$A_j = \mathrm{diag}(0, \ldots, -1, \ldots, 0) \oplus (-|v_k^j\rangle\langle v_k^j|) \quad 1 \leqslant j \leqslant d^N, \tag{A.5}$$

$$A_j = \mathrm{diag}(0, \ldots, -1, \ldots, 0) \oplus (|v_k^{j-d^N}\rangle\langle v_k^{j-d^N}|) \quad d^N < j \leqslant 2d^N, \tag{A.6}$$

where the -1 is in the jth position. Thus we have $s = 2d^N$ constraints. Note that one can always write the constraints of (A.2) with equalities by introducing additional slack variables [23]. Finally,

$$b = (-\lambda_1, \ldots, -\lambda_{d^N}, \lambda_1, \ldots, \lambda_{d^N})^{\mathrm{T}}. \tag{A.7}$$

We remark that the SDP is strongly feasible for $k < m$. To see that, note that the vectors $\{|v_k^i\rangle\}_{i=1}^k$ are linearly independent, hence there exists a biorthogonal basis $\{|w_k^i\rangle\}_{i=1}^k$. We use them to define $R = \sum_{i=1}^k |w_k^i\rangle\langle w_k^i|$. This R is positive definite ($R \succ 0$) and we choose z to be strictly positive ($z > 0$). This point is in the interior of the feasible region. It follows that there exists no duality gap in the SDP [23].

We have implemented the SDP with SeDuMi 1.3 [24] and the add-on Yalmip [25], with default options.

Appendix B. sos polynomials for the exponentially decaying distribution of eigenvalues

Here we prove result 8. That is, we show that the distance from the sos polynomial of equation (57) (which passes through the largest $k-2$ eigenvalues and 0) to the exponentially decaying distribution $\lambda_j = a\mathrm{e}^{-bj}$ (with $a, b > 0$) decays exponentially with k,

$$\sum_{i=1}^{n} |\lambda_i - p_k(\lambda_i)| \leqslant O(\mathrm{e}^{-k}). \tag{B.1}$$

We first upper bound the distance using the triangle inequality

$$\sum_{i=1}^{n} |\lambda_i - p_k(\lambda_i)| = \sum_{i=k-1}^{n} |\lambda_i - p_k(\lambda_i)| \leqslant \sum_{i=k-1}^{n} \lambda_i + \sum_{i=k-1}^{n} p_k(\lambda_i). \tag{B.2}$$

The first term in the last expression is equal to

$$\frac{\mathrm{e}^{-b(k-1)} - \mathrm{e}^{-b(n+1)}}{1-\mathrm{e}^{-b}} \tag{B.3}$$

and thus it decreases exponentially with k. In the remaining of the appendix, we will show that the last term also decreases exponentially with k.

First note that the last term of (B.2) can be upper bounded by

$$\sum_{i=k-1}^{n} p_k(\lambda_i) \leqslant \sum_{i=k-1}^{n} \lambda_i^2 \sum_{r=1}^{k-2} \frac{l_r(0)}{\lambda_r}, \tag{B.4}$$

where the first sum can again be computed exactly

$$\sum_{i=k-1}^{n} \lambda_i^2 = a^2 \frac{\mathrm{e}^{-2b(k-1)} - \mathrm{e}^{-2b(n+1)}}{1-\mathrm{e}^{-2b}}. \tag{B.5}$$

Now we consider $l_r(0)$,

$$l_r(0) = \prod_{j=1, j\neq r}^{k-2} \frac{1}{(\mathrm{e}^{-b(r-j)} - 1)^2} \tag{B.6}$$

and split the product into a term with $j < r$ times a term with $j > r$. To upper bound the term with $j < r$, we lower bound its denominator as

$$\prod_{j=1}^{r-1} (1-\mathrm{e}^{-b(r-j)})^2 \geqslant \prod_{j=-\infty}^{r-1} (1-\mathrm{e}^{-b(r-j)})^2 \tag{B.7}$$

$$= \left[\exp\left(\sum_{x=1}^{\infty} \ln(1-\mathrm{e}^{-bx})\right)\right]^2, \tag{B.8}$$

where we have defined the variable $x = r - j$. Now, note that $0 \geqslant \ln(1-\mathrm{e}^{-bx}) \geqslant \ln(1-\mathrm{e}^{-b})$. We lower bound this function by a straight line of the variable e^{-bx}, i.e. $\ln(1-\mathrm{e}^{-bx}) \geqslant -\alpha\mathrm{e}^{-bx}$,

where α is the slope of the function. We determine α by imposing $\ln(1-\mathrm{e}^{-b}) = -\alpha \mathrm{e}^{-b}$, thus yielding

$$\left[\exp\left(\sum_{x=1}^{\infty} \ln(1-\mathrm{e}^{-bx})\right)\right]^2 \geqslant \left[\exp\left(-\alpha \sum_{x=1}^{\infty} \mathrm{e}^{-bx}\right)\right]^2 \tag{B.9}$$

$$= (1-\mathrm{e}^{-b})^{2/(1-\mathrm{e}^{-b})} =: C. \tag{B.10}$$

Alternatively, this can be lower bounded by noting that the leftmost term is the Pochhammer function $((\mathrm{e}^{-b};\mathrm{e}^{-b})_\infty)^2$, where $(a;q)_n$ is defined as $(a;q)_n = \prod_{x=0}^{n}(1-aq^x)$. The function $f(b) = ((\mathrm{e}^{-b};\mathrm{e}^{-b})_\infty)^2$ is lower bounded by a finite value if b is larger than 0).

Now we consider the term with $j > r$ and lower bound its denominator as

$$\prod_{j=r+1}^{k-2} (\mathrm{e}^{-b(r-j)}-1)^2 \geqslant \prod_{j=r+1}^{j^\star} (1-\mathrm{e}^{-b(r-j)})^2 =: D, \tag{B.11}$$

where $j^\star$ is the smallest j such that $\mathrm{e}^{-b(r-j)} - 1 \geqslant 1$, that is, $j^\star = r + \lceil (\ln 2)/b \rceil$ (for example, for $b \geqslant 1$, $j^\star = r+1$). Note that C and D are independent of k, and the number of terms in D is independent of r (it only depends on b).

Using the lower bounds for the denominators of the parts with $j < r$ and that with $j > r$, we find

$$l_r(0) \leqslant \frac{1}{CD}. \tag{B.12}$$

We are interested in the sum of $l_r(0)/\lambda_r$, which is

$$\sum_{r=1}^{k-2} \frac{l_r(0)}{\lambda_r} = (\mathrm{e}^{b(k-1)} - \mathrm{e}^{b}) \frac{1}{(\mathrm{e}^b-1)aCD}. \tag{B.13}$$

Finally, putting together (B.5) and (B.13) we find

$$\sum_{i=k}^{n} p_k(\lambda_i) \leqslant (\mathrm{e}^{-2b(k-1)} - \mathrm{e}^{-2b(n+1)})\,(\mathrm{e}^{b(k-1)} - \mathrm{e}^{b}) \frac{a}{(1-\mathrm{e}^{-2b})(\mathrm{e}^b-1)CD}$$

$$= A\mathrm{e}^{-bk} + O(1), \tag{B.14}$$

where we have defined the constant A as the prefactor of e^{-bk} and the rest as independent of k. This proves the claim (B.1) and thus result 8.

References

[1] Hastings M 2006 *Phys. Rev.* B **73** 085115
[2] Fannes M, Nachtergaele B and Werner R F 1992 *Commun. Math. Phys.* **144** 443
[3] Perez-Garcia D, Verstraete F, Wolf M M and Cirac J I 2007 *Quantum Inform. Comput.* **7** 401
[4] Cirac J I, Michalakis S, Pérez-García D and Schuch N 2013 *Phys. Rev.* B **88** 115108
[5] Wahl T B, Tu H H, Schuch N and Cirac J I 2013 *Phys. Rev. Lett.* at press (arXiv:1308.0316v2)
[6] Wahl T B, Pérez-García D and Cirac J I 2012 *Phys. Rev.* A **86** 062314

[7] Schuch N, Pérez-García D and Cirac I 2011 *Phys. Rev.* B **84** 165139
[8] Pérez-García D, Wolf M M, Sanz M, Verstraete F and Cirac J I 2008 *Phys. Rev. Lett.* **100** 167202
[9] Schuch N, Perez-Garcia D and Cirac J I 2010 *Ann. Phys.* **325** 2153
[10] Verstraete F, Garcia-Ripoll J and Cirac J I 2004 *Phys. Rev. Lett.* **93** 207204
[11] Zwolak M and Vidal G 2004 *Phys. Rev. Lett.* **93** 207205
[12] Gouveia J, Parrilo P A and Thomas R 2013 *Math. Oper. Res.* **38** 248–64 (arXiv:1111.3164v2)
[13] Fiorini S, Massar S, Pokutta S, Tiwary H R and de Wolf R 2012 *STOC'12: Proc. 14th Symp. on Theory of Computing* pp 95–106 (arXiv:1111.0837v3)
[14] Jain R, Shi Y, Wei Z and Zhang S 2013 Efficient protocols for generating bipartite classical distributions and quantum states *IEEE Trans. Inf. Theory* **59** 5171–8 (arXiv:1203.1153v1)
[15] Cohen J E and Israel T 1993 *Linear Algebra and Its Applications* **190** 149–68
[16] Ziegler G M 2007 *Lectures on Polytopes* (Berlin: Springer)
[17] Moore C and Mertens S 2011 *The Nature of Computation* (Oxford: Oxford University Press)
[18] Schön C, Solano E, Verstraete F, Cirac J and Wolf M 2005 *Phys. Rev. Lett.* **95** 110503
[19] Blekherman G, Parrilo P A and Thomas R R (ed) 2013 *Semidefinite optimization and convex algebraic geometry* (*MOS-SIAM Series on Optimization*)
[20] Verstraete F and Cirac J I 2004 arXiv:cond-mat/0407066
[21] Wolf M M and Cirac J I 2008 *Commun. Math. Phys.* **279** 147–68
[22] Wolf M, Eisert J, Cubitt T and Cirac J 2008 *Phys. Rev. Lett.* **101** 150402
[23] Vandenberghe L and Boyd S 1996 *SIAM Rev.* **38** 49–95
[24] Sturm J F 1999 *Optimization methods and Software* **11–2** 625–53
[25] Lofberg J 2004 YALMIP: a toolbox for modeling and optimization in MATLAB *2004 IEEE Int. Symp. on Computer Aided Control Systems Design* pp 284–9