LEOPOLD-FRANZENS UNIVERSITÄT INNSBRUCK

MASTER THESIS

Approximate tensor decompositions: disappearance of many separations

Author:
Andreas KLINGLER

Supervisor:
Ass.-Prof. Dr. Gemma DE LAS CUEVAS

A thesis submitted in fulfillment of the requirements
for the degree of Master of Science

in the

Institut für Theoretische Physik

April 19, 2020

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Stellen, die wörtlich oder inhaltlich den angegebenen Quellen entnommen wurden, sind als solche kenntlich gemacht. Die vorliegende Arbeit wurde bisher in gleicher oder ähnlicher Form noch nicht als Magister-/Master-/Diplomarbeit/Dissertation eingereicht.

19. April 2020

Datum *Unterschrift*

LEOPOLD-FRANZENS UNIVERSITÄT INNSBRUCK

Abstract

Fakultät für Mathematik, Informatik und Physik
Institut für Theoretische Physik

Master of Science

Approximate tensor decompositions: disappearance of many separations

by Andreas KLINGLER

The field of tensor decompositions and matrix factorizations has become very popular in recent years since it has a wide range of applications in the description of complex systems. Although natural matrix factorizations are quite well understood, there are a lot of open questions in the field of matrix decompositions with additional restrictions. The picture is similar to the area of tensor decompositions. In this thesis, we will introduce different well-known notions of decompositions and factorizations and their corresponding ranks, which denote the number of necessary summands. We will extend this concept to a very general framework of tensor decompositions based on simplicial complexes with which we will study relations between the introduced ranks. It is shown that there exist separations between the different notions, namely one can be arbitrarily much larger than the other. For this purpose, we will introduce an idea of approximate ranks again based on simplicial complexes. We will define them as the minimal (exact) rank attained in the ε-neighborhood of the investigated element. We will study this concept on the tensor product space of matrices and the tensor product space of vectors using the Schatten p-norm and the entrywise p-norm, respectively. We will show that many separations between the ranks will disappear in the approximate case. This work opens, for example, the door to effective and useful descriptions of mixed states which is a critical issue in many applications.

Acknowledgements

First and foremost, I want to express my sincere gratitude to my research supervisor, Ass.-Prof. Dr. Gemma De las Cuevas, for giving me the opportunity to do research in this fascinating field and to work on this challenging and exciting project. She taught me the methodology to carry out the research and to present the research works as clearly as possible. She has always been encouraging in discussions, which kept up my joy and motivation working on this project.

I am also extremely grateful to Univ.-Prof. Dr. Tim Netzer for the immense support during this project. He always had an open door and a lot of patience, when I was struggling and even if I had trivial mathematical questions. He also inspired me with many fascinating discussions and challenging problems.

I would also like to extend my thanks to my parents for supporting me and making it possible to study physics and mathematics.

Finally, I would like to thank all my friends and colleagues who accompanied me through my studies for all the inspiring discussions.

Contents

Introduction

Many descriptions of complex systems, such as quantum systems, suffer from the problem that the number of parameters increases exponentially with the system size. This immense growth of complexity is one of the main limiting factors in the classical simulation of such systems. It hinders the theoretical understanding of processes like high-temperature superconductivity or studies of complex models like the 2-dimensional Hubbard model [31]. The mathematical framework describing these systems are mainly matrices and tensors, and various decompositions of these objects lead to efficient descriptions.

In the study of matrices, the *rank*, namely the number of linearly independent columns or rows, is a useful parameter to characterize the amount of information stored in the matrix. The rank of an $n \times m$ matrix specifies the number of vectors $a \in \mathbb{C}^n$ and $b \in \mathbb{C}^m$ necessary to decompose it into a sum of terms $a \cdot b^T$. One application of such decompositions is low-rank approximations of matrices. It turns out that the truncation of the singular value decomposition (i.e. neglecting the smallest singular values) gives the best fixed-rank approximation of the given matrix with respect to the Frobenius norm (i.e. the entrywise 2-norm). This result is known as the Eckart-Young theorem [17]. The field of matrix decompositions can be extended into two directions, the study of matrix decompositions with additional constraints and the field of tensor decompositions.

In recent years the study of decompositions with additional restrictions on the summands a, b have gained significant importance due to their extensive applicability in the fields of information theory, algebraic geometry, and optimization, to name a few [19]. In contrast to the natural rank-decomposition of matrices, much less is known about decompositions with additional constraints. For example, it is not clear for every notion of restricted decomposition whether a matrix can attain a particular decomposition with a fixed rank [3]. Nonetheless, there are many results regarding upper and lower bounds of ranks for particular classes of matrices with respect to different notions of ranks [19,21,22]. This leads to the concepts of separations between different ranks, namely one rank can be arbitrarily much larger than the other.

Another direction in generalizing the study of decompositions is the extension of matrix spaces to tensor product spaces. Such spaces are of particular interest not only in quantum many-body physics but also in electrical engineering, data analysis, or concerning machine learning [8, 20]. Investigations of particular examples, for instance, matrix product states in the context of 1-dimensional quantum many-body systems have led to useful descriptions of those systems. Nonetheless, a general theory of tensor decompositions presents a lot of challenges and limitations which arise from the limitations of matrix decompositions with additional restrictions. These have hindered progress to more sophisticated systems.

In this thesis, we introduce different known notions of matrix factorizations and tensor decompositions and extend these notions to approximate decompositions. We will investigate connections between these notions and properties of their different ranks. In particular, we will study the space of matrices

$$\mathcal{M}_d \otimes \mathcal{M}_d \otimes \cdots \otimes \mathcal{M}_d \tag{1}$$

where $\mathcal{M}_d$ denotes the space of complex $d \times d$ matrices and the space of tensors

$$\mathbb{C}^d \otimes \mathbb{C}^d \otimes \cdots \otimes \mathbb{C}^d. \tag{2}$$

We will examine several decompositions by considering different arrangements of summation indices and symmetries in the summation terms. These are motivated by the fact that many applications, for example, tensor networks, show that a reasonable arrangement implies a significant reduction of the rank and hence a much more effective description. The symmetries of the described system also translate to symmetries in the decompositions and further imply a reduction of the number of free parameters.

Another concept, we want to embed in the different decompositions, is the notion of positivity. In many areas of research, the objects which describe the system apply a notion of positivity. In quantum physics, for example, states will be represented by positive semidefinite (psd) matrices. We will investigate different local certificates of the fact that the global element is positive in tensor product spaces. It can be shown that effective decompositions of psd matrices in the space (1) suffer from the fact that there exists no sign of this global positivity in the parts of the decomposition. On the other hand, decompositions with a local certificate of positivity can be arbitrarily much more costly than the former decompositions. This behavior is called *separation between ranks* and has several implications. This discussion is not exclusive to the space of positive semidefinite matrices. Different notions of ranks of nonnegative tensors (i.e. elements in (2) where all entries are nonnegative) exhibit similar properties.

In the last part of this thesis, the main research project, we introduce notions of approximate decompositions of the tensor product spaces (1) and (2) and its corresponding ranks. Using recent results about approximate versions of the Carathéodory theorem, which gives upper bounds on the number of coefficients for convex combinations, we will investigate its implications to approximate decompositions. In particular, we will see that the separations between different notions of ranks, which appear in the exact case, will disappear in the approximate case.

The disappearance of the separations has many applications in different fields. We expect, for example, that this result opens the door for an approximate canonical form of matrix product density operators. A further interesting question regards local approximations. In this work, we only study approximate ranks with respect to the Schatten p-norm and the ℓ_p-norm for $p > 1$, which approximate the global element. Another perspective would be local approximations which only approximate the local elements of the tensor product decomposition.

Chapter 1

Tensor Networks and Matrix Decompositions

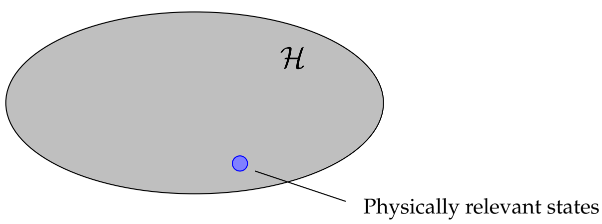
In order to describe quantum systems, the postulates of quantum mechanics constitute the basic mathematical framework. One feature of the axiomatization is the description of interacting systems utilizing tensor products. In particular, the Hilbert space of two interacting systems $\mathcal{H}_1, \mathcal{H}_2$ is given by its tensor product space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. This is the basic mathematical resource describing entanglement between the interacting systems. Extending this simple structure to more than two interacting systems illustrates one of the main challenges in the development of an applicable, manageable theory of many-body quantum systems: the size of the Hilbert space increases exponentially in the number of interacting systems. In particular, every state of an n-fold tensor product system, each of dimension d, can be described by

$$|\psi\rangle = \sum_{i_0,\ldots,i_{n-1}=1}^{d} c_{i_0,\ldots,i_{n-1}} |i_0, \ldots, i_{n-1}\rangle \tag{1.1}$$

where $c_{i_0,\ldots,i_n} \in \mathbb{C}$ are d^n different complex parameters constrained by the normalization condition

$$\sum_{i_0,\ldots,i_n=1}^{d} |c_{i_0,\ldots,i_{n-1}}|^2 = 1.$$

At first glance, it seems not entirely clear whether physically relevant quantum states are spread over this vast Hilbert space and exploit the whole range of degrees of freedom, or whether the situation is rather this one:



In quantum many-body physics, it is especially interesting to describe local Hamiltonians, namely, systems with restricted interaction length. The study of these systems employing tensor network states, for example, *matrix product states* led to a tremendous insight into this topic in the last 20 years. In particular, tensor network states turned out to be an adequate description of the tiny sub-manifold containing the physical states of local Hamiltonians (as visualized in the above figure).

The notion of *matrix product states (MPS)*, that is, descriptions of the form

$$|\psi_{\mathrm{MPS}}\rangle := \sum_{i_0,\ldots,i_n=1}^{d} \operatorname{tr}\left(A_{i_0}^{[0]} \cdot A_{i_1}^{[1]} \cdots A_{i_{n-1}}^{[n-1]}\right) |i_0, \ldots, i_{n-1}\rangle \tag{1.2}$$

where $A_{i_k}^{[k]}$ are $r \times r$ matrices turned out to be a practical framework to describe 1-dimensional noncritical systems (i.e. systems with an energy gap between ground state and first excited states) with an exponential decay in correlation. This is due to the fact that ground states of local gapped Hamiltonians obey an area law [23]. In particular, the entanglement entropy between two connected regions scales with the boundary of the regions which is a constant in the 1d case [18].

The central parameter in the MPS description, characterizing the efficiency, is the size of the matrices

$$A_{i_k}^{[k]} \in \mathcal{M}_r.$$

This parameter is usually called *bond dimension* of the system in physics literature [30]. It is shown that local critical and noncritical systems can be approximated by a MPS with bond dimension r depending polynomially in the system size n and the approximation error ε with respect to the 2-norm [33]. In particular, this implies an exponential improvement of the description compared to the representation of the form (1.1). Hence, tensor networks are state-of-the-art in many situations, efficiently describing quantum many-body systems.

In recent years the idea of tensor network decompositions describing pure states was adapted to tensor decompositions of mixed states [34] which are needed for the description of open systems. The matrix product density operator form (short MPDO) is also relevant for boundary theories in 2d systems [6]. One profound difference between the description of pure states with vectors $|\psi\rangle$ and the description of mixed states with density matrices ρ is the fact that a density matrix has to be additionally positive semidefinite (psd). Embedding this behavior to tensor network decompositions leads to several challenges in the field. For example, the MPDO form does not have a local certificate of positivity; namely, one cannot tell from the local matrices whether the global element is psd or not. This behavior implies that the description in this form cannot be interpreted as a mixed state. Hence, many numerical methods developed for pure states do not apply to the MPDO-form. Another approach that contains a local certificate of positivity is the idea of purification. It is a well-known fact that every mixed state attains a description as a pure state in a larger Hilbert space. Nonetheless, we will see that this approach leads to an arbitrarily more costly description of the system. We will study this behavior in the following sections more thoroughly.

The principal methodology to show this behavior will be a correspondence between diagonal psd matrices and entrywise nonnegative matrices. We will introduce and study this correspondence in Section 2.4. Decompositions of nonnegative matrices received a lot of attention in recent years and have, besides this correspondence to decompositions of psd matrices, a considerable variety of applications in, for example, algebraic geometry or information theory, to name a few. We will study the applications to communication complexity more thoroughly in Section 3.

The goal of this chapter is to introduce different known tensor network decompositions of psd matrices and matrix decompositions in their natural way and also in the language of tensor decompositions, namely as a sum of elementary tensors with a special arrangement of the summation indices and different local certificates of positivity. This will be a starting point for generalized notions of tensor decompositions based on simplicial complexes in Chapter 2. In addition, we will study the different notions of ranks and their applications.

1 Decompositions of positive semidefinite matrices

In the following we will extend the idea of the matrix product state decompositions, as already mentioned in (1.2) to mixed states. In particular, we define different notions of tensor network decompositions for density matrices and will relate them to particular tensor decompositions. We will denote the space of all complex $d \times d$ matrices by $\mathcal{M}_d$, the hermitian conjugate by † and the complex conjugate by *. Recall that a matrix $\rho \in \mathcal{M}_d$ is called a density matrix if it is hermitian, i.e. $\rho^\dagger = \rho$, positive semidefinite (denoted by $\rho \geq 0$), i.e. all eigenvalues are nonnegative, and $\mathrm{tr}(\rho) = 1$, i.e. all eigenvalues sum up to 1.

In Section 1.1, we will introduce the *matrix product density operator form* as an analog to the matrix product state for density matrices. Further, we will show a relation of this decomposition to a special type of tensor decompositions, which will be a starting point for further investigations in this section and will also give rise to extensions of tensor decompositions on simplicial complexes. This more sophisticated decomposition framework will be studied in Chapter 2. In Section 1.2, we will introduce the *local purification form* as a decomposition with a *local certificate of positivity* of the global element. Section 1.3 examines the defined notions of decomposition defined in Section 1.1 and Section 1.2 with an additional constraint in the symmetry of the decompositions. We will investigate the *translational invariant* tensor decompositions, which can also be seen as a starting point for the extensions to more sophisticated symmetric decompositions in Chapter 2.

1.1 The matrix product density operator form

The *matrix product density operator (MPDO) form*, first introduced in [34], is a canonical generalization of the matrix product state decomposition to density matrices. It is defined by

$$\rho := \sum_{i_0,\ldots,i_{n-1},j_0,\ldots,j_{n-1}=1}^{d} \mathrm{tr}\left(\rho^{[0]}_{(i_0,j_0)} \cdot \rho^{[1]}_{(i_1,j_1)} \cdots \rho^{[n-1]}_{(i_{n-1},j_{n-1})}\right) |i_0,\ldots,i_{n-1}\rangle\langle j_0,\ldots,j_{n-1}| \quad (1.3)$$

where $\rho^{[k]}_{(i_k,j_k)} \in \mathcal{M}_r$ and $k \in \{0,\ldots,n-1\}$. The visualization of this decomposition as a tensor network is shown in Figure 1.1.

Considering the visual tensor network representation, the drawn connections denoted by the indices $\alpha_0,\ldots,\alpha_{n-1}$ describe the contractions of the different matrices $\rho^{[k]}_{(i_k,j_k)}$.

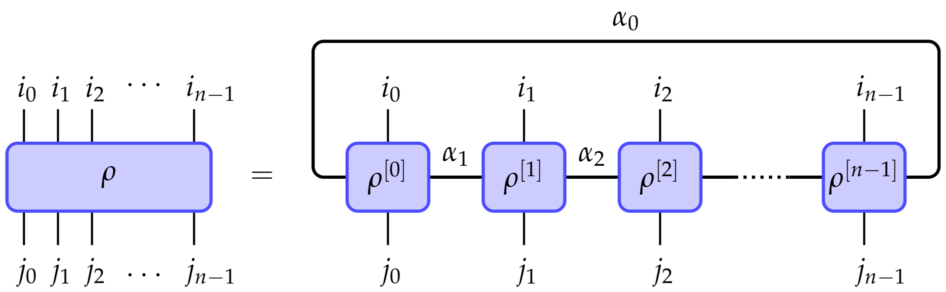


FIGURE 1.1: Tensor network representation of the *matrix product density operator form*.

Hence, these indices range from 1 to r, which is the dimension of the matrices. The contractions with respect to the indices $\alpha_1, \ldots, \alpha_{n-1}$ correspond to the matrix multiplications, the contraction with respect to index α_0 corresponds to the trace. In physics literature, r is usually called the *bond dimension* and is a parameter for the efficiency of the description [30]. In the following definition, we will (re-)define this decomposition in the language of *tensor decompositions*. Based on this decomposition, we will now consider different notions of ranks corresponding to the bond dimension following [14].

Definition 1.1 (MPDO-form)**.** *Let $\rho \in \mathcal{M}_d \otimes \cdots \otimes \mathcal{M}_d \cong \mathcal{M}_{d^n}$. A matrix product density operator form is given by*

$$\rho = \sum_{\alpha_0,\ldots,\alpha_{n-1}=1}^{r} \rho^{[0]}_{\alpha_0,\alpha_1} \otimes \rho^{[1]}_{\alpha_1,\alpha_2} \otimes \cdots \otimes \rho^{[n-1]}_{\alpha_{n-1},\alpha_0} \tag{1.4}$$

where $\rho^{[k]}_{\alpha,\beta} \in \mathcal{M}_d$ for all $\alpha, \beta \in \{1, \ldots, r\}$ and $k \in \{0, \ldots, n-1\}$.

The minimal integer r attaining a decomposition of the form (1.4) *is called* operator Schmidt rank *of ρ, in short* $\mathrm{osr}(\rho)$.

Note that every representation in (1.3) given by the family of matrices

$$\left\{\rho^{[k]}_{(i_k,j_k)} \in \mathcal{M}_r : k \in \{0, \ldots, n-1\} \text{ and } i_k, j_k \in \{1, \ldots, d\}\right\}$$

corresponds to a representation in (1.4) given by

$$\left\{\rho^{[k]}_{\alpha_k,\alpha_{k+1}} \in \mathcal{M}_d : k \in \{0, \ldots, n-1\} \text{ and } \alpha_k, \alpha_{k+1} \in \{1, \ldots, r\}\right\}$$

using the relation

$$\left(\rho^{[k]}_{\alpha_k,\alpha_{k+1}}\right)_{i_k,j_k} = \left(\rho^{[k]}_{(i_k,j_k)}\right)_{\alpha_k,\alpha_{k+1}} \tag{1.5}$$

for all $k \in \{0, \ldots, n-1\}$, $\alpha_k, \alpha_{k+1} \in \{0, \ldots, r\}$ and $i_k, j_k \in \{0, \ldots, d\}$.

Further note that the tensor network representation in Figure 1.1 also gives an interpretation of the tensor decomposition. It describes the arrangement of the indices in the decomposition (1.4). In particular, every local matrix in the elementary tensors shares a joint index with its neighbors. This arrangement will be generalized to more sophisticated geometries in Chapter 2.

The matrix product density operator form can also be defined with open boundary conditions. This decomposition differs from the MPDO form in the fact that the connection denoted by index α_0 disappears. In other words, there is no connection between the first and the last local space.

Remark 1.2 (MPDO-form with open boundary conditions)**.** Let $\rho \in \mathcal{M}_d \otimes \cdots \otimes \mathcal{M}_d \cong \mathcal{M}_{d^n}$. A *MPDO form with open boundary conditions (obc)* is a decomposition

$$\rho := \sum_{i_0,\ldots,i_{n-1},j_0,\ldots,j_{n-1}=1}^{d} \rho^{[0]}_{(i_0,j_0)} \cdot \rho^{[1]}_{(i_1,j_1)} \cdots \rho^{[n-1]}_{(i_{n-1},j_{n-1})} |i_0,\ldots,i_{n-1}\rangle\langle j_0,\ldots,j_{n-1}|$$

where $\rho^{[0]}_{(i_0,j_0)} \in \mathbb{C}^{1\times r}$, $\rho^{[n-1]}_{(i_{n-1},j_{n-1})} \in \mathbb{C}^{r\times 1}$ and $\rho^{[k]}_{(i_k,j_k)} \in \mathcal{M}_r$ for $1 \leq k \leq n-2$. Using the correspondence in (1.5), we also can write the MPDO in a tensor decomposition form.

$$\rho = \sum_{\alpha_1,\ldots,\alpha_{n-1}=1}^{r} \rho^{[0]}_{\alpha_1} \otimes \rho^{[1]}_{\alpha_1,\alpha_2} \otimes \cdots \otimes \rho^{[n-2]}_{\alpha_{n-2},\alpha_{n-1}} \otimes \rho^{[n-1]}_{\alpha_{n-1}}$$

The above defined decomposition can again be described as a tensor network shown in Figure 1.2. △

Note that the MPDO-decomposition with open boundary conditions corresponds in the bipartite case (i.e. only one tensor product) to the trivial tensor decomposition

$$\rho = \sum_{\alpha=1}^{r} \rho^{[0]}_{\alpha} \otimes \rho^{[1]}_{\alpha}.$$

We will use this fact later when studying the correspondence between decompositions of bipartite psd matrices and factorizations of nonnegative tensors in Section 2.4.

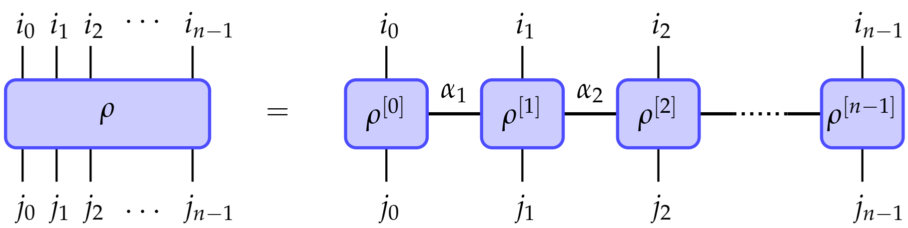


FIGURE 1.2: Tensor network representation of the MPDO form with open boundary conditions.

One major problem of the above-defined decompositions is the lack of a local certificate of positivity. In other words, for a given $\rho \geq 0$ it is not possible to detect the fact of global positivity in the local matrices $\rho_{\alpha,\alpha'}$. For example, it was shown in [11] that for an MPDO-decomposition of a psd matrix ρ it is in general undecidable to determine whether an extended version of the same MPDO to an arbitrarily larger tensor product spaces remains positive semidefinite. It follows that there exists no algorithm which determines an answer to this particular question. This implies that there cannot exist a computable certificate of positivity in general in the MPDO form.

Nonetheless, in the special case of separable states, it is possible to construct an MPDO-form with a local certificate of positivity. Recall that a state $\rho \in \mathcal{M}_d \otimes \cdots \otimes \mathcal{M}_d$ is called *separable* if and only if it can be written in the form

$$\rho = \sum_{\alpha=1}^{r} p_\alpha \cdot \rho_\alpha^{[0]} \otimes \rho_\alpha^{[1]} \otimes \cdots \otimes \rho_\alpha^{[n-1]} \qquad p_\alpha \geq 0 \text{ and } \sum_{\alpha=1}^{r} p_\alpha = 1$$

where $\rho_\alpha^{[0]} \otimes \rho_\alpha^{[1]} \otimes \cdots \otimes \rho_\alpha^{[n-1]} \geq 0$ are product states.

Definition 1.3 (Separable MPDO-decomposition)**.** *Let $\rho \in \mathcal{M}_d \otimes \cdots \otimes \mathcal{M}_d$ be separable. A* separable matrix product density operator decomposition *is a decomposition of the form* (1.4) *additionally fulfilling the condition*

$$\rho_{\alpha,\alpha'}^{[i]} \geq 0 \quad \text{for all } \alpha, \alpha' \in \{1, \ldots, r\} \text{ and } i \in \{0, \ldots, n-1\}.$$

Note that every state is separable if and only if there exists a separable MPDO-decomposition [14]. Hence, the fact that separable states are decomposable into product states can be generalized to MPDO-forms.

In the end, we want to present a generalization of the matrix product density operator form, the so-called *matrix product operator* form. This decomposition is valid for arbitrary operators and will be especially necessary to define the *local purification form*, which is a decomposition with a local certificate of positivity for all psd matrices.

Definition 1.4 (Matrix product operator (MPO) form)**.** *Consider an operator*

$$\sigma \in \mathcal{M}_{d_0,d'_0} \otimes \mathcal{M}_{d_1,d'_1} \otimes \cdots \otimes \mathcal{M}_{d_{n-1},d'_{n-1}}$$

where $\mathcal{M}_{d_i,d'_i}$ denotes the space of all complex $d_i \times d'_i$ matrices. The matrix product operator *form is a decomposition of the form*

$$\sigma = \sum_{\alpha_0,\ldots,\alpha_{n-1}=1}^{r} \sigma_{\alpha_0,\alpha_1}^{[0]} \otimes \sigma_{\alpha_1,\alpha_2}^{[1]} \otimes \cdots \otimes \sigma_{\alpha_{n-1},\alpha_0}^{[n-1]}. \tag{1.6}$$

where $\sigma_{\alpha_k,\alpha_{k+1}}^{[k]} \in \mathcal{M}_{d_k,d'_k}$.

The minimum integer r attaining a matrix product operator decomposition of σ is again called operator Schmidt rank of σ, denoted $\operatorname{osr}(\sigma)$.

Note that in the computational basis notation, similar to (1.3), the MPO form again consists of a family of $r \times r$ matrices. The only difference is the fact that the indices i_k range in $\{1, \ldots, d_k\}$ and the indices j_k in $\{1, \ldots, d'_k\}$.

1.2 The local purification form

In the following section, we will introduce the *local purification form*, which is a decomposition similar to the matrix product density operator form but with a local certificate of positivity. The local purification form is strongly related to the purification of mixed states.

To begin, recall the fact that every mixed state can be interpreted as a reduced state (by tracing out the ancillary degrees of freedom) of a pure state in a bigger Hilbert space. This fact formulated in the following well-known theorem [29].

Theorem 1.5. *Let $\rho \in \mathcal{M}_d$ be a mixed state. Then there exists a* purification *of ρ, i.e. a Hilbert space $\mathcal{H}_E$ and a pure state $|\psi\rangle \in \mathbb{C}^d \otimes \mathcal{H}_E$ such that*

$$\rho = \mathrm{tr}_{\mathcal{H}_E} |\psi\rangle\langle\psi|$$

Moreover, $\mathcal{H}_E$ can be chosen with $\dim(\mathcal{H}_E) = d$.

Proof. Since ρ is positive semidefinite, diagonalization yields

$$\rho = \sum_{i=1}^{d} \lambda_i |v_i\rangle\langle v_i|$$

where $\lambda_i \geq 0$ and $\{|v_i\rangle\}_{i=1}^d$ is an orthonormal basis. Define

$$|\psi\rangle = \sum_{i=1}^{d} \sqrt{\lambda_i} |v_i\rangle |v_i\rangle \in \mathbb{C}^d \otimes \mathcal{H}_E$$

where $\mathcal{H}_E = \mathbb{C}^d$. We obtain that $\rho = \mathrm{tr}_{\mathcal{H}_E} |\psi\rangle\langle\psi|$ where $\mathrm{tr}_{\mathcal{H}_E}$ denotes the partial trace over the space $\mathcal{H}_E$. Moreover, since $\mathrm{tr}\rho = 1$, we have $\sum_{i=1}^d \lambda_i = 1$ and hence $\langle\psi|\psi\rangle = 1$. □

Note that for an isometry $V : \mathcal{H}_E \to \mathcal{H}_{E'}$ (i.e. $V^\dagger V = id$) and a purification $|\psi\rangle$ of ρ, the state

$$(id_{\mathbb{C}^d} \otimes V)|\psi\rangle \in \mathbb{C}^d \otimes \mathcal{H}_{E'}$$

is again a purification of ρ. Moreover, all purifications of ρ are related by an isometry of the above type [29]. Instead of defining purifications as a pure state together with a partial trace for the reconstruction of ρ, we can equivalently write a purification as an operator.

Remark 1.6 (Purification as operator)**.** Since we can write every purification of the mixed state $\rho \in \mathcal{M}_d$ as

$$|\psi\rangle = \sum_{i=1}^{d} \sum_{j=1}^{m} \psi_{i,j} |i\rangle |v_j\rangle \in \mathbb{C}^d \otimes \mathcal{H}_E$$

where $\{|v_j\rangle\}_{j=1}^m$ is an orthonormal basis of $\mathcal{H}_E$, we can define the operator

$$\sigma = \sum_{i=1}^{d} \sum_{j=1}^{m} \psi_{i,j} |v_j\rangle\langle i| \in \mathcal{M}_{m,d}.$$

The reduction to ρ (i.e. the partial trace over $\mathcal{H}_E$ applied to $|\psi\rangle$) is given in this case by

$$\sigma^\dagger \sigma = \sum_{i,j,i',j'} \psi^*_{i',j'} \psi_{i,j} |i'\rangle\langle v_{j'}|v_j\rangle\langle i| = \mathrm{tr}_{\mathcal{H}_E} |\psi\rangle\langle\psi| = \rho.$$

△

This remark implies that the existence of a purification for psd matrices relates to the fact that a matrix ρ is psd if and only if $\rho = \sigma^\dagger \sigma$ for some matrix $\sigma \in \mathcal{M}_{m,d}$. Moreover, the non-uniqueness of purifications relates to the fact that for a given isometry V on an m-dimensional space we have $(V\sigma)^\dagger V\sigma = \sigma^\dagger V^\dagger V \sigma = \sigma^\dagger \sigma$. The constructed purification $|\psi_0\rangle$ in the proof of Theorem 1.5 can be associated with the unique positive semidefinite square root of ρ

$$\sigma_0 = \sum_{i=1}^{d} \sqrt{\lambda_i} |v_i\rangle\langle v_i|$$

where $|v_i\rangle$ is the ith eigenstate of ρ with corresponding eigenvalue $\lambda_i \geq 0$.

For the rest of this thesis, we will always see a purification as an operator, that is, in the language of Remark 1.6.

In the following, we will define an analog to the MPDO form, the *local purification form*. We will see that this decomposition has, in contrast to the MPDO form, a local certificate of positivity for all positive semidefinite matrices.

Definition 1.7 (Local purification form). *Let $\rho \in \mathcal{M}_d \otimes \cdots \otimes \mathcal{M}_d \cong \mathcal{M}_{d^n}$. A* local purification form *of ρ is given by a purification*

$$\sigma \in \mathcal{M}_{d_0,d} \otimes \cdots \otimes \mathcal{M}_{d_{n-1},d}$$

in the matrix product operator form (1.6). *The corresponding tensor network decomposition is depicted in Figure 1.3.*

The minimum integer r among all purifications and MPO decompositions is called purification-rank. *In other words,*

$$\textit{puri-rank}(\rho) = \min\{\textit{osr}(\sigma) : \rho = \sigma^\dagger \sigma\}.$$

Remark 1.8. By construction, every purification σ can be written as

$$\sigma = V\sigma_0$$

with an isometry $V : \mathcal{H}_E \to \mathcal{H}_{E'}$ where $\mathcal{H}_E := \left(\mathbb{C}^d\right)^{\otimes n}$ and the purification σ_0 is the positive square root of ρ given by

$$\sigma_0 = \sum_{j=0}^{m} \sqrt{\lambda_j} |v_j\rangle\langle v_j| \in \mathcal{M}_{d^n}$$

where $|v_j\rangle$ is the jth eigenstate of ρ with eigenvalue $\lambda_j \geq 0$. Hence, the purification rank can be equivalently written as

$$\text{puri-rank}(\rho) = \min\{\text{osr}(V\sigma_0) : V \text{ isometry}\}.$$

△

The advantage of the local purification form is the existence of a local certificate of positivity. In the following, we want to study this behavior more thoroughly.

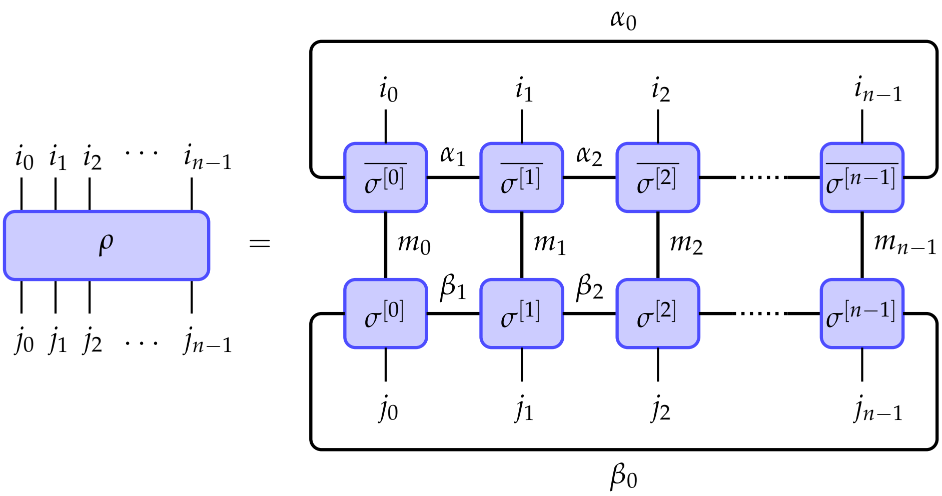


FIGURE 1.3: Tensor network of a purification $\rho = \sigma^\dagger \sigma$ where σ attains a matrix product operator form. The contractions with index m_k show the matrix multiplication of $\sigma^\dagger\sigma$. The order-4 tensors in the upper row are given by $\left(\overline{\sigma^{[k]}}\right)_{\alpha,\beta,i,j} := \sigma^{[k]}_{\alpha,\beta,j,i}$.

Let $\rho \in \mathcal{M}_d \otimes \cdots \otimes \mathcal{M}_d \cong \mathcal{M}_{d^n}$ be a positive semidefinite matrix and

$$\sigma = \sum_{\alpha_0,\ldots,\alpha_{n-1}=1}^{r} \sigma^{[0]}_{\alpha_0,\alpha_1} \otimes \sigma^{[1]}_{\alpha_1,\alpha_2} \otimes \cdots \otimes \sigma^{[n-1]}_{\alpha_{n-1},\alpha_0}$$

a purification of ρ. Performing the matrix multiplication $\rho = \sigma^\dagger \sigma$ leads to a decomposition

$$\rho = \sum_{\alpha_0,\ldots\alpha_{n-1},\beta_0,\ldots,\beta_{n-1}=1}^{r} \rho^{[0]}_{\alpha_0,\alpha_1,\beta_0,\beta_1} \otimes \rho^{[1]}_{\alpha_1,\alpha_2,\beta_1,\beta_2} \otimes \cdots \otimes \rho^{[n-1]}_{\alpha_{n-1},\alpha_0,\beta_{n-1},\beta_0}$$

where we define

$$\rho^{[i]}_{\alpha,\alpha',\beta,\beta'} := \left(\sigma^{[i]}_{\alpha,\alpha'}\right)^\dagger \cdot \left(\sigma^{[i]}_{\beta,\beta'}\right) \tag{1.7}$$

Rearranging the indices of the tensor $\rho^{[i]}_{\alpha,\alpha',\beta,\beta',k,l}$ to a square matrix as

$$\rho^{[i]} = \sum_{k,l,\alpha,\alpha',\beta,\beta'} \left(\rho^{[i]}\right)^{k,l}_{\alpha,\alpha',\beta,\beta'} |k,\alpha,\alpha'\rangle\langle l,\beta,\beta'| \in \mathcal{M}_{d\cdot r^2}$$

and rearranging the indices of the local tensors $\sigma^{[i]}_{\beta,\beta',m,l}$ to a matrix

$$\sigma^{[i]} = \sum_{m,l,\beta,\beta'} \left(\sigma^{[i]}\right)^{m,l}_{\beta,\beta'} |m\rangle\langle l,\beta,\beta'| \in \mathcal{M}_{M,d\cdot r^2}$$

where M is the dimension of the local ancillary space indexed by m, makes clear the local certificate of positivity. This finally leads to the necessary and sufficient condition for positive semidefiniteness

$$\rho^{[i]} = \left(\sigma^{[i]}\right)^\dagger \cdot \left(\sigma^{[i]}\right) \geq 0 \quad \text{for all } i \in \{0,\ldots,n-1\}$$

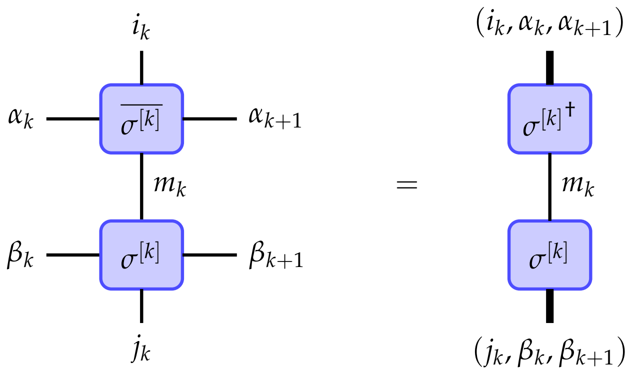


FIGURE 1.4: Visualization of the local certificate of positivity for a purification. This figure shows one local part in the purification of Figure 1.3.

where we have used Relation (1.7). Figure 1.4 shows this fact in the language of tensor networks.

Although the local purification form has a local certificate of positivity, its rank is in general larger than the operator Schmidt rank. More precisely, some examples have a constant operator Schmidt rank and a purification rank increasing in the dimension of the system. This behavior appears even in the simplest nontrivial case, namely bipartite matrices diagonal in the computational basis [12, 14, 15]. We will study this behavior in the more general setting of tensor decompositions on simplicial complexes in Chapter 2.

1.3 Translationally invariant decompositions

Symmetries play a central role in theoretical physics. In particular, imposing symmetries often leads to a massive reduction of free parameters and hence a much more apt description of the physical system. In the following section, we want to introduce translational invariance, a canonical symmetry in the study of 1d systems, which can be described by matrix product states and matrix product density operators respectively for particular cases. In Chapter 2 we will extend the notion of translational invariance to arbitrary symmetries on simplicial complexes.

At the start, we introduce the notion of a translational invariant for an element in a tensor product space.

Definition 1.9. *Let $\rho \in \mathcal{M}_d \otimes \cdots \otimes \mathcal{M}_d \cong \mathcal{M}_{d^n}$. We call ρ translational invariant (t.i.), if $T\rho T^\dagger = \rho$, where T denotes the unitary translation operator of the physical indices*

$$T = \sum_{i_0,\ldots,i_{n-1}=1}^{d} |i_1, i_2, \ldots, i_{n-1}, i_0\rangle\langle i_0, i_1, \ldots, i_{n-2}, i_{n-1}|.$$

In other words, ρ is translational invariant if it commutes with the translation operator T. This implies that ρ and T are simultaneously diagonalizable [25]. The above-defined invariance can be equivalently defined by introducing the permutation

$$g : \{0, \ldots, n-1\} \to \{0, \ldots, n-1\} : i \mapsto i+1$$

as a group action where addition is meant to be modulo n. In particular, we can define the linear mapping $T(\cdot)T^\dagger$ by its action on the generating set of $\mathcal{M}_d \otimes \cdots \otimes \mathcal{M}_d$, namely all elementary tensors as

$$\begin{aligned} T\Big(\rho^{[0]} \otimes \cdots \otimes \rho^{[n-1]}\Big)T^\dagger &= \rho^{[g0]} \otimes \cdots \otimes \rho^{[g(n-1)]} \\ &= \rho^{[1]} \otimes \rho^{[2]} \otimes \cdots \otimes \rho^{[n-1]} \otimes \rho^{[0]}. \end{aligned} \tag{1.8}$$

In addition, translational invariance of ρ implies invariance for all permutations in the group generated by g,

$$G := \{g^n : n \in \mathbb{N}\} \subseteq S_n := \Big\{\pi : \{0, \ldots, n-1\} \to \{0, \ldots, n-1\} \text{ permutation}\Big\}$$

where g^n is the n-fold composition of the translation g.

Let us now define the notion of translational invariant decompositions. A translational invariant decomposition is given by the same notions of decomposition defined in Section 1.1 and Section 1.2 with the additional constraint that *every* elementary tensor in the decomposition is translationally invariant. This implies that the dependence of the tensor product site index in every matrix in the elementary tensor disappears.

Definition 1.10 (Translational invariant decompositions). *Let ρ be a matrix in the space $\mathcal{M}_d \otimes \cdots \otimes \mathcal{M}_d \cong \mathcal{M}_{d^n}$.*

(i) A translational-invariant matrix product density operator *(t.i.-MPDO) decomposition is given by*

$$\rho = \sum_{\alpha_0, \ldots, \alpha_{n-1}=1}^{r} \rho_{\alpha_0,\alpha_1} \otimes \rho_{\alpha_1,\alpha_2} \otimes \cdots \otimes \rho_{\alpha_{n-1},\alpha_0} \tag{1.9}$$

The minimal integer r among all t.i.-MPDO forms is called translational invariant operator Schmidt rank*, denoted*

$$\textit{ti-osr}(\rho).$$

(ii) The separable t.i.-MPDO *form of ρ is given by a decomposition of the form (1.9) with the additional property*

$$\rho_{\alpha,\alpha'} \geq 0 \quad \textit{for all } \alpha, \alpha' \in \{1, \ldots r\}$$

The minimal integer r among all t.i. separable decompositions is called t.i. separable rank*, denoted*

$$\textit{ti-sep-rank}(\rho).$$

(iii) A translational invariant local purification *of ρ is given by a purification $\rho = \sigma^\dagger\sigma$ where*

$$\sigma \in \mathcal{M}_{d_0,d} \otimes \cdots \otimes \mathcal{M}_{d_{n-1},d}$$

is given in the translational invariant MPO-form

$$\sigma = \sum_{\alpha_0, \ldots, \alpha_{n-1}=1}^{r} \sigma_{\alpha_0,\alpha_1} \otimes \sigma_{\alpha_1,\alpha_2} \otimes \cdots \otimes \sigma_{\alpha_{n-1},\alpha_0}. \tag{1.10}$$

The minimal r among all purifications and t.i. MPO forms is called t.i. purification rank, *denoted*

$$\textit{ti-puri-rank}(\rho).$$

Note that the difference between non-t.i.- and t.i.-decompositions is the fact that the local tensors of the latter do not depend on the tensor product index. Further note that ρ is translational invariant if and only if there exists a t.i.-MPDO form. In particular, a t.i.-MPDO form can be constructed from a general MPDO form. Similar results hold for psd t.i. ρ and t.i. local purifications and separable t.i. ρ and t.i. separable decompositions (see [14] for details).

Although every translational invariant state attains a translational invariant decomposition, the corresponding t.i.-rank can be much larger than the non-t.i. rank.

Example 1.11 (The GHZ- and the W-state)**.** In the following example we want to study the different notions of decompositions on the GHZ- and W-state which are both translational invariant.

(i) Consider the GHZ-state given by

$$|GHZ\rangle := \left(\frac{|0\rangle^{\otimes n} + |1\rangle^{\otimes n}}{\sqrt{2}}\right)$$

where $|0\rangle^{\otimes n}$ and $|1\rangle^{\otimes n}$ denote the n-fold tensor product state of $|0\rangle$ and $|1\rangle$ respectively. We consider the corresponding density matrix

$$\rho_{GHZ} := |GHZ\rangle\langle GHZ|$$

Since ρ_{GHZ} is a pure state (i.e. as a product of a column and a row vector a rank-1 matrix), a MPS decomposition of $|GHZ\rangle$ is already a valid purification. Further it is easy to verify that ρ_{GHZ} is translational invariant. A t.i.-MPS decomposition of $|GHZ\rangle$, i.e. a decomposition of the form

$$|\psi_{\text{t.i.-MPS}}\rangle := \sum_{i_0,\ldots,i_n=1}^{d} \operatorname{tr}\left(A_{i_0} \cdot A_{i_1} \cdots A_{i_{n-1}}\right) |i_0, \ldots, i_{n-1}\rangle \tag{1.11}$$

is given by

$$A_0 := 2^{-\frac{1}{2n}} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad A_1 := 2^{-\frac{1}{2n}} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Hence,

$$\text{puri-rank}(\rho_{GHZ}) = \text{ti-puri-rank}(\rho_{GHZ}) = 2.$$

Since for pure states we have $\text{ti-puri-rank}(\rho)^2 = \text{ti-osr}(\rho)$ (similar to the non-t.i. case, see [14, Prop. 47] for details) we have

$$\text{osr}(\rho_{GHZ}) = \text{ti-osr}(\rho_{GHZ}) = 4.$$

A realization of an MPDO (in the form of (1.3)) with rank 4 is given by the 4×4 matrices

$$\rho_{(0,0)} := 2^{-\frac{1}{n}} \cdot E_{11}, \quad \rho_{(0,1)} := 2^{-\frac{1}{n}} \cdot E_{22}, \quad \rho_{(1,0)} := 2^{-\frac{1}{n}} \cdot E_{33}, \quad \rho_{(1,1)} := 2^{-\frac{1}{n}} \cdot E_{44}$$

where E_{ij} is the matrix which is 1 in the (i, j)-entry and zero elsewhere.

(ii) Consider the W-state given by

$$|W\rangle := \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \sigma_x^{(j)} |0\rangle^{\otimes n}$$

where $\sigma_x^{(j)}$ acts as $\sigma_x := |0\rangle\langle 1| + |1\rangle\langle 0|$ on the jth local space and as an identity elsewhere. For example in the case of $n = 4$ we have

$$|W\rangle = \frac{1}{2}\Big(|0,0,0,1\rangle + |0,0,1,0\rangle + |0,1,0,0\rangle + |1,0,0,0\rangle\Big).$$

We consider again the corresponding density matrix $\rho_W := |W\rangle\langle W|$. Similar to (i), $|W\rangle$ is a purification which can be decomposed by a (non t.i.) MPS decomposition (i.e. of the form (1.2)) by

$$A_0^{[0]} := \frac{1}{\sqrt{n}} \cdot \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad A_1^{[0]} = \frac{1}{\sqrt{n}} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

and for $i \in \{1, \ldots, n-1\}$

$$A_0^{[i]} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_1^{[i]} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Hence puri-rank$(\rho_W) = 2$ and since ρ_W is a pure state, osr$(\rho_W) = 4$ by [14, Prop. 47]. On the other hand, it is shown in Ref. [14] that

$$\text{ti-puri-rank}(\rho_W) \geq \sqrt{n}.$$

(iii) Consider a mixed state version of the W-state, i.e.

$$\rho := \frac{1}{n} \sum_{j=0}^{n-1} \sigma_x^{(j)} (|0\rangle\langle 0|)^{\otimes n} \sigma_x^{(j)}.$$

For example in the case of $n = 3$ the density matrix has the form

$$\rho = \frac{1}{3}\Big(|0,0,1\rangle\langle 0,0,1| + |0,1,0\rangle\langle 0,1,0| + |1,0,0\rangle\langle 1,0,0|\Big).$$

By definition, ρ is separable and translational invariant but the upper decomposition is not explicitly translational invariant. Similar to (ii) it is shown in Ref. [14] that

$$\text{sep-rank}(\rho) = 2 \quad \text{and} \quad \text{ti-sep-rank}(\rho) \geq \sqrt{n}$$

△

2 Factorizations of nonnegative matrices

We have seen in Section 1 different notions of decompositions on the tensor product space $\mathcal{M}_d \otimes \cdots \otimes \mathcal{M}_d$ with the canonical notion of positivity given by positive semidefinite matrices. In the following section, we introduce a different approach to decompositions on the space of nonnegative matrices, namely matrices with nonnegative entries. In the language of matrix analysis, the rank of a matrix is the number of linearly independent columns or rows, respectively, of the given matrix. We will see that this parameter also corresponds to the number of rank-1 matrices (i.e. matrices, which can be written as $v \cdot w^t$ for two vectors v, w) necessary for a valid decomposition. Based on this decomposition, further factorizations with additional constraints on the elementary elements of the decomposition can be defined.

This section is organized as follows: In Section 2.1 and Section 2.2, we introduce different notions of factorizations, which in general only exist for nonnegative matrices. We also extend these notions to symmetric decompositions (Section 2.3). Finally, we will relate these factorizations with decompositions of psd matrices on the bipartite tensor product space $\mathcal{M}_d \otimes \mathcal{M}_d$ (Section 2.4).

2.1 Minimal and nonnegative factorization

In the following, we introduce the notions of minimal and nonnegative factorizations, i.e. writing a given matrix as a matrix product of two other matrices whose size correspond to the rank of the given matrix. Recall that for a given $p \times q$ matrix $M \in \mathcal{M}_{p,q}$ the rank of the matrix M, denoted $\mathrm{rank}(M)$, is the number of linearly independent rows or columns, respectively. An equivalent characterization is the dimension of the range of linear map induced by the matrix M,

$$\mathrm{Im}(M) := \{Ax : x \in \mathbb{C}^q\}$$

In other words,

$$\dim(\mathrm{Im}(M)) = \mathrm{rank}(M).$$

Further recall that every matrix M admits a singular value decomposition. In particular, there exist isometries $U \in \mathcal{M}_{p,r}$, $V \in \mathcal{M}_{q,r}$ (i.e. $U^\dagger U = id, V^\dagger V = id$) with $r = \mathrm{rank}(M) \leq \min\{p, q\}$ and $\Sigma \in \mathcal{M}_r$ a diagonal matrix with *positive* entries such that

$$M = U \cdot \Sigma \cdot V^\dagger.$$

Proposition and Definition 1.12 (Minimal factorization [7]). *Let $M \in \mathcal{M}_{p,q}$. There exists $A \in \mathcal{M}_{p,r}$ and $B \in \mathcal{M}_{q,r}$ where $r = rank(M)$ such that*

$$M = A \cdot B^T.$$

The above factorization is called minimal factorization *of M.*

Proof. Let $M = U \cdot \Sigma \cdot V^\dagger$ a singular value decomposition of M. Setting

$$A = U \in \mathcal{M}_{p,r} \quad \text{and} \quad B = \left(\Sigma \cdot V^\dagger\right)^T \in \mathcal{M}_{q,r}$$

shows the statement. □

The definition of the minimal factorization as a product of two rectangular matrices can be used as a starting point to study factorizations with additional constraints. For example, in the following, we define the notions of nonnegative factorizations. In Section 2.3, we will study similar factorizations with additional constraints in symmetry.

Definition 1.13 (Nonnegative factorization)**.** *Let $M \in \mathcal{M}_{p,q}$. A* nonnegative factorization *of M is a minimal factorization*

$$M = A \cdot B^T$$

where $A \in \mathcal{M}_{p,r}$ and $B \in \mathcal{M}_{q,r}$ are nonnegative matrices, i.e. all entries are nonnegative. The minimal r attaining a nonnegative factorization is called nonnegative rank*, denoted*

$$\text{nn-rank(M)}.$$

Note that the above notions, minimal and nonnegative factorization, can be equivalently written as a decomposition in the form

$$M = \sum_{l=1}^{r} a_l \cdot b_l^T = \sum_{l=1}^{r} a_l \otimes b_l$$

where a_l and b_l are the lth column of A and B respectively and $a_l \otimes b_l := a_l \cdot b_l^T$ is a realization of the tensor product space $\mathbb{C}^p \otimes \mathbb{C}^q \cong \mathcal{M}_{p,q}$. This expression shows that the rank of a tensor product decomposition coincides in this particular case with the rank of a matrix. Furthermore, the upper decomposition is nonnegative if and only if all vectors a_l, b_l have nonnegative entries.

The minimal and nonnegative factorization can also be equivalently written as

$$M_{ij} = \widetilde{a}_i^T \cdot \widetilde{b}_j = \langle \widetilde{a}_i, \widetilde{b}_j \rangle \tag{1.12}$$

where both $\widetilde{a}_i$ and $\widetilde{b}_j \in \mathbb{C}^r$ are the ith and jth row of A and B respectively.

Note that a nonnegative decomposition exists if and only if M is nonnegative. In particular, for a nonnegative matrix M it holds that

$$\text{rank}(M) \leq \text{nn-rank}(M) \leq \min\{p, q\},$$

where the former inequality is trivial and the latter inequality holds due to the fact that $M = I_p \cdot M$ and $M = M \cdot I_q$ are valid nonnegative decompositions where I_p and I_q are the identity matrices of size p and q respectively. The first inequality can also be strict for nonnegative matrices. For example, the matrix

$$M = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

has $\text{rank}(M) = 3$ and $\text{nn-rank}(M) = 4$ [7]. We will discuss a further example in the context of separations in Theorem 2.16.

Note that the relation between minimal and nonnegative factorizations, and tensor decompositions on $\mathbb{C}^p \otimes \mathbb{C}^q$ as described in (i) also gives rise to an extension of decompositions to *nonnegative tensors* on the space $\mathbb{C}^d \otimes \cdots \otimes \mathbb{C}^d$. This extension will be studied in the context of decompositions on simplicial complexes in Chapter 2 more thoroughly.

2.2 Positive semidefinite factorization

In the following we want define a non-commutative extension of the notion of non-negative decompositions, namely the positive semidefinite decomposition. Note that for a nonnegative factorization, Relation (1.12) shows a representation of this decomposition as a scalar product on the cone $\mathbb{R}^d_+$. The positive semidefinite factorization is a generalization to the cone of real psd matrices together with the scalar product

$$\langle A, B\rangle := \mathrm{tr}\left(A \cdot B^T\right).$$

Nevertheless, we also allow in the following definition complex psd matrices.

Definition 1.14 (Positive semidefinite decomposition [21]). *Let $M \in \mathcal{M}_{p,q}$ be nonnegative. A* positive semidefinte definite (psd) factorization *of M is given by*

$$M_{ij} = \mathrm{tr}\left(E_i \cdot F_j^T\right)$$

where $E_i, F_j \in \mathcal{M}_r$ are psd matrices for $i \in \{1, \ldots, p\}$ and $j \in \{1, \ldots, q\}$. The minimal r which attains a psd factorization is called positive semidefinite rank of M, denoted

$$\text{psd-rank(M)}.$$

Note that for a nonnegative matrix $M \in \mathcal{M}_{p,q}$ the following relation holds [19]

$$\frac{1}{2}\sqrt{1 + 8\mathrm{rank}(M)} - \frac{1}{2} \leq \text{psd-rank}(M) \leq \text{nn-rank}(M) \tag{1.13}$$

where the second inequality is trivial because every nonnegative decomposition can be realized as a positive semidefinite decomposition. This is because the scalar product used for nonnegative vectors is equal to the scalar product of psd matrices restricted to diagonal matrices. Further note that in this definition the transpose makes no difference since positive semidefiniteness persists under transposition. We leave in a transpose for later convenience.

We end this section with a simple example from [19] which shows the relations between the different ranks.

Example 1.15. Consider the nonnegative matrix

$$M = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

It is clear that $\mathrm{rank}(M) = \text{nn-rank}(M) = 3$. This implies by Relation (1.13) that

$$2 \leq \text{psd-rank}(M) \leq 3.$$

A psd-factorization is given by

$$E_1 = F_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_2 = F_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, E_3 = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \text{ and } F_3 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Hence, $\text{psd-rank}(M) = 2$. △

2.3 Symmetric factorizations

In the following, we introduce notions similar to the minimal, nonnegative, and positive semidefinite factorization with an additional constraint in symmetry. These factorizations are defined only for square matrices due to obvious reasons in the definition.

Definition 1.16 (Symmetric factorizations)**.** *Let $M \in \mathcal{M}_d$ be a square matrix.*

(i) A symmetric factorization *of M is given by*

$$M = A \cdot A^T$$

where $A \in \mathcal{M}_{d,r}$. The minimal integer r attaining such a decomposition is called symmetric rank *of A, denoted*

$$\text{symm-rank(M)}.$$

(ii) A completely positive (cp) factorization *of M is given by a symmetric factorization*

$$M = A \cdot A^T$$

where $A \in \mathcal{M}_{d,r}$ is nonnegative*. The minimal r attaining such a decomposition is called* completely positive rank*, denoted*

$$\text{cp-rank(M)}.$$

(iii) A completely positive semidefinite transposed (cpsdt) factorization *of M is a factorization of the form*

$$M_{ij} = \text{tr}\left(E_i \cdot E_j^T\right)$$

where $E_i \in \mathcal{M}_r$ is psd for all $i \in \{1, \ldots, d\}$. The minimal r among all such decompositions is denoted

$$\text{cpsdt-rank(M)}.$$

A symmetric factorization of M exists if and only if M is symmetric (i.e. $M^T = M$). It can be computed for example by Takagi's factorization [25] which consists of a unitary $U \in \mathcal{M}_d$ (i.e. $U^\dagger U = UU^\dagger = id$) and a diagonal matrix Σ containing the nonnegative square roots of $M \cdot M^\dagger$ such that

$$M = U \cdot \Sigma \cdot U^T.$$

Obviously, $A = U \cdot \sqrt{\Sigma}$ is a valid symmetric factorization. By the definition of Σ it holds that

$$\text{symm-rank}(M) \leq d.$$

Note that this decomposition is very similar to a decomposition of psd matrices. The only difference is the third factor, which is here U^T instead of $U^\dagger$.

A nonnegative matrix M admits a cpsdt factorization if and only if M is symmetric. This fact can be shown using the later introduced relation (Theorem 1.18) between the cpsdt factorization and the symmetric purification form (Definition 1.17). For details, we refer to [14].

Further note that in the case of cpsdt factorization, the transposition in the second factor makes a difference. The factorization without transposition is called *completely positive semidefinite* (cpsd) factorization. In contrast to all other defined factorizations, very little is known about this factorization. For example, there does not even exist an upper bound of the cpsd-rank dependent on the dimension of the studied matrices [22].

Note that not every nonnegative square matrix M has a cp factorization. A necessary condition for the existence of a cp factorization is that M is positive semidefinite. One necessary and sufficient condition for M to have a cp factorization is the relation between nonnegative matrices and symmetric separable decompositions (Definition 1.17) given in Theorem 1.18. In the topic of completely positive factorizations, there are still many open questions concerning deciding membership or the geometry of the cp-cone [3].

2.4 Correspondence between decompositions and factorizations

In the last two parts, Section 1 and Section 2, we have defined different notions of decompositions in two different spaces. The former notions, namely the decompositions for psd matrices, are of interest in the study of quantum many-body systems and their effective descriptions. The latter, namely the factorizations of nonnegative matrices, are applied in the field of conic optimization and were studied a lot in recent years. In this part, we will introduce the notion of symmetric, bipartite decompositions of psd matrices, which are a symmetric analog to the MPDO with open boundary conditions (Remark 1.2). We will show a correspondence between these decompositions and the factorizations of nonnegative matrices. Theorem 1.18 will be extended to weighted simplicial complexes and proved in this general setting in Chapter 2.

Recall that a matrix product density operator with open boundary conditions (Remark 1.2) in the bipartite case has the shape

$$\rho = \sum_{\alpha=1}^{r} \rho_\alpha^{[0]} \otimes \rho_\alpha^{[1]}. \tag{1.14}$$

In particular, since the bipartite MPDO with open boundary conditions has only one index, it is equivalent to the trivial bipartite tensor decomposition. Moreover, similar to the t.i.-MPDO case, we define a symmetric analog to the tensor decomposition. In the language of (1.8) the symmetric tensor decomposition is an explicitly invariant decomposition for the full permutation group S_n. This leads to the following definition of the different notions of ranks as a symmetric tensor decomposition.

Definition 1.17 (Symmetric decomposition [8])**.** *Let* $\rho \in \mathcal{M}_d \otimes \cdots \otimes \mathcal{M}_d \cong \mathcal{M}_{d^n}$.

(i) A symmetric decomposition *is given by*

$$\rho = \sum_{\alpha=1}^{r} \rho_\alpha \otimes \rho_\alpha \otimes \cdots \otimes \rho_\alpha$$

(ii) A symmetric separable decomposition *is given by a symmetric decomposition with the additional property*

$$\rho_\alpha \geq 0 \quad \textit{for all } \alpha \in \{1, \ldots, r\}.$$

(iii) *A* symmetric local purification form *is given by* $\rho = \sigma^\dagger \sigma$, *where*

$$\sigma \in \mathcal{M}_{d',d} \otimes \cdots \otimes \mathcal{M}_{d',d}$$

admits a symmetric decomposition of the form

$$\sigma = \sum_{\alpha=1}^{r} \sigma_\alpha \otimes \sigma_\alpha \otimes \cdots \otimes \sigma_\alpha$$

where $\sigma_\alpha \in \mathcal{M}_{d',d}$.

Note that a symmetric tensor decomposition exists if and only if ρ is invariant with respect to arbitrary permutations of the tensor product spaces [8]. In other words, for all permutations $\pi : \{0, \ldots, n-1\} \to \{0, \ldots, n-1\}$ we have

$$\rho_{(i_0,\ldots,i_{n-1},j_0,\ldots,j_{n-1})} = \rho_{(i_{\pi(0)},\ldots,i_{\pi(n-1)},j_{\pi(0)},\ldots,j_{\pi(n-1)})}.$$

Note that the symmetric bipartite decomposition

$$\rho = \sum_{\alpha=1}^{r} \rho_\alpha \otimes \rho_\alpha$$

corresponds to a symmetric version of the bipartite MPDO-form with open boundary conditions (1.14) but is different than the bipartite t.i.-MPDO-form (1.9)

$$\rho = \sum_{\alpha_0,\alpha_1=1}^{r} \rho_{\alpha_0,\alpha_1} \otimes \rho_{\alpha_1,\alpha_0}.$$

In the following, assume ρ describes a bipartite system diagonal in a local basis, for example the computational basis, i.e.

$$\rho = \sum_{i=1}^{p} \sum_{j=1}^{q} m_{ij} |i,j\rangle\langle i,j| \in \mathcal{M}_p \otimes \mathcal{M}_q. \tag{1.15}$$

In other words, ρ describes a classically correlated quantum system. We define the corresponding $p \times q$ matrix by

$$M = \sum_{i=1}^{p} \sum_{j=1}^{q} m_{ij} |i\rangle\langle j| \in \mathcal{M}_{p,q}.$$

Note that M is nonnegative (i.e. entrywise nonnegative) if and only if ρ is psd.

The following theorem relates the decompositions of diagonal bipartite psd matrices ρ to the decompositions of the corresponding nonnegative matrices M. This implies that the corresponding ranks of these decompositions are equal. We will prove this correspondence in Chapter 2 in a more general setting (see Theorem 2.12 for details). For a proof of this particular case, we refer to [14].

Theorem 1.18. *Let* $\rho \in \mathcal{M}_p \otimes \mathcal{M}_q$ *be a bipartite, diagonal matrix of the form* (1.15) *and* $M \in \mathcal{M}_{p,q}$ *the corresponding nonnegative matrix. Then the following correspondences between decompositions of psd matrices and factorizations of nonnegative matrices are true:*

	Decomposition of ρ	*Factorization of M*
(i)	*MPDO-form with obc*	*minimal factorization*
(ii)	*separable MPDO-form with obc*	*nonnegative factorization*
(iii)	*local purification form with obc*	*psd factorization*
(iv)	*symmetric decomposition*	*symmetric factorization*
(v)	*symmetric separable decomposition*	*completely positive factorization*
(vi)	*symmetric local purification form*	*cpsdt factorization*

Note that above theorem is a special case of Theorem 2.12 presented below, which shows the correspondence between decompositions of psd matrices and nonnegative tensors on arbitrary simplicial complexes Ω with valid group actions G. Using Theorem 2.12, Theorem 1.18 is an immediate implication using the line of length 1, namely $\Omega = \Lambda_1$, and the group actions $G = \{id\} \subseteq S_2$ and $G = S_2$.

3 Nonnegative factorizations and correlation complexities

The study of factorizations of nonnegative matrices has many applications in different fields — for example, in semidefinite and linear programming or algebraic geometry [19]. In this section, we want to study the relation of factorizations to correlation complexities more thoroughly. Finally, we show the separation between the quantum correlation complexity and the classical correlation complexity.

Assume Alice and Bob want to generate random variables X and Y respectively based on a probability distribution P taking finitely many values. We denote the probability of Alice generating $X = i$ and Bob generating $Y = j$ by $P(X = i, Y = j)$. Hence the distribution of the random variable X taking values in $\{1, \dots p\}$ and Y taking values in $\{1, \dots, q\}$ can be represented by a *normalized* (i.e. all entries sum to 1), *nonnegative* matrix $M \in \mathcal{M}_{p,q}$ setting

$$M_{ij} := P(X = i, Y = j).$$

The condition

$$\sum_{i=1}^{p} \sum_{j=1}^{q} M_{ij} = 1$$

corresponds to the normalization of the probability distribution.

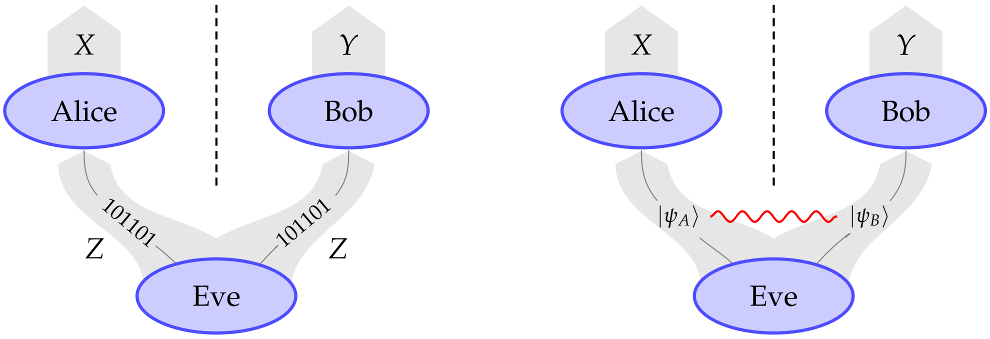
Assume, Alice and Bob are *spatially separated*, i.e. they cannot communicate and also cannot observe the generated random variable of the other party. This lack of communication implies that the generated random variables X and Y are independent, in other words

$$P(X = i, Y = j) = P(X = i) \cdot P(Y = j).$$

In the language of nonnegative matrices, this is equivalent to M being a rank-1 matrix,

$$M = v \cdot w^T = v \otimes w$$

where $v \in \mathbb{C}^p$ with $v_i = P(X = i)$ and $w \in \mathbb{C}^q$ with $w_j = P(Y = j)$.



(A) Realization of RCorr. Eve sends the same classical bit-string to Alice and Bob. X and Y are generated by probability distributions conditioned on the value of Z.

(B) Realization of QCorr. Eve generates a quantum state and sends one part to Alice and one part to Bob. Generation of X and Y by a positive operator valued measure (POVM).

FIGURE 1.5: Visualization of the quantum and classical correlation complexities.

The natural question arising in this discussion is the following: Is there a measure of correlation of P? We will see that the answer to this question is positive and is connected to the nonnegative and positive semidefinite rank of the matrix M.

This section is organized as follows: In Section 3.1 we will show a relation between the classical correlation complexity of P and the nonnegative rank of M. In Section 3.2 we will present a similar relation between the quantum correlation complexity of P and the positive semidefinite rank of M. Finally, we will show in Section 3.3 the appearance of a separation between these notions of correlation complexity, i.e. the former can be arbitrarily larger than the latter.

3.1 Interpretation of the nonnegative rank

Let us now extend this discussion to the configuration shown in Figure 1.5. Assume there is a third party, Eve, which generates a random variable Z taking values $\{1, \dots, r\}$ and sends the value of Z as a (classical) bit-string to both Alice and Bob. For a rigorous mathematical discussion, we need, therefore, the notion of conditional probability. The conditional probability distribution of X under condition Z is defined as

$$P(X = i | Z = l) := \frac{P(X = i, Z = l)}{P(Z = l)} \quad \text{for } i \in \{1, \dots, p\} \text{ and } l \in \{1, \dots r\}.$$

It describes the probability distribution of X if the result of Z is determined for the given probability distribution of the tuple (X, Z).

We are now able to classify all possible distribution functions P realized after fixing a strategy and a fixed value r of outcomes for Z, which limits the shared information.

Theorem 1.19 (Characterization of the nonnegative rank [7]). *Let X, Y be random variables taking values in $\{1, \dots p\}$ and $\{1, \dots, q\}$ respectively with joint probability distribution P. Let M be the corresponding nonnegative matrix. The following statements are equivalent:*

(i) There exists a rank-r nonnegative factorization

$$M = \sum_{l=1}^{r} v_l \cdot w_l^T$$

with $v_l \in \mathbb{R}_+^p$ *and* $w_l \in \mathbb{R}_+^q$.

(ii) The probability distribution P can be generated by a random variable Z mapping to $\{1, \dots, r\}$ *and* X, Y *independent random variables conditioned on Z, i.e.*

$$P(X = i, Y = j) = \sum_{l=1}^{r} P(X = i | Z = l) \cdot P(Y = j | Z = l) \cdot P(Z = l)$$

Proof. We will start with the easier direction.
(ii) $\Longrightarrow$ (i): We have to show that M attains a rank-r decomposition. Hence, define $(v_l)_i = P(X = i | Z = l) \cdot P(Z = l)$ and $(w_l)_j = P(Y = j | Z = l)$. By definition of the conditional probability we get the desired result

$$M_{ij} = P(X = i, Y = j) = \sum_{l=1}^{r} (v_l)_i \cdot (w_l)_j = \sum_{l=1}^{r} v_l \cdot w_l^T.$$

(i) $\Longrightarrow$ (ii): Let

$$M = \sum_{l=1}^{r} v_l \cdot w_l^T$$

be a nonnegative rank-r factorization. Define the probability distribution of the random variable Z as

$$P(Z = l) := \|v_l\|_{\ell_1} \cdot \|w_l\|_{\ell_1} \geq 0$$

where $l \in \{1, \dots, r\}$ and $\| \cdot \|_{\ell_1}$ is the sum of all entries. This is a valid probability distribution since

$$\begin{aligned} \sum_{l=1}^{r} P(Z = l) &= \sum_{l=1}^{r} \|v_l\|_{\ell_1} \cdot \|w_l\|_{\ell_1} = \sum_{i,j} \sum_{l=1}^{r} (v_l)_i \cdot (w_l)_j \\ &= \sum_{i,j} \sum_{l=1}^{r} \left(v_l \cdot w_l^T \right)_{ij} = \sum_{i,j} M_{ij} = 1. \end{aligned}$$

In the last step, we used the normalization of the probability matrix M_{ij}. Further, we define the normalized vectors

$$\widetilde{v}_l := \frac{v_l}{\|v_l\|_{\ell_1}} \quad \text{and} \quad \widetilde{w}_l := \frac{w_l}{\|w_l\|_{\ell_1}}$$

and define the conditional probability distribution as

$$P(X = i | Z = l) := (\widetilde{v}_l)_i \quad \text{and} \quad P(Y = j | Z = l) := (\widetilde{w}_l)_j$$

which are well defined by the normalization of the vectors $\tilde{v}_l$ and $\tilde{w}_l$ respectively. Finally, we obtain

$$\begin{aligned} P(X=i, Y=j) &= \sum_{l=1}^{r} (v_l)_i \cdot (w_l)_j \\ &= \sum_{l=1}^{r} P(X=i|Z=l) \cdot P(Y=j|Z=l) \cdot P(Z=l) \end{aligned}$$

which shows the statement. In particular, this shows that X and Y are independent when conditioned on Z. □

In the configuration of Figure 1.5, the length of the bit-string, which contains the realization of Z and is sent to Alice and Bob to generate a probability distribution P is called *random correlation complexity*, denoted

$$\mathsf{RCorr}(P).$$

Theorem 1.19 gives a direct characterization of the random correlation complexity showed in the following corollary.

Corollary 1.20. *Let X and Y be random variables with joint probability distribution P and corresponding nonnegative matrix M. Then the following relation holds:*

$$\mathit{RCorr}(P) = \left\lceil \log_2 \textit{nn-rank}(M) \right\rceil.$$

3.2 Interpretation of the positive semidefinite rank

In the following, we introduce a quantum analog to the random correlation complexity shown in Figure 1.5. We will first recall the notion of positive operator valued measures (POVM), the mathematical framework describing the measurement of quantum states. Recall that for a Hilbert space $\mathcal{H}$, the set of all bounded linear operators mapping $\mathcal{H}$ to $\mathcal{H}$ is denoted $\mathcal{B}(\mathcal{H})$. In the case of finite-dimensional Hilbert spaces $\mathcal{B}(\mathcal{H})$ corresponds to the set of all linear operators on $\mathcal{H}$.

Definition 1.21. *A* positive operator valued measure *(POVM) on a Hilbert space $\mathcal{H}$ is a family of positive semidefinite operators $E_i \in \mathcal{B}(\mathcal{H})$ where $i \in \{1, \ldots, k\}$ such that*

$$\sum_{i=1}^{k} E_i = id_{\mathcal{H}}.$$

Note that a POVM $\{E_i\}_{i=1}^k$ together with a state $|\psi\rangle \in \mathcal{H}$ gives rise to a probability distribution P, assigning every operator a different measurement result

$$P(X=i) := \mathrm{tr}(E_i |\psi\rangle\langle\psi|).$$

Further note that for given POVMs

$$\left\{E_i^{[0]}\right\}_{i=1}^{k} \quad \text{and} \quad \left\{E_i^{[1]}\right\}_{i=1}^{k}$$

on $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, the pairwise tensor product

$$\left\{E_i^{[0]} \otimes E_j^{[1]}\right\}_{i,j=1}^{k}$$

again defines a valid POVM on $\mathcal{H}_1 \otimes \mathcal{H}_2$. We call this POVM a *local POVM* since all measurements are done locally.

Assume that Eve can generate a quantum state in a Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ where $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = r$ instead of a probability distribution of integer values $\{1, \dots, r\}$. Assume also that the random variables X and Y generated by Alice and Bob arise from local measurements of $|\psi\rangle$. Hence the probability distribution corresponds to

$$P(X = i, Y = j) = \mathrm{tr}\big((F_i \otimes G_j)|\psi\rangle\langle\psi|\big)$$

where $\{F_i\}_{i=1}^p$ is the POVM generating the random variable X taking values in $\{1, \dots, p\}$ and $\{G_j\}_{j=1}^q$ the POVM generating the random variable Y taking values in $\{1, \dots, q\}$.

The following theorem characterizes the dimension r of the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ and the psd-rank of the nonnegative matrix induced by the joint probability distribution.

Theorem 1.22. *Let X and Y be random variables taking values in $\{1, \dots p\}$ and $\{1, \dots, q\}$, respectively, with joint probability distribution P. Let M be the corresponding nonnegative matrix. The following statements are equivalent:*

(i) M admits a psd factorization of size r.

(ii) There exists a state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ where $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = r$ and two POVMs $\{F_i\}_{i=1}^p$ and $\{G_j\}_{j=1}^q$ on $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively such that

$$P(X = i, Y = j) = \mathrm{tr}((F_i \otimes G_j)|\psi\rangle\langle\psi|)$$

Proof. We again start with the easier direction.

(ii) $\Longrightarrow$ (i): Let $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$. Using the Schmidt-decomposition (see [29] for details)

$$|\psi\rangle = \sum_{k=1}^{r} \lambda_k |u_k\rangle \otimes |v_k\rangle$$

where $\{|u_k\rangle\}_{k=1}^r$ and $\{|v_k\rangle\}_{k=1}^r$ are orthonormal bases of $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively and $\lambda_k \geq 0$, we define

$$U = \sum_{k=1}^{r} \sqrt{\lambda_k}|k\rangle\langle u_k| \quad \text{and} \quad V = \sum_{k=1}^{r} \sqrt{\lambda_k}|k\rangle\langle v_k|$$

where $\{|k\rangle\}_{k=0}^r$ is an orthonormal basis of $\mathbb{C}^r$. Using these matrices we obtain the generating psd matrices of the psd factorization by setting

$$A_i = U \cdot F_i \cdot U^\dagger \geq 0 \quad \text{and} \quad B_j = \left(V \cdot G_i \cdot V^\dagger\right)^T \geq 0$$

and obtain

$$\operatorname{tr}\left(A_i \cdot B_j^T\right) = \sum_{k,k'} \lambda_k \lambda_{k'} \langle u_k|F_i|u_{k'}\rangle \langle v_k|G_j|v_{k'}\rangle = \operatorname{tr}\big((F_i \otimes G_j)|\psi\rangle\langle\psi|\big)$$

(i) $\Longrightarrow$ (ii): Define

$$W_A = \sum_{i=1}^{p} A_i \quad \text{and } W_B = \sum_{i=1}^{q} B_j$$

Obviously, W_A and W_B are as a sum of psd matrices psd and without loss of generality, all eigenvalues of W_A and W_B are strictly positive. Else all A_i or all B_j would have a joint eigenspace with eigenvalue 0, i.e. there exists a positive semidefinite factorization with smaller a rank. We denote the square root of W_A by

$$(W_A)^{1/2} = \sum_{k=1}^{r} \sqrt{\lambda_k^{(A)}} \left|\psi_k^{(A)}\right\rangle \left\langle \psi_k^{(A)}\right|$$

where $\lambda_k^{(A)} > 0$ are the eigenvalues and $\left|\psi_k^{(A)}\right\rangle$ the eigenstates of W_A. Similarly we denote the square root of W_B by

$$(W_B)^{1/2} = \sum_{k=1}^{r} \sqrt{\lambda_k^{(B)}} \left|\psi_k^{(B)}\right\rangle \left\langle \psi_k^{(B)}\right|.$$

The definition

$$|\psi\rangle = (W_A^{1/2} \otimes W_B^{1/2}) \sum_{l=1}^{r} |l, l\rangle$$

yields a well defined state, since

$$\begin{aligned}
\langle\psi|\psi\rangle &= \sum_{k,l=1}^{r} \langle k,k|W_A \otimes W_B|l,l\rangle = \sum_{k,l=1}^{r} \langle k|W_A|l\rangle\langle l|W_B^T|k\rangle \\
&= \operatorname{tr}(W_A \cdot W_B^T) = \sum_{i=1}^{p}\sum_{j=1}^{q} \operatorname{tr}(A_i \cdot B_j^T) = \sum_{i,j=1}^{d} M_{ij} = 1
\end{aligned}$$

Further defining

$$F_i := (W_A)^{-1/2} \cdot A_i \cdot (W_A)^{-1/2} \geq 0 \quad \text{and} \quad G_j := (W_B)^{-1/2} \cdot B_j \cdot (W_B)^{-1/2} \geq 0$$

yield two valid POVMs $\{F_i\}_{i=1}^{p}$ and $\{G_j\}_{j=1}^{q}$ with the desired property

$$\begin{aligned}
\operatorname{tr}((F_i \otimes G_j)|\psi\rangle\langle\psi|) &= \langle\psi|F_i \otimes G_j|\psi\rangle = \sum_{k,l=1}^{r} \langle k,k|A_i \otimes B_j|l,l\rangle \\
&= \operatorname{tr}\left(A_i \cdot B_j^T\right) = M_{ij}
\end{aligned}$$

□

Similar to the random correlation complexity $\mathsf{RCorr}(P)$ we also can define for this configuration a measure of correlation. We define the *quantum correlation complexity* $\mathsf{QCorr}(P)$ as the minimum number of qubits Eve has to send to each, Bob and Alice to generate the joint probability distribution P in the configuration of Figure 1.5 (B). In other words, $2^{\mathsf{QCorr}(P)}$ indicates the dimension of the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. Theorem 1.22 again immediately characterizes this complexity measure.

Corollary 1.23. *Let X, Y a random variables with joint probability distribution P and corresponding nonnegative matrix M. Then the following relation holds:*

$$QCorr(P) = \left\lceil \log_2 psd\text{-}rank(M) \right\rceil.$$

3.3 Separation of correlation complexities

The characterizations of RCorr and QCorr in Corollary 1.20 and Corollary 1.23 have several broad implications. Note that for all probability distributions P, Relation (1.13) implies that

$$\mathsf{QCorr}(P) \leq \mathsf{RCorr}(P),$$

hence the quantum method in this configuration is always more efficient. Besides, for nn-rank and psd-rank there appears a separation, i.e. there exists *no* function $f : \mathbb{N} \to \mathbb{N}$ such that

$$\text{nn-rank}(M) \leq f(\text{psd-rank}(M)) \quad \text{for all } d \in \mathbb{N} \text{ and } M \in \mathcal{M}_d \text{ nonnegative.}$$

In other words, the nonnegative rank cannot be upper bounded by the positive semidefinite rank alone – there is also a dependence in the dimension of the matrix M. This can be deduced with the following example, and further implies a separation between $\mathsf{QCorr}(P)$ and $\mathsf{RCorr}(P)$.

Example 1.24 (Separation of QCorr and RCorr)**.** Let $M_d \in \mathcal{M}_d$ be the Euclidean distance matrix, that is, we define

$$(M_d)_{ij} = (i-j)^2 \quad \text{for all } i, j \in \{1, \dots, d\}.$$

It was shown in [19] that

$$\text{psd-rank}(M_d) = 2 \quad \text{and} \quad \text{nn-rank}(M_d) \geq \log_2(d).$$

Define the joint probability distribution of X and Y corresponding to a normalized version of M_d as

$$P(X = i, Y = j) = \frac{6 \cdot (i-j)^2}{d^2 \cdot (d^2 - 1)} \quad \text{for } i, j \in \{1, \dots, d\}.$$

Using the characterizations of Corollary 1.20 and Corollary 1.23 immediately implies for the joint probability distribution P

$$\mathsf{QCorr}(P) = 1 \quad \text{and} \quad \mathsf{RCorr}(P) \geq \left\lceil \log_2 \log_2 d \right\rceil$$

△

The above example implies that for this class of probability distributions, the number of classical bits which has to be sent to Alice and Bob respectively depends on the dimension parameter d. In contrast, each probability distribution in this class can be generated by only sending one qubit to each party.

Chapter 2

Generalization of Decompositions to Weighted Simplicial Complexes

In this chapter, we will present an extension of tensor decompositions based on *weighted simplicial complexes* (wsc) introduced in [12]. This framework relates the summation indices to facets of the wsc. Roughly speaking, a weighted simplicial complex is a generalization of a graph which additionally allows:

(i) Multiple, identical connections;

(ii) Connections between more than two vertices (called *facets*); and

(iii) All connections have a weight

A graph only allowing extension (i) is called *multigraph*, and graphs allowing (i) and (ii) are called *multihypergraph* [4]. In principle, the extension of tensor decompositions can also be formulated for multihypergraphs, but the notion of wsc is easier to define and work with.

We will also present the notions of symmetries in tensor decompositions with group actions G on a wsc, which will be denoted Ω. These group actions give rise to a generalization of translational invariant decompositions discussed in Section 1.3.

In this chapter, we will present a brief overview of the construction of the framework of weighted simplicial complexes and group actions (Section 1), and then we will introduce the notion of (Ω, G)-*decompositions* together with the notions of (Ω, G)-*ranks* (Section 2). We will in particular discuss the application to psd matrices (Section 2.2) and nonnegative tensors (Section 2.3). Finally, we will study the relations of the different ranks in Section 3.2.

1 Weighted simplicial complexes and group actions

The goal of this section is to present the notions of weighted simplicial complexes (wsc) Ω and group actions G on Ω. In the beginning, we briefly introduce the notation used in the following sections. Later, we will see some simple examples to motivate the definition of wsc. For each index $i \in [n] := \{0, \ldots, n\}$, we fix a $\mathbb{C}$-vector space $\mathcal{V}_i$ (called the *local vector space*). We denote the *global vector space* as the tensor product space

$$\mathcal{V} := \mathcal{V}_0 \otimes \cdots \otimes \mathcal{V}_n.$$

By definition every $v \in \mathcal{V}$ can be expressed as a sum of elementary tensors

$$v^{[0]} \otimes \cdots \otimes v^{[n]}.$$

Similar to Chapter 1, we are interested in v as a *tensor network*, hence we define decompositions with different arrangements of the summation indices. We will provide some relevant and straightforward examples in Example 2.2.

Further, if v is invariant with respect to permutations of the indices $[n]$ of the elementary tensors, we want to find explicitly invariant decompositions, that is, decompositions whose elementary tensors themselves are invariant under the permutation of the indices $[n]$. This can be formalized by a group action.

Definition 2.1. *Let G be a group and X be a finite set. A group action G on X is a mapping*

$$G \times X \to X : (g, x) \mapsto g \cdot x$$

with the properties

(i) $\forall x \in X : e \cdot x = x$ *(where e is the neutral element of G)*
(ii) $\forall x \in X, \forall g, h \in G : (gh) \cdot x = g \cdot (h \cdot x)$

We will write in the following gx instead of $g \cdot x$ for simplicity. Note that for a fixed $g \in G$ the mapping

$$X \to X : x \mapsto gx$$

is bijective. Hence, for $X = [n]$ we can identify G as a subgroup of the symmetric group

$$S_n := \{\pi : [n] \to [n] : \pi \text{ bijective}\}$$

together with the concatenation as the multiplication.

In relation to symmetric tensor decompositions, we consider for a given group action G on the set $[n]$ the induced linear group action on $\mathcal{V}$, i.e.

$$g : v^{[0]} \otimes \cdots \otimes v^{[n]} \mapsto v^{[g0]} \otimes \cdots \otimes v^{[gn]}$$

for $g \in G$. An element $v \in \mathcal{V}$ is called *G-invariant* if it is invariant under the action of G on $\mathcal{V}$. The subspace of invariant elements is denoted $\mathcal{V}_{\text{inv}}$. We are now able to revisit some examples, already encountered in Chapter 1, with this notation.

Example 2.2. (i) Consider a decomposition of the form

$$v = \sum_{\alpha=1}^{r} v_\alpha^{[0]} \otimes \cdots \otimes v_\alpha^{[n]}.$$

The minimal number r of elementary tensors is called the *tensor rank* of v. If we set $\mathcal{I} = \{1, \ldots, r\}$ and $\mathcal{F} = \{\{0, \ldots, n\}\}$ we can equivalently write the decomposition as a sum over all possible functions $\alpha : \mathcal{F} \to \mathcal{I}$, denoted $\alpha \in \mathcal{I}^{\mathcal{F}}$, in particular

$$v = \sum_{\alpha \in \mathcal{I}^{\mathcal{F}}} v_{\alpha_{|0}}^{[0]} \otimes \cdots \otimes v_{\alpha_{|n}}^{[n]} \tag{2.1}$$

where $\alpha_{|i}$ denotes the restriction of α to the set $\mathcal{F}_i = \{F \in \mathcal{F} : i \in F\}$. Intuitively, every element $F \in \mathcal{F}$ corresponds to a summation index α_i, and elements of F correspond to the positions where α_i appears.

(ii) Let G be a transitive group action on $[n]$, i.e. there is only one orbit, $Gi = [n]$ for all $i \in [n]$. For an element v which is invariant under the group action G we consider the G-invariant decomposition

$$v = \sum_{\alpha=1}^{r} v_\alpha \otimes \cdots \otimes v_\alpha$$

which is called *symmetric tensor decomposition*. The smallest number r among all possible decompositions is called *symmetric tensor rank*.

(iii) Consider a decomposition of the form

$$v = \sum_{\alpha_0,\ldots,\alpha_{n-1}=1}^{r} v^{[0]}_{\alpha_0} \otimes v^{[1]}_{\alpha_0,\alpha_1} \otimes \cdots \otimes v^{[n-1]}_{\alpha_{n-2},\alpha_{n-1}} \otimes v^{[n]}_{\alpha_{n-1}}.$$

Such a decomposition is called *matrix product operator form* and the minimal r among all possible decompositions the *operator Schmidt rank*. By choosing

$$\mathcal{F} = \Big\{\{0,1\},\{1,2\},\ldots,\{n-1,n\}\Big\}$$

and $\mathcal{I} = \{1,\ldots,r\}$ this corresponds to the decomposition of Equation (2.1).

(iv) Consider a decomposition of the form

$$v = \sum_{\alpha_0,\ldots,\alpha_{n-1}=1}^{r} v^{[0]}_{\alpha_0,\alpha_1} \otimes v^{[1]}_{\alpha_1,\alpha_2} \otimes \cdots \otimes v^{[n-2]}_{\alpha_{n-2},\alpha_{n-1}} \otimes v^{[n-1]}_{\alpha_{n-1},\alpha_0}.$$

This is similar to the decomposition in (iii), but with periodic boundary conditions. Hence, this decomposition corresponds to a choice of

$$\mathcal{F} = \Big\{\{i,i+1\} : i \in \{0,\ldots,n-2\}\Big\} \cup \Big\{\{n-1,0\}\Big\}$$

in the form of Equation (2.1). Additionally considering the symmetry operation given by the cyclic group $G = C_n$ (i.e. the group generated by a mapping $i \mapsto i+1$, where addition is modulo n) we obtain the *t.i. matrix product operator form* [14]

$$v = \sum_{\alpha_0,\ldots,\alpha_{n-1}=1}^{r} v_{\alpha_0,\alpha_1} \otimes v_{\alpha_1,\alpha_2} \otimes \cdots \otimes v_{\alpha_{n-2},\alpha_{n-1}} \otimes v_{\alpha_{n-1},\alpha_0}.$$

△

In the following, we will define the notion of *weighted simplicial complexes* (wsc) and relate them to the sets $\mathcal{F}$ in Example 2.2. Relation (2.1), together with the definition of wsc will be the critical parts of (Ω, G)-decompositions.

1.1 Weighted simplicial complexes

In this section, we give a rigorous definition of the notions of weighted simplicial complexes (wsc) Ω based on [12]. This concept will formalize various arrangements of indices in one consistent definition.

Recall that we write $\mathcal{P}_n$ for the power set $\mathcal{P}([n])$ (the set of all subsets of $[n]$, which has 2^{n+1} elements). For a more detailed discussion of the properties of wsc, we refer to [10].

In the following, we denote the set of all natural numbers including 0 by $\mathbb{N}_0$.

Definition 2.3. *(i) A* weighted simplicial complex (wsc) on $[n]$ *is a function*

$$\Omega : \mathcal{P}_n \to \mathbb{N}_0$$

such that $S_1 \subseteq S_2 \in \mathcal{P}_n$ implies that $\Omega(S_1)$ divides $\Omega(S_2)$.

(ii) A set $S \in \mathcal{P}_n$ is called simplex of Ω *if $\Omega(S) \neq 0$. In the following, we will assume that each singleton $\{i\} \in \mathcal{P}_n$ is a simplex, and call every element $i \in [n]$ a* vertex *of the wsc. A maximal simplex (with respect to inclusion) is called a* facet of Ω. *The set of all facets is denoted*

$$\mathcal{F} := \{F \in \mathcal{P}_n : F \text{ facet of } \Omega\}$$

and the set of all facets which contain the vertex i is denoted

$$\mathcal{F}_i := \{F \in \mathcal{F} : i \in F\}.$$

The restriction of Ω to $\mathcal{F}$ gives rise to a multiset, called $\widetilde{\mathcal{F}}$, which contains $F \in \mathcal{F}$ precisely $\Omega(F)$-times. We define the multiset $\widetilde{\mathcal{F}}_i$ for $i \in [n]$ analogously. There exists a canonical collapse map

$$c : \widetilde{\mathcal{F}} \to \mathcal{F} \quad \text{and} \quad c : \widetilde{\mathcal{F}}_i \to \mathcal{F}_i$$

mapping all copies of a facet to the underlying facet.

(iii) Two vertices i, j are neighbors *if*

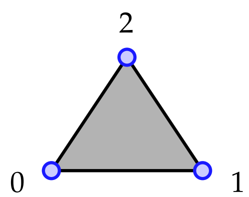
$$\mathcal{F}_i \cap \mathcal{F}_j \neq \{\}.$$

Further, two vertices $i, j \in [n]$ are connected *if there exists a sequence of vertices $i_0, i_1, \ldots, i_k$ such that i_m and i_{m+1} are neighbors for all $m \in [k-1]$ and $i = i_0$ and $j = i_k$.*

Note that a wsc whose image is $\{0, 1\}$ is called simplicial complex (sc). In contrast to its weighted counterpart, an sc cannot contain multiple, identical facets. Hence, a simplicial complex is a special case of a hypergraph containing single facets.

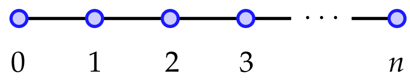
In the following, we give some straightforward examples of simplicial complexes and relate their facets to the sets $\mathcal{F}$ constructed in Example 2.2.

Example 2.4. (i) The simplicial complex (sc) Σ_n that maps each subset of $[n]$ to 1 is called the *n-simplex*. For $n = 2$ this can be depicted as follows:



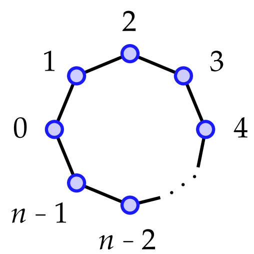
It only has one (multi-)facet, i.e. $\mathcal{F} = \widetilde{\mathcal{F}} = \{[n]\}$ and hence gives rise to the tensor decomposition of Example 2.2 (i) and (ii).

(ii) For $n \geq 1$, the *line of length* n (i.e. composed of $n+1$ points) is the sc Λ_n corresponding to the following graph:



The set $\mathcal{F} = \widetilde{\mathcal{F}}$ has n elements. In particular, it coincides with $\mathcal{F}$ in Example 2.2 (iii) and hence generates a matrix product operator form. Intuitively, the sc shows the connections between the local vector spaces through shared summation indices.

(iii) For $n \geq 3$, the *circle of length* n is the sc Θ_n corresponding to the following graph:



It has n facets and generates the decomposition given in Example 2.2 (iv).

△

1.2 Group actions

The next step is the definition of group actions on weighted simplicial complexes. Intuitively, these are group actions which respect the geometry of the wsc. First, we will introduce the notions of G-linearity and G-invariance, which are necessary for the classification of valid group actions on weighted simplicial complexes.

Definition 2.5. *Let G be a group action on the sets X,Y. A function $f : X \to Y$ is called* G-linear *if*

$$f(gx) = gf(x)$$

for all $x \in X$ and $g \in G$. If g acts trivially on Y, we call f G-invariant.

We say that the group action of G on X is *free*, if $\mathrm{Stab}(x) = \{e\}$ for all $x \in X$, where

$$\mathrm{Stab}(x) := \{g \in G : gx = x\}.$$

To define a group action on Ω we first consider a group action of G on the set $[n]$. Every group action on $[n]$ canonically induces a group action on $\mathcal{P}_n$. Further, if Ω is G-invariant, G also induces a group action on $\mathcal{F}$, since for $F \in \mathcal{F}$, $g \in G$ and $F \subsetneq S$ holds $\Omega(gS) = \Omega(S) = 0$. These properties motivate the following definition.

Definition 2.6. *A group action G on the wsc Ω consists of the following:*

(i) An action G on $[n]$ such that Ω is G-invariant with respect to the induced action on $\mathcal{P}_n$. This induces an action of G on $\mathcal{F}$.

(ii) An action of G on $\widetilde{\mathcal{F}}$, such that the collapse map

$$c : \widetilde{\mathcal{F}} \to \mathcal{F}$$

is G-linear (we also say the action of G on $\widetilde{\mathcal{F}}$ refines the action of G on $\mathcal{F}$).

Further, an action of G on the wsc Ω is called free *if the action of G on $\widetilde{\mathcal{F}}$ is free.*

Note that in order to obtain a group action on a wsc, one has to provide additional information compared to a group action on $\mathcal{F}$, namely, one must specify how elements $g \in G$ permute the different copies of facets in the multiset when passing from a facet F to the facet gF. It is shown in [12, Prop. 7] that a group action on $\mathcal{F}$ can always be refined to a free group action on $\widetilde{\mathcal{F}}$.

In the following, we always associate a group action G as a group action on a wsc Ω, that is, fulfilling Definition 2.6.

2 Invariant tensor decompositions and ranks

We are now ready to introduce the different types of (Ω, G)-decompositions and their corresponding ranks. In Section 2.1, we will define the (Ω, G)-decomposition without any additional restrictions to the elements in the local vector spaces. Later on, we will use this notion in the space of matrices and the space of tensors. This will be the basis for other decompositions in the cone of psd matrices (Section 2.2) and nonnegative tensors (Section 2.3), where additional constraints characterize various decompositions.

From now on, we fix a connected wsc Ω and a free group action G on Ω.

2.1 Decompositions in general vector spaces

We now present the notion of (Ω, G)-decomposition and its corresponding (Ω, G)-rank. Recall that

$$\mathcal{V} := \mathcal{V}_0 \otimes \cdots \otimes \mathcal{V}_n.$$

The local dimensions $d_i := \dim(\mathcal{V}_i)$ can be chosen differently as long as there are no further restrictions given by the group action of G on $\mathcal{V}$; that is, whenever i, j are in the same orbit of the group action of G on $[n]$, then $d_i = d_j$. The construction in this definition is based on the observations in Example 2.2.

Definition 2.7. *For $v \in \mathcal{V}$, an (Ω, G)*-decomposition *is given by a finite index set $\mathcal{I}$ and families*

$$V^{[i]} := \left(v_\beta^{[i]}\right)_{\beta \in \mathcal{I}^{\widetilde{\mathcal{F}}_i}},$$

where $v_\beta^{[i]} \in \mathcal{V}_i$ for $i \in [n]$ such that:

(a) We have

$$v = \sum_{\alpha \in \mathcal{I}^{\widetilde{\mathcal{F}}}} v_{\alpha_{|0}}^{[0]} \otimes v_{\alpha_{|1}}^{[1]} \otimes \cdots \otimes v_{\alpha_{|n}}^{[n]}. \tag{2.2}$$

(b) For all $i \in [n]$, $g \in G$ and $\beta \in \mathcal{I}^{\widetilde{\mathcal{F}}_i}$ it holds

$$v_\beta^{[i]} = v_{{}^g\beta}^{[gi]},$$

where ${}^g\beta \in \mathcal{I}^{\widetilde{\mathcal{F}}_{gi}}$ and ${}^g\beta(F) := \beta(g^{-1}F)$ for $F \in \widetilde{\mathcal{F}}_{gi}$.

The smallest cardinality of $\mathcal{I}$ among all possible (Ω, G)-decompositions of v is called the (Ω, G)-rank of v, *denoted*

$$\mathrm{rank}_{(\Omega,G)}(v).$$

If no (Ω, G)-decomposition of v exists, we set $\mathrm{rank}_{(\Omega,G)}(v) := \infty$.

The mapping ${}^g\beta$ is well defined due to the fact that G is a group action on Ω (in particular, it extends to $\widetilde{\mathcal{F}}$ and maps facets to facets). Note that if G is a free group action on a connected wsc Ω, and $v \in \mathcal{V}_{\mathrm{inv}}$, there always exists an (Ω, G)-decomposition of v [12, Thm. 13].

For simplicity, if G is the trivial group, we will call an (Ω, G)-decomposition just an Ω-decomposition, and write rank_Ω for the rank. The same simplification will also be used for all ranks defined in the following two subsections.

2.2 Decompositions in the cone of psd matrices

Separability (or its negation, entanglement), and purifications are central notions in quantum information theory. In the next two definitions, we will formulate these notions in the framework of (Ω, G)-decompositions.

We will assume that the local vector space is given by

$$\mathcal{V}_i := \mathcal{M}_{d_i}$$

and hence

$$\mathcal{V} := \mathcal{M}_{d_0} \otimes \cdots \otimes \mathcal{M}_{d_n} \cong \mathcal{M}_{d_0 \cdots d_n},$$

whose hermitian part is denoted by $\mathrm{Her}_{d_0} \otimes \cdots \otimes \mathrm{Her}_{d_n} \cong \mathrm{Her}_{d_0 \cdots d_n}$. The d_i can again be chosen differently as long as there are no further restrictions given by the group action of G. Further, we will define the *cone of psd matrices* as $\mathcal{M}^+_{d_0 \cdots d_n}$. If $\rho \in \mathcal{M}^+_{d_0 \cdots d_n}$ fulfills $\mathrm{tr}(\rho) = 1$, we call it a *state*.

Let us now define (Ω, G)-purifications and (Ω, G)-square root decompositions.

Definition 2.8. *Let $\rho \in \mathcal{M}^+_{d_0 \cdots d_n}$ be a psd matrix.*

(i) An (Ω, G)-purification *is an element*

$$\sigma \in \mathcal{M}_{d'_0, d_0} \otimes \cdots \otimes \mathcal{M}_{d'_n, d_n}$$

with

$$\rho = \sigma^\dagger \sigma \ \text{ and } \ \mathrm{rank}_{(\Omega,G)}(\sigma) < \infty,$$

where $\mathcal{M}_{d'_i, d_i}$ denotes the space of all complex $d'_i \times d_i$ matrices and † the hermitian conjugate. The smallest (Ω, G)-rank among all (Ω, G)-purifications is called (Ω, G)-purification rank of ρ, *denoted*

$$\text{puri-rank}_{(\Omega,G)}(\rho).$$

(ii) $\sigma \in Her_{d_0} \otimes \cdots \otimes Her_{d_n}$ *is called* square root of ρ *if* $\sigma^2 = \rho$. *We call the smallest* (Ω, G)*-rank among all square roots of* ρ *the* (Ω, G)-quantum square root rank of ρ *, denoted*

$$\text{q-sqrt-rank}_{(\Omega,G)}(\rho).$$

Remark 2.9. (i) Note that not every matrix σ which fulfills $\sigma^2 = \rho$ is automatically hermitian. Using the spectral decomposition of ρ setting $\rho = UDU^\dagger$ with $D = \text{diag}(\lambda_1, \lambda_2, \ldots)$, one can see that all its hermitian square roots are of the form [24]

$$\sigma = UD^{1/2}U^\dagger, \quad D^{1/2} = \text{diag}\Big(\pm\sqrt{\lambda_1}, \pm\sqrt{\lambda_2}, \ldots\Big).$$

(ii) If G is a free group action on a connected wsc Ω and $\rho \in \mathcal{V}_{\text{inv}}$ psd, there always exists an (Ω, G)-purification of ρ and a square root with a finite (Ω, G)-rank (see Theorem 27 in [12] for details). △

The next step is the definition of the separable (Ω, G)-rank. We call the matrix $\rho \in \mathcal{M}^+_{d_0\cdots d_n}$ *separable* if it admits a decomposition

$$\rho = \sum_j \rho_j^{[0]} \otimes \cdots \otimes \rho_j^{[n]}$$

where $\rho_j^{[i]} \in \mathcal{M}^+_{d_i}$. Equivalently, the set of all separable matrices for fixed local dimensions $d_0, \ldots, d_n$ is given by

$$\mathcal{M}^+_{d_0} \otimes \cdots \otimes \mathcal{M}^+_{d_n}.$$

If additionally $\text{tr}(\rho) = 1$, we call ρ a *separable state*. From now on, we will denote the set of separable states

$$\text{SEP}_{d_0,d_1,\ldots,d_n} := \{\rho \in \mathcal{M}^+_{d_0\cdots d_n} : \rho \text{ separable state}\}.$$

If a state $\rho \in \mathcal{M}^+_{d_0\cdots d_n}$ is not contained in $\text{SEP}_{d_0,d_1,\ldots,d_n}$, we call it *entangled*. Further, if $d_i = d$ for all $i \in [n]$, we will write for simplicity $\text{SEP}_{n,d} := \text{SEP}_{d_0,\ldots,d_n}$.

Definition 2.10. *A* separable (Ω, G)-decomposition of $\rho \in \mathcal{M}_{d_0\cdots d_n}$ *is given by an* (Ω, G)*-decomposition*

$$\rho = \sum_{\alpha \in \mathcal{I}^{\tilde{\mathcal{F}}}} \rho^{[0]}_{\alpha_{|0}} \otimes \rho^{[1]}_{\alpha_{|1}} \otimes \cdots \otimes \rho^{[n]}_{\alpha_{|n}}$$

in which $\rho^{[i]}_\beta \in \mathcal{M}^+_{d_i}$ *for all* $\beta \in \mathcal{I}^{\tilde{\mathcal{F}}_i}$ *and* $i \in [n]$. *The smallest cardinality of an index set* $\mathcal{I}$ *among all possible separable* (Ω, G)*-decompositions of* ρ *is called the* separable (Ω, G)-rank of ρ, *denoted*

$$\text{sep-rank}_{(\Omega,G)}(\rho).$$

If there exists no separable (Ω, G)*-decomposition of* v*, we set* $\text{sep-rank}_{(\Omega,G)}(v) := \infty$.

Note that if G is a free group action on a connected wsc Ω and $\rho \in \mathcal{V}_{\text{inv}}$ is separable, there always exists a separable (Ω, G)-decomposition of ρ [12, Thm. 21].

2.3 Decompositions in the cone of nonnegative tensors

In the following we will consider the set of nonnegative tensors and define different notions of (Ω, G)-ranks based on [12, Sec. 5]. Before presenting the definitions of different decompositions and ranks we will introduce the necessary notation.

For simplicity we consider the local space $\mathcal{V}_i = \mathbb{C}^d$ and define the global space $\mathcal{V} := \mathcal{K}_{n,d}$, where

$$\mathcal{K}_{n,d} := \bigotimes_{i=0}^{n} \mathbb{C}^d.$$

If n and d are clear from the context, we write $\mathcal{K}$ instead of $\mathcal{K}_{n,d}$. Any element $M \in \mathcal{K}$ can be uniquely written as

$$M = \sum_{i_0,\dots,i_n} m_{i_0,\dots,i_n} e_{i_0} \otimes \cdots \otimes e_{i_n}$$

where e_j denotes the j-th standard basis vector in the corresponding vector space $\mathbb{C}^d$. M is said to be *nonnegative* if $m_{i_0,\dots,i_n} \geq 0$ for all $i_0, \dots, i_n$.

We now give a brief definition of notions of different (Ω, G)-decompositions and the corresponding ranks. For a more detailed discussion we refer to [12].

Definition 2.11. *Let $M \in \mathcal{K}_{n,d}$ be a tensor.*

(i) A nonnegative (Ω, G)-decomposition of *M is an (Ω, G)-decomposition (Equation (2.2)) where all $v^{[i]}_{\alpha_{|i}} \in \mathbb{C}^d$ have nonnegative entries. The corresponding rank is called the* nonnegative (Ω, G)-rank of *M, denoted*

$$\text{nn-rank}_{(\Omega,G)}(M).$$

(ii) A positive semidefinite (Ω, G)-decomposition of *M consists of psd matrices*

$$E^{[i]}_j \in \mathcal{M}^+_{k_i}$$

where $k_i = \left|\mathcal{I}^{\widetilde{\mathcal{F}}_i}\right|$ for $i \in [n]$ and $j \in \{1, \dots, d\}$, such that

$$\left(E^{[i]}_j\right)_{\beta,\beta'} = \left(E^{[gi]}_j\right)_{{}^g\beta,{}^g\beta'}$$

for all i, g, j, β, β', where ${}^g\beta(F) := \beta(g^{-1}F)$ for $F \in \widetilde{\mathcal{F}}_{gi}$, and

$$m_{i_0,\dots,i_n} = \sum_{\alpha,\alpha' \in \mathcal{I}^{\widetilde{\mathcal{F}}}} \left(E^{[0]}_{i_0}\right)_{\alpha_{|0},\alpha'_{|0}} \cdots \left(E^{[n]}_{i_n}\right)_{\alpha_{|n},\alpha'_{|n}}$$

for all $i_0, \dots, i_n$. The smallest cardinality of an index set $\mathcal{I}$ among all possible positive semidefinite (Ω, G)-decompositions is called the positive semidefinite (Ω, G)-rank of *M, denoted*

$$\text{psd-rank}_{(\Omega,G)}(M).$$

(iii) $N \in \mathcal{K}_{n,d}$ with $n_{i_0,\dots,i_n} \in \mathbb{R}$ is called a square root of *M, if $M = N \circ N$, where $M = N \circ N$ denotes the Hadamard-product (i.e. entrywise multiplication, $m_{i_0,\dots,i_n} = n^2_{i_0,\dots,i_n}$). The smallest (Ω, G)-rank among all square roots of M is called* (Ω, G)-square root rank of

M, denoted

$$\text{sqrt-rank}_{(\Omega,G)}(M).$$

Using the realization of the tensor product $v \otimes w := v \cdot w^T$ where $v \in \mathbb{C}^{d_0}$ and $w \in \mathbb{C}^{d_1}$, these notions of ranks are a natural extension of the ranks defined in Section 2 of Chapter 1. For the nonnegative decomposition this is an obvious fact. For the positive semidefinite decomposition we consider $\Omega = \Lambda_1$ (a line of length 1, equivalently Θ_1 or Σ_1, see Example 2.4 for details) and G the trivial group. Then $\mathcal{F} = \widetilde{\mathcal{F}} = \{\{0,1\}\}$ and hence

$$\begin{aligned} m_{i_0,i_1} &= \sum_{\alpha,\alpha' \in \mathcal{I}^{\mathcal{F}}} \left(E_{i_0}^{[0]}\right)_{\alpha_{|0},\alpha'_{|0}} \cdot \left(E_{i_1}^{[1]}\right)_{\alpha_{|1},\alpha'_{|1}} \\ &= \sum_{\alpha,\alpha'=1}^{r} \left(E_{i_0}^{[0]}\right)_{\alpha,\alpha'} \cdot \left(E_{i_1}^{[1]}\right)_{\alpha,\alpha'} = \mathrm{tr}\left(\left(E_{i_0}^{[0]}\right) \cdot \left(E_{i_1}^{[1]}\right)^T\right) \end{aligned}$$

which corresponds to the nonnegative rank defined in Section 2.2 of Chapter 1.

3 Relations between the different ranks

In this section we will prove relations between the different ranks. First, we will show a correspondence between nonnegative tensors and diagonal psd matrices. Using this correspondence we will relate the different notions of tensor ranks with the different matrix ranks.

We will also study inequalities between different matrix ranks and similar relations between tensor ranks (Section 3.2).

Then we will turn to the question of separations between different ranks, i.e. if some ranks can be arbitrarily larger than others. This will be studied in Section 3.3. As we will later see, one of the main results of Chapter 3 is the disappearance of these separations in the approximate case.

3.1 Correspondence between nonnegative tensors and psd matrices

Here, we present the correspondence between nonnegative tensors and psd matrices for arbitrary (Ω, G)-decompositions [12].

Recall that every tensor $M \in \mathcal{K}_{n,d}$, written as

$$M = \sum_{i_0,\ldots,i_n} m_{i_0,\ldots,i_n} e_{i_0} \otimes \cdots \otimes e_{i_n}$$

can be associated with a diagonal matrix $\sigma \in \mathcal{M}_d \otimes \cdots \otimes \mathcal{M}_d \cong \mathcal{M}_{d^{n+1}}$ by setting

$$\sigma = \sum_{i_0\ldots,i_n} m_{i_0,\ldots,i_n} E_{i_0 i_0} \otimes \cdots \otimes E_{i_n i_n}, \tag{2.3}$$

where E_{jk} is the $d \times d$ matrix which has value 1 on position (j,k) and 0 elsewhere. Obviously σ is psd if and only if M is nonnegative.

Using this correspondence, we can relate the notions of ranks for nonnegative tensors with the ranks of psd matrices, as shown in the following Theorem based on [12].

Theorem 2.12. *Let Ω be a connected wsc and G a group action on Ω. Further let $M \in \mathcal{K}_{n,d}$ and $\sigma \in \mathcal{M}_{d^{n+1}}$ the corresponding diagonal matrix. Then the following holds:*

(i)	$\text{rank}_{(\Omega,G)}(M)$	$=$	$\text{rank}_{(\Omega,G)}(\sigma)$
(ii)	$\text{nn-rank}_{(\Omega,G)}(M)$	$=$	$\text{sep-rank}_{(\Omega,G)}(\sigma)$
(iii)	$\text{psd-rank}_{(\Omega,G)}(M)$	$=$	$\text{puri-rank}_{(\Omega,G)}(\sigma)$
(iv)	$\text{sqrt-rank}_{(\Omega,G)}(M)$	$=$	$\text{q-sqrt-rank}_{(\Omega,G)}(\sigma)$

Proof. We show each point separately.

(i) We start with an (Ω, G)-decomposition

$$M = \sum_{\alpha \in \mathcal{I}^{\widetilde{\mathcal{F}}}} m^{[0]}_{\alpha_{|0}} \otimes \cdots \otimes m^{[n]}_{\alpha_{|n}}$$

where $m^{[i]}_\beta \in \mathbb{C}^d$ for all $\beta \in \mathcal{I}^{\widetilde{\mathcal{F}}_i}$ where $i \in \{0, \cdots, n\}$. Define the local diagonal matrix

$$\sigma^{[i]}_\beta := \text{diag}(m^{[i]}_\beta) \in \mathcal{M}_d. \tag{2.4}$$

By the choice of the tensor product it is immediate that

$$\sigma = \sum_{\alpha \in \mathcal{I}^{\widetilde{\mathcal{F}}}} \sigma^{[0]}_{\alpha_{|0}} \otimes \cdots \otimes \sigma^{[n]}_{\alpha_{|n}} \in \mathcal{M}_{d^{n+1}}$$

is diagonal and equal to the related matrix in Equation (2.3). Conversely for a given diagonal matrix σ, the nonnegative tensor can be obtained again by Relation (2.4).

(ii) is similar to (i) using the fact that M is nonnegative if and only if σ is psd.

(iii) Let $E^{[i]}_j \in \mathcal{M}^+_{k_i}$ be a psd matrix given from a psd (Ω, G)-decomposition of M as defined in Definition 2.11 (ii). Then there exists a matrix $A^{[i]}_j \in \mathcal{M}_{k_i}$ such that

$$E^{[i]}_j = \left(A^{[i]}_j\right)^\dagger \cdot \left(A^{[i]}_j\right)$$

Such a matrix can be obtained for example by computing a hermitian square root $\sqrt{E^{[i]}_j}$. Denoting $a^{[i]}_{j,\beta} \in \mathbb{C}^{\mathcal{I}^{\widetilde{\mathcal{F}}_i}}$ as the column β of $A^{[i]}_j$, we can write for $\beta, \beta' \in \mathcal{I}^{\widetilde{\mathcal{F}}_i}$

$$\left(E^{[i]}_j\right)_{\beta,\beta'} = \left(a^{[i]}_{j,\beta}\right)^\dagger \cdot \left(a^{[i]}_{j,\beta'}\right).$$

Further, we obtain from the symmetry properties of $E^{[i]}_j$ that

$$a^{[gi]}_{j,g\beta} = a^{[i]}_{j,\beta}$$

for all $g \in G$. Defining

$$\xi^{[i]}_\beta := \sum_{j=1}^{d} a^{[i]}_{j,\beta} \otimes E_{jj}$$

yields a valid (Ω, G)-purification since

$$\begin{aligned}\sigma &= \xi^\dagger \xi = \sum_{\alpha,\alpha' \in \mathcal{I}^{\widetilde{\mathcal{F}}}} \left((\xi^{[0]}_{\alpha_{|0}})^\dagger \cdot (\xi^{[0]}_{\alpha'_{|0}}) \otimes \cdots \otimes (\xi^{[n]}_{\alpha_{|n}})^\dagger \cdot (\xi^{[n]}_{\alpha'_{|n}}) \right) \\ &= \sum_{\alpha,\alpha' \in \mathcal{I}^{\widetilde{\mathcal{F}}}} \sum_{j_0,\dots,j_n=1}^{d} \left(E^{[0]}_{j_0}\right)_{\alpha_{|0}\alpha'_{|0}} \cdots \left(E^{[n]}_{j_n}\right)_{\alpha_{|n}\alpha'_{|n}} E_{j_0 j_0} \otimes \cdots \otimes E_{j_n j_n},\end{aligned}$$

where we have used that $E_{jj} \cdot E_{ii} = \delta_{i,j} \cdot E_{jj}$. Here, $\delta_{i,j}$ denotes the Kronecker delta (in particular, $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ else). Conversely, let

$$\xi \in \mathcal{M}_{d'_0,d_0} \otimes \cdots \otimes \mathcal{M}_{d'_n,d_n}$$

be a (Ω, G)-purification of σ. We denote the local matrices of the (Ω, G)-decomposition of ξ by $\xi^{[i]}_\beta$ for $\beta \in \mathcal{I}^{\widetilde{\mathcal{F}}_i}$ and $i \in \{0, \dots, n\}$ and define $E^{[i]}_j \in \mathcal{M}_{k_i}$ by

$$\left(E^{[i]}_j\right)_{\beta,\beta'} := \left(\left(\xi^{[i]}_\beta\right)^\dagger \cdot \left(\xi^{[i]}_{\beta'}\right) \right)_{jj}$$

A similar calculation as above shows that this definition provides a positive semidefinite (Ω, G)-decomposition of M.

(iv) Let $N \in \mathcal{K}_{n,d}$ be a square-root of the tensor M, i.e.

$$M = N \circ N$$

and $n^{[i]}_\beta \in \mathbb{C}^d$ the local vectors realizing a (Ω, G)-decomposition of N. Hence, we can write

$$M = \sum_{\alpha,\gamma \in \mathcal{I}^{\widetilde{\mathcal{F}}}} \left(n^{[0]}_{\alpha_{|0}} \cdot n^{[0]}_{\gamma_{|0}} \right) \otimes \cdots \otimes \left(n^{[n]}_{\alpha_{|n}} \cdot n^{[n]}_{\gamma_{|n}} \right)$$

By defining

$$\xi^{[i]}_\beta := \sum_{j=1}^{d} \left(n^{[i]}_\beta\right)_j \cdot E_{jj}$$

we obtain a valid (Ω, G)-quantum-square-root decomposition, since

$$\begin{aligned}\xi^2 &= \sum_{\alpha,\gamma \in \mathcal{I}^{\widetilde{\mathcal{F}}}} \sum_{j_0,\dots,j_n=1}^{d} \left(n^{[0]}_{\alpha_{|0}}\right)_{j_0} \cdot \left(n^{[0]}_{\gamma_{|0}}\right)_{j_0} \cdots \left(n^{[n]}_{\alpha_{|n}}\right)_{j_n} \cdot \left(n^{[n]}_{\gamma_{|n}}\right)_{j_n} E_{j_0 j_0} \otimes \cdots \otimes E_{j_n j_n} \\ &= \sum_{j_0,\cdots,j_n=1}^{d} m_{j_0,\cdots,j_n} E_{j_0 j_0} \otimes \cdots \otimes E_{j_n j_n} = \sigma.\end{aligned}$$

The other direction can be showed by backwards calculation and the fact that the local matrices $\xi^{[i]}_\beta$ of the diagonal matrix $\xi \in \mathcal{M}_{d_0 \cdots d_n}$ are again diagonal.

□

3.2 Inequalities

In the following, we show several relations between the (Ω, G)-ranks of matrices. To prove these inequalities, we construct for a given decomposition (specified through the larger rank) another decomposition with the same rank. This immediately shows that the latter decomposition has a smaller rank.

Theorem 2.13. *For each $\rho \in \mathcal{M}_{d_0} \otimes \cdots \otimes \mathcal{M}_{d_n}$ we have*

(i)	$\text{rank}_{(\Omega,G)}(\rho)$	$\leq$	$\text{sep-rank}_{(\Omega,G)}(\rho)$
(ii)	$\text{puri-rank}_{(\Omega,G)}(\rho)$	$\leq$	$\text{sep-rank}_{(\Omega,G)}(\rho)$
(iii)	$\text{rank}_{(\Omega,G)}(\rho)$	$\leq$	$\text{puri-rank}_{(\Omega,G)}(\rho)^2$
(iv)	$\text{puri-rank}_{(\Omega,G)}(\rho)$	$\leq$	$\text{q-sqrt-rank}_{(\Omega,G)}(\rho)$

Proof. (i) Since every separable (Ω, G)-decomposition is a (Ω, G)-decomposition this statement is obvious.

(ii) Let $\mathcal{I}$ be an index set which realizes a separable (Ω, G)-decomposition

$$\rho = \sum_{\alpha \in \mathcal{I}^{\tilde{\mathcal{F}}}} \rho^{[0]}_{\alpha_{|0}} \otimes \cdots \otimes \rho^{[n]}_{\alpha_{|n}}$$

Since $\rho^{[i]}_\beta$ is psd there exists a hermitian square root

$$\tau^{[i]}_\beta := \sqrt{\rho^{[i]}_\beta}$$

for $\beta \in \mathcal{I}^{\tilde{\mathcal{F}}_i}$ and $i \in \{0, \dots n\}$. Consider the matrix

$$\xi^{[i]}_\beta := \left(\tau^{[i]}_\gamma \cdot \delta_{\gamma,\beta}\right)_{\beta,d} \in \mathcal{M}_{k_i,d}$$

where $k_i = \left|\mathcal{I}^{\tilde{\mathcal{F}}_i}\right|$ and $\delta_{\gamma,\beta}$ is the Kronecker delta. By definition of $\tau^{[i]}_\beta$ we have

$$\xi^{[gi]}_{g\beta} = \xi^{[i]}_\beta.$$

Hence these local matrices provide a valid (Ω, G)-decomposition of an element

$$\xi \in \mathcal{M}_{k_0,d_0} \otimes \cdots \otimes \mathcal{M}_{k_n,d_n}.$$

with $\text{rank}_{(\Omega,G)}(\xi) \leq |\mathcal{I}|$. Since

$$\left(\xi^{[i]}_\beta\right)^\dagger \cdot \left(\xi^{[i]}_\gamma\right) = \delta_{\beta,\gamma} \cdot \rho^{[i]}_\beta$$

we obtain

$$\begin{aligned} \xi^\dagger \cdot \xi &= \sum_{\alpha,\gamma \in \mathcal{I}^{\tilde{\mathcal{F}}}} \left(\xi^{[0]}_{\alpha_{|0}}\right)^\dagger \cdot \left(\xi^{[0]}_{\gamma_{|0}}\right) \otimes \cdots \otimes \left(\xi^{[n]}_{\alpha_{|n}}\right)^\dagger \cdot \left(\xi^{[n]}_{\gamma_{|n}}\right) \\ &= \sum_{\alpha \in \mathcal{I}^{\tilde{\mathcal{F}}}} \rho^{[0]}_{\alpha_{|0}} \otimes \cdots \otimes \rho^{[n]}_{\alpha_{|n}} = \rho. \end{aligned}$$

(iii) First we will show that the (Ω, G)-rank of the product of two matrices is smaller or equal than the product of the two individual (Ω, G)-ranks. Therefore, let $\sigma, \tau \in \mathcal{M}_{d_0 \ldots d_n}$. If σ attains a (Ω, G)-decomposition on an index set $\mathcal{I}$ and τ on an index set $\mathcal{J}$ then $\sigma \cdot \tau$ attains an (Ω, G)-decomposition on $\mathcal{L} := \mathcal{I} \times \mathcal{J}$. More precisely, consider the two projections $p_1 : \mathcal{L} \to \mathcal{I}$ and $p_2 : \mathcal{L} \to \mathcal{J}$ and define

$$\zeta_\beta^{[i]} := \sigma_{p_1 \circ \beta}^{[i]} \cdot \tau_{p_2 \circ \beta}^{[i]}$$

for $\beta \in \mathcal{L}^{\widetilde{\mathcal{F}}_i}$ and $i \in \{0, \ldots, n\}$. These tensors provide a (Ω, G)-decomposition of $\sigma \cdot \tau$. Hence

$$\text{rank}_{(\Omega,G)}(\sigma \cdot \tau) \leq \text{rank}_{(\Omega,G)}(\sigma) \cdot \text{rank}_{(\Omega,G)}(\tau).$$

Now let $\rho \in \mathcal{M}_{d_0 \ldots d_n}$ with (Ω, G)-purification ξ. Then we obtain

$$\begin{aligned} \text{rank}_{(\Omega,G)}(\rho) &= \text{rank}_{(\Omega,G)}(\xi^\dagger \cdot \xi) \\ &\leq \text{rank}_{(\Omega,G)}(\xi^\dagger) \cdot \text{rank}_{(\Omega,G)}(\xi) = \text{puri-rank}_{(\Omega,G)}(\rho)^2 \end{aligned}$$

(iv) Clear, since every (Ω, G)-square-root is an (Ω, G)-purification.

□

Using the correspondence given in Equation (2.3) and its implications in Theorem 2.12, we can easily prove a similar result to Theorem 2.13.

Corollary 2.14. *Let $M \in \mathcal{K}_{n,d}$. Then the following is true:*

(i)	$\textit{rank}_{(\Omega,G)}(M)$	$\leq$	$\textit{nn-rank}_{(\Omega,G)}(M)$
(ii)	$\textit{psd-rank}_{(\Omega,G)}(M)$	$\leq$	$\textit{nn-rank}_{(\Omega,G)}(M)$
(iii)	$\textit{rank}_{(\Omega,G)}(M)$	$\leq$	$\textit{psd-rank}_{(\Omega,G)}(M)^2$
(iv)	$\textit{psd-rank}_{(\Omega,G)}(M)$	$\leq$	$\textit{sqrt-rank}_{(\Omega,G)}(M)$

Proof. We only prove the first point, the other statements (ii)-(iv) are similar.

(i) Let σ be the diagonal matrix corresponding to M given in Equation (2.3). Then we obtain using Theorem 2.12 and Theorem 2.13

$$\text{rank}_{(\Omega,G)}(M) = \text{rank}_{(\Omega,G)}(\sigma) \leq \text{sep-rank}_{(\Omega,G)}(\sigma) = \text{nn-rank}_{(\Omega,G)}(M)$$

□

3.3 Separations of the ranks

In Section 3.2 we have seen that some ranks can be upper bounded by others (Theorem 2.13 and 2.14). A natural question is whether these can also be related by a function depending on the other ranks. To explain this concretely with the ranks given in Theorem 2.13 (i) let us consider the following problem.

Question 2.15. *Given a wsc Ω and a group action G on Ω. Does there exist a function $F : \mathbb{N} \to \mathbb{N}$ such that*

$$\textit{sep-rank}_{(\Omega,G)}(\rho) \leq F\Big(\textit{rank}_{(\Omega,G)}(\rho)\Big) \quad \textit{for all } \rho \in \textit{SEP}_{d_0,d_1,\ldots,d_n}$$

Obviously we have to restrict this question to the set of separable states $\mathrm{SEP}_{d_0,d_1,\dots,d_n}$, since $\text{sep-rank}_{(\Omega,G)}(\rho) = \infty$ for non-separable matrices. Nevertheless, the answer to this question is negative. Similar results appear for other relations as we will show in the following theorem. We say two ranks show a separation, denoted by $\ll$, if one rank cannot be upper bounded by a function only depending on the other rank.

Theorem 2.16. *Let $\Omega = \Lambda_1$ the line of length one (i.e. it contains two vertices) and G be the trivial group.*

(a) The following separations are true for the cone of nonnegative matrices:

$$\begin{array}{llcl} (i) & \text{rank}_{\Lambda_1} & \ll & \text{psd-rank}_{\Lambda_1} \\ (ii) & \text{psd-rank}_{\Lambda_1} & \ll & \text{sqrt-rank}_{\Lambda_1} \\ (iii) & \text{nn-rank}_{\Lambda_1} & \ll & \text{sqrt-rank}_{\Lambda_1} \\ (iv) & \text{sqrt-rank}_{\Lambda_1} & \ll & \text{nn-rank}_{\Lambda_1} \end{array}$$

(b) The following separations are true for psd matrices:

$$\begin{array}{llcl} (i) & \text{rank}_{\Lambda_1} & \ll & \text{puri-rank}_{\Lambda_1} \\ (ii) & \text{puri-rank}_{\Lambda_1} & \ll & \text{q-sqrt-rank}_{\Lambda_1} \end{array}$$

(c) The following separations are true for separable states:

$$\begin{array}{llcl} (i) & \text{puri-rank}_{\Lambda_1} & \ll & \text{sep-rank}_{\Lambda_1} \end{array}$$

Note that $\ll$ is a transitive operation, i.e. $A \ll B$ and $B \ll C$ implies $A \ll C$.

Proof. (a) We show the separations using results from [19].

(i) Let $S_d \in \mathcal{M}_d \cong \mathcal{K}_{2,d}$ be the slack matrix of a d-gon in the plane. Then it holds

$$\text{rank}_{\Lambda_1}(S_d) = 3 \quad \text{for all } d \in \mathbb{N}$$

and

$$\lim_{d\to\infty} \text{nn-rank}_{\Lambda_1}(S_d) = \infty.$$

(ii)-(iii) Let $n_1, n_2, \dots$ be a sequence of integers such that $2n_i - 1$ is prime for all $i \in \mathbb{N}$. We define the prime matrix $P_d \in \mathcal{M}_d \cong \mathcal{K}_{2,d}$ by

$$(P_d)_{i,j} = n_i + n_j - 1$$

Then it holds

$$\text{nn-rank}_{\Lambda_1}(P_d) = \text{psd-rank}_{\Lambda_1}(P_d) = 2 \quad \text{for all } d \in \mathcal{M}_d$$

and

$$\text{sqrt-rank}_{\Lambda_1}(P_d) = d \quad \text{for all } d \in \mathbb{N}$$

(iv) Let $M_d \in \mathcal{M}_d \cong \mathcal{K}_{2,d}$ be the Euclidean distance matrix, i.e.

$$(M_d)_{i,j} = (i-j)^2 \quad \text{for } i,j \in \{1,\dots,d\}$$

Then it holds

$$\text{rank}_{\Lambda_1}(M_d) = \text{psd-rank}_{\Lambda_1}(M_d) = \text{sqrt-rank}_{\Lambda_1}(M_d) = 2 \quad \text{for all } d \in \mathbb{N}$$

and

$$\text{nn-rank}_{\Lambda_1}(M_d) \geq \log_2(d) \quad \text{for all } d \in \mathbb{N}.$$

(b) This is immediate due to Theorem 2.12 and (a). □

This discussion can be extended to more sophisticated geometries. A construction of a state which shows a separation between rank and puri-rank on Λ_n is given in [15].

In the next chapter, we will study these separations from a new perspective. More precisely, we will show that the separations will disappear in the approximate case for both psd matrices (Corollary 3.21) and nonnegative tensors (Corollary 3.25).

Chapter 3

Approximate Decompositions on Weighted Simplicial Complexes

In Chapter 1 we have seen that most ranks can be upper bounded by a function of the dimension. For example, for a nonnegative matrix $M \in \mathcal{M}_d$, we always have

$$\operatorname{rank}(M) \leq \text{nn-rank}(M) \leq d.$$

One reason for finding such upper bounds, for example in the case of the matrix rank or matrix nonnegative rank, is the fact that M is a linear operator on a d-dimensional vector space. Hence, M can always be considered by its action on linear independent vectors spanning in the worst case the whole space, which obviously has at most d elements. This argument holds as long as we can decompose the matrix or tensor into arbitrary elements of a vector space, but often it cannot be extended when asking for additional properties of the elements in the decompositions like the local certificate of positivity.

One structure all spaces with a notion of positivity have in common is the conic structure of the set of all positive elements. In the case of nonnegative tensors, we know that for a nonnegative tensor T, the tensor $\lambda \cdot T$ is again nonnegative for all $\lambda \geq 0$. This particular structure also implies convexity, in particular, for nonnegative tensors T, S the convex combination $\lambda \cdot T + (1-\lambda) \cdot S$ is also nonnegative for $0 \leq \lambda \leq 1$.

Hence the set of all nonnegative tensors can be written as a convex combination of a (possibly infinite) generating set. In this case, the generating set can be chosen as the infinite set of summands in the decompositions, namely the set of elementary tensors. By definition of nonnegative tensors, it immediately holds that

$$\mathcal{K}^+_{n,d} = \operatorname{conv}\Big(\Big\{v^{[0]} \otimes \cdots \otimes v^{[n-1]} : v^{[i]} \in \big(\mathbb{R}^+\big)^d \quad \text{for all } i \in \{0, \ldots, n-1\}\Big\}\Big)$$

where $\mathbb{R}^+ := [0, \infty)$.

Analogously, the set of all n-partite separable states can be characterized by convex combinations of product states, that is,

$$\mathrm{SEP}_{n,d} = \operatorname{conv}\Big(\Big\{\rho^{[0]} \otimes \cdots \otimes \rho^{[n-1]} : \rho^{[i]} \in \mathcal{M}^+_d \text{ state} \quad \text{for all } i \in \{0, \ldots, n-1\}\Big\}\Big).$$

This equality holds directly by the definition of separable states, namely, $\rho \in \mathrm{SEP}_{n,d}$ if there exist states $\rho^{[i]}_\alpha \in \mathcal{M}^+_d$ for $i \in \{0, \ldots, n-1\}$, $\alpha \in \{1, \ldots, r\}$ and positive real numbers λ_α summing to 1 such that

$$\rho = \sum_{\alpha=1}^{r} \lambda_\alpha \cdot \rho_\alpha^{[0]} \otimes \cdots \otimes \rho_\alpha^{[n-1]}.$$

The integer r is an upper bound of the tensor rank (Definition 2.11) and can also give upper bounds to various (Ω, G)-ranks.

The famous Carathéodory theorem (see for example [37]) characterizes the maximal number r (dependent on the ambient space) necessary to write any element of a convex set as a convex combination of elements in the generator set. Hence, it is of interest in the study of ranks.

Theorem 3.1 (Carathéodory). *Let $\mathcal{V}$ be a d-dimensional vector space and $S \subseteq \mathcal{V}$ a set. Further let $a \in conv(S)$ be an element of the convex hull generated by S. Then there exist $s_0, \ldots, s_d \in S$ and $\lambda_0, \ldots, \lambda_d \geq 0$ nonnegative real numbers summing to 1 such that*

$$a = \sum_{i=0}^{d} \lambda_i s_i.$$

In other words, Carathéodory's theorem guarantees that every element can be written as a convex combination of at most $d+1$ generating points (i.e. elements of the set S). In the above mentioned examples, the application of Theorem 3.1 implies that every element in $\text{SEP}_{n,d}$ can be written as a convex combination of at most $d^{2n}+1$ product psd matrices and every element in $\mathcal{K}_{n,d}$ can be written as a convex combination of at most d^n+1 nonnegative elementary tensors. It follows that

$$\text{nn-rank}_{\Sigma_n}(M) \leq d^{2n} + 1$$

for all nonnegative tensors in $M \in \mathcal{K}_{n,d}$ as shown in [14].

The results implied from Carathéodory's theorem can be used for upper bounds of (Ω, G)-ranks which are, in general, not optimal. For example, the upper bounds of the ranks of matrix-factorizations shown in Section 2 of Chapter 1 are smaller than the upper bounds one would obtain with Carathéodory's theorem. Nonetheless, we want to go in another direction concerning approximate decompositions and their ranks. Specifically, we show in Section 1 an approximate version of Theorem 3.1, which, roughly speaking, says that there always exists an element ε-close to the original element which can be written as a convex combination of a number of generating elements *independent* of the dimension of the ambient space.

In Section 2, we will define the notions of approximate ranks based on (Ω, G)-ranks of Chapter 2. Afterwards, we will apply the approximate version of Carathéodory's theorem to show that all notions of approximate (Ω, G)-ranks can be upper bounded independently of the system dimension. These results have a considerable impact on the study of separations in the approximate case: We will show in Section 3 that many separations disappear in the approximate case. To end this chapter, we provide in Section 4 an algorithm to compute an approximate decomposition fulfilling the dimension independent upper bounds.

This chapter is mainly a review of paper [13].

1 The approximate Carathéodory Theorem

In this section, we introduce an approximate Carathéodory Theorem for *uniformly smooth Banach spaces* to finally derive a version applicable to *Schatten p-classes* (i.e. matrix spaces equipped with the Schatten p-norm) and ℓ_p*-spaces*, which will be a central tool in this chapter. The main result of this section is given in Theorem 3.9 and will be used later in Section 2.

Before starting with the derivation of this theorem, we will introduce some necessary definitions and notions of approximation.

Recall that $\mathcal{M}_d$ denotes the set of complex $d \times d$ matrices. Further $\|\cdot\|_p$ denotes the unnormalized Schatten p-norm, in particular, for $A \in \mathcal{M}_d$ we have

$$\|A\|_p := \operatorname{tr}(|A|^p)^{1/p} = \left(\sum_{i=1}^{d} s_i(A)^p\right)^{1/p}, \tag{3.1}$$

where $|A| := \sqrt{A^\dagger A}$ is the psd square root of $A^\dagger A$. Moreover, $\{s_i(A)\}_{i=1}^d$ denotes the set of singular values of A. Note that the definition in (3.1) is a norm for $1 \leq p < \infty$. Note that this norm can be written for A hermitian as

$$\|A\|_p = \left(\sum_{i=1}^{d} |\lambda_i(A)|^p\right)^{1/p}$$

where $\{\lambda_i(A)\}_{i=1}^d$ denotes the set of all eigenvalues of A. The Schatten p-norm has a broad spectrum of applications for different values of p as we will see in the following example.

Example 3.2. Let $A \in \mathcal{M}_d$ be a hermitian matrix.

(i) The Schatten 1-norm

$$\|A\|_1 = \sum_{i=1}^{d} |\lambda_i(A)|$$

is called the trace-norm and can be used to characterize the set of density matrices as all psd matrices in the unit ball with respect to $\|\cdot\|_1$.

(ii) The Schatten 2-norm can be equivalently written as

$$\|A\|_2 = \sqrt{\operatorname{tr}(A^\dagger A)}.$$

It is the norm induced by the Hilbert-Schmidt inner product $\langle A, B\rangle := \operatorname{tr}(A^\dagger B)$. Note that the Schatten 2-norm is the only Schatten norm induced by an inner product.

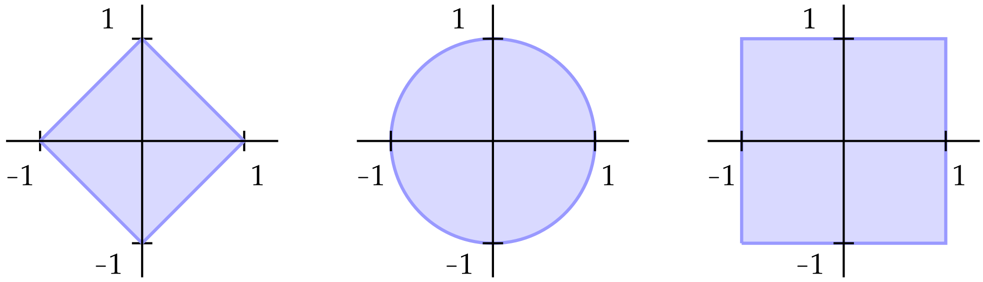
(iii) The Schatten ∞-norm

$$\|A\|_\infty = \lim_{p\to\infty} \|A\|_p = \max_i |\lambda_i(A)|$$

is equal to the operator norm of A and hence important in the study of C^**-algebras*.

The visualization of the different unitballs concerning the corresponding eigenvalues in the case of $d = 2$ is shown in Figure 3.1. △

Before showing a version of the approximate Carathéodory Theorem for matrix



(A) $p = 1$ (B) $p = 2$ (C) $p = \infty$

FIGURE 3.1: Unit balls of the eigenvalues in the space of all hermitian 2×2 matrices for the different values of p studied in Example 3.2.

spaces equipped with the Schatten p-norm, we will start with a more general version of the approximate Carathéodory Theorem valid for a larger class of Banach spaces, namely the class of uniformly smooth Banach spaces. This class can be characterized, roughly speaking, as the class of all vector spaces equipped with a differentiable norm.

1.1 On uniformly smooth Banach spaces

We start with a version of the approximate Carathéodory Theorem [26], which holds for uniformly smooth Banach spaces. We start with the definition of the notion of uniformly smooth. Also, we introduce the *modulus of smoothness*, which will be necessary later on.

Definition 3.3. *Let X with $\|\cdot\|$ be a Banach space. The* modulus of smoothness *$\rho_X : [0,\infty] \to [0,\infty]$ is given by*

$$\rho_X(t) := \sup\left\{\frac{1}{2}\Big(\|x+ty\| + \|x-ty\|\Big) - 1 : \|x\| = \|y\| \leq 1\right\} \text{ for } t \in [0,\infty].$$

A Banach space is called uniformly smooth *if $\rho_X(t) = o(t)$, i.e. $\rho_X(t)/t \to 0$ as $t \to 0$.*

Note that a finite-dimensional Banach space X with $\|\cdot\|$ is uniformly smooth if and only if for all $x, y \in X$ with $\|x\| = \|y\| = 1$ the limit

$$\lim_{t\to 0} \frac{\|x+ty\| - \|x\|}{t}$$

exists. In other words, for all $x \neq 0$ all directional derivatives of the norm exist [28]. This implies that the unit ball of uniformly smooth spaces is smooth.

Note that ρ_X is convex, strictly increasing and $\rho_X(0) = 0$. The first property is an immediate consequence of the triangle inequality applied to $\|\cdot\|$. The second property follows from the fact that $\rho_X(\varepsilon) > 0$ for $\varepsilon > 0$ and the convexity of ρ_X. These three characteristics imply that there exists the inverse function ρ_X^{-1}.

Theorem 3.4 (Approximate Carathéodory [26]). *Let S be a bounded set in a uniformly smooth Banach space X equipped with the norm $\|\cdot\|$, and $a \in \mathrm{conv}(S)$. Then there exists a* *sequence* $\{x_i\}_{i=1}^{\infty} \subseteq S$ *such that for* $a_k = \frac{1}{k}\sum_{i=1}^{k} x_i$ *the following inequality holds*

$$\|a - a_k\| \leq \frac{2\exp(2)}{k \cdot \rho_X^{-1}(1/k)} \cdot \mathrm{diam}(S)$$

where $\rho_X(\cdot)$ *is the modulus of smoothness.*

Note that Theorem 3.4 gives an upper bound for the approximation error independent of a, and the only dependence on the space X is given through the diameter of S and the inverse of the modulus of smoothness. We will see that for the particular case of $X = \mathcal{M}_d$ equipped with the Schatten p-norm the upper bound is also dimension independent.

1.2 On Schatten-classes

In the following, we denote the modulus of smoothness of $\mathcal{M}_d$ with Schatten p-norm by $\rho_p(\cdot)$. We now want to evaluate an upper bound for the expression $1/\rho_p^{-1}(1/k)$ following [26, 27] to finally give a version of the approximate Carathéodory theorem valid for Schatten p-classes.

A necessary tool for the following calculation are *Hanner's inequalities*.

Theorem 3.5 (Hanner's inequalities for Schatten norms [1])**.** *Let* $A, B \in \mathcal{M}_d$*. For* $4 \leq p < \infty$ *the following inequality holds:*

$$\Big(\|A\|_p + \|B\|_p\Big)^p + \Big|\|A\|_p - \|B\|_p\Big|^p \geq \|A+B\|_p^p + \|A-B\|_p^p.$$

For $1 \leq p \leq 4/3$ *the inequality is reversed and for* $p = 2$ *equality holds.*

Note that this theorem is an extension of the original Hanner's inequalities for ℓ_p-spaces, i.e. spaces $X = \mathbb{C}^d$ equipped with the entrywise p-norm [27]. It is widely believed that Theorem 3.5 is valid for $1 < p < \infty$ similar to the ℓ_p case. Nonetheless, it is only proven on this restricted range of p for all matrices. Before showing upper bounds for $\rho_p(t)$ we want to prove some necessary inequalities in the following preliminary lemma.

Lemma 3.6. *Let* $a, b \geq 0$, $0 \leq t \leq 1$ *and* $p \geq 1$, $q \geq 2$*. Then the following inequalities are true:*

$$\begin{aligned}
&\textit{(i)} & (a+b)^p &\leq 2^{p-1}(a^p + b^p) \\
&\textit{(ii)} & (1+t^p)^{1/p} - 1 &\leq \frac{t^p}{p} \\
&\textit{(iii)} & \left(\frac{(1+t)^q + (1-t)^q}{2}\right)^{1/q} - 1 &\leq \frac{q-1}{2}t^2
\end{aligned}$$

Proof. (i) Since $x \mapsto x^p$ is convex on $[0, \infty]$ for $p \geq 1$ we have for $a, b \geq 0$:

$$\left(\frac{1}{2}a + \frac{1}{2}b\right)^p \leq \frac{1}{2}(a^p + b^p).$$

Multiplying both sides with 2^p shows the statement.

(ii) Set

$$f(t) = \left(1 + \frac{t^p}{p}\right)^p - (1 + t^p).$$

We have to show that $f(t) \geq 0$ for all $t \geq 0$. It obviously holds that $f(0) = 0$ and

$$f'(t) = \left[\left(1 + \frac{t^p}{p}\right)^{p-1} - 1\right] \cdot p \cdot t^{p-1} \geq 0$$

Hence $f(t) \geq 0$ for all $t \geq 0$.

(iii) Let $0 \leq t \leq 1$. Then the statement follows from a particular version of the hypercontractivity inequality (see for example Lemma 1 in [2]) which reads

$$\left(\frac{(1+\rho\cdot\varepsilon)^q + (1-\rho\cdot\varepsilon)^q}{2}\right)^{1/q} \leq \left(\frac{(1+\varepsilon)^p + (1-\varepsilon)^p}{2}\right)^{1/p} \iff \rho^2 \leq \frac{p-1}{q-1}$$

for all $\rho, \varepsilon \geq 0$. Setting $t = \rho \cdot \varepsilon$, $p = 2$ and $\rho^2 = 1/(q-1)$ we obtain

$$\begin{aligned}\left(\frac{(1+t)^q + (1-t)^q}{2}\right)^{1/q} &\leq \left(\frac{(1+\varepsilon)^2 + (1-\varepsilon)^2}{2}\right)^{1/2} \\ &= \left(1+\varepsilon^2\right)^{1/2} \leq 1 + \frac{\varepsilon^2}{2} = 1 + \frac{q-1}{2} \cdot t^2\end{aligned}$$

This shows the statement. □

Corollary 3.7. *The following inequalities hold for* $0 \leq t \leq 1$*:*

$$\rho_p(t) \leq \begin{cases} \frac{1}{p} \cdot t^p & \text{if } 1 \leq p \leq 4/3 \\ \frac{p-1}{2} \cdot t^2 & \text{if } p = 2 \text{ or } 4 \leq p < \infty \end{cases}$$

This implies that for $1 < p \leq 4/3$, $p = 2$ *and* $4 \leq p < \infty$, $\mathcal{M}_d$ *equipped with the Schatten p-norm is uniformly smooth. Furthermore,* $\mathcal{M}_d$ *with the Schatten* 1*-norm is not uniformly smooth.*

Proof. (i) Let $1 \leq p \leq 4/3$. To show the inequality, we apply Hanner's inequality for $A := X + tY$ and $B := X - tY$ and Lemma 3.6 to obtain

$$\begin{aligned}\rho_p(t) &= \sup\left\{\frac{1}{2}(\|X+tY\|_p + \|X-tY\|_p) - 1 : \|X\|_p = \|Y\|_p = 1\right\} \\ &\leq (1+t^p)^{1/p} - 1 \leq \frac{t^p}{p}.\end{aligned}$$

(ii) Let $p = 2$ or $4 \leq p < \infty$. Using Lemma 3.6 (i), (iii) and Hanner's inequality for $A := X$ and $B := tY$ we obtain

$$\rho_p(t) = \sup\left\{\frac{1}{2}(\|X + tY\|_p + \|X - tY\|_p) - 1 : \|X\|_p = \|Y\|_p = 1\right\}$$

$$\leq \sup\left\{\left(\frac{\|X + tY\|_p^p + \|X - tY\|_p^p}{2}\right)^{1/p} - 1 : \|X\|_p = \|Y\|_p = 1\right\}$$

$$\leq \left(\frac{(1 + t)^p + |1 - t|^p}{2}\right)^{1/p} - 1 \leq \frac{p - 1}{2} \cdot t^2.$$

To prove that the Schatten 1-class is not uniformly smooth we refer to the ℓ_p-case (see [27] for details). □

Using the fact that the modulus of smoothness is increasing with t and upper bounded by a power of t, we can formulate the approximate Carathéodory Theorem for several Schatten classes with an upper bound independent of the modulus of smoothness.

Theorem 3.8 (Approximate Carathéodory for Schatten classes)**.** *Let $\mathcal{S}$ be a bounded set in $\mathcal{M}_d$ equipped with Schatten p-norm, where $1 < p \leq 4/3$, $p = 2$ or $4 \leq p < \infty$, and $A \in \operatorname{conv}(\mathcal{S})$. Then there exists a sequence $\{X_i\}_{i=1}^{\infty} \subseteq \mathcal{S}$ such that for the convex combination*

$$A_k = \frac{1}{k}\sum_{i=1}^{k} X_i$$

the following inequalities hold:

(a) $\|A - A_k\|_p \leq \dfrac{2\exp(2)}{p^{1/p}} \cdot k^{1/p-1} \cdot \operatorname{diam}(\mathcal{S})$ *if* $1 < p \leq 4/3$

(b) $\|A - A_k\|_p \leq \exp(2) \cdot \sqrt{\dfrac{2(p-1)}{k}} \cdot \operatorname{diam}(\mathcal{S})$ *if* $p = 2$ *or* $4 \leq p < \infty$

Proof. Using the fact that ρ_p is increasing in its argument, and Theorem 3.4 together with the fact that $0 \leq 1/k \leq 1$ for $k \in \mathbb{N}$, we obtain

$$\rho_p^{-1}(1/k) \geq \begin{cases} \sqrt[p]{\frac{p}{k}} & \text{if } 1 < p \leq 4/3 \\ \sqrt{\frac{2}{k\cdot(p-1)}} & \text{if } p = 2 \text{ or } 4 \leq p < \infty \end{cases}$$

This proves the statement. □

Note that the proof of Theorem 3.4 is constructive, hence the elements of the sequence $\{X_i\}_{i=1}^{\infty}$ can be computed by a deterministic algorithm. We will study this more thoroughly in Section 4.

By fixing some approximation error $\varepsilon > 0$ and p, the previous result can be equivalently formulated as an upper bound on the number of summands k necessary to attain an ε-approximation with respect to the Schatten p-norm.

Theorem 3.9. *Let $\mathcal{S} \subseteq \mathcal{M}_d$ be bounded, $A \in conv(\mathcal{S})$ and $\varepsilon > 0$ be given. Then in the ε-ball with respect to the Schatten p-norm around A there is a point B which is a convex combination of at most*

$$\text{(a)} \quad \left\lceil C_p \cdot \left(\frac{diam(\mathcal{S})}{\varepsilon}\right)^{\frac{p}{p-1}} \right\rceil \qquad \text{if } 1 < p \leq 4/3$$

$$\text{(b)} \quad \left\lceil D_p \cdot \left(\frac{diam(\mathcal{S})}{\varepsilon}\right)^{2} \right\rceil \qquad \text{if } p = 2 \text{ or } 4 \leq p < \infty$$

points from S, where

$$C_p := \left(\frac{2\exp(2)}{p^{1/p}}\right)^{\frac{p}{p-1}} \quad \text{and} \quad D_p := 2(p-1)\cdot\exp(4).$$

Proof. It follows directly from Theorem 3.8. □

Let us highlight the main features of Theorem 3.9. For a fixed approximation error $\varepsilon > 0$ and a norm parameter p, the above theorem differs from Theorem 3.1 only by the fact that we consider approximations. In both cases (a) and (b), we obtain, in contrast to Theorem 3.1, a dimension independent upper bound. Further, the upper bound is best for $p = 2$ and increases up to infinity for fixed $\varepsilon > 0$ if we approach $p = 1$. Further, the bounds also diverge if we take the limit $\varepsilon \to 0$ for arbitrary p, which is reasonable since the dimension dependent upper bound of Theorem 3.1 is optimal. Hence, for applying this result, the approximation error has always to be fixed.

The above discussion can be repeated for entrywise ℓ_p-norms which will be important as a natural norm later on for nonnegative tensors.

Remark 3.10. If we assume that $\mathcal{M}_d$ is equipped with the ℓ_p-norm $\|\cdot\|_{\ell_p}$, i.e. for $A \in \mathcal{M}_d$ we have

$$\|A\|_{\ell_p} := \left(\sum_{i,j=1}^{d} |A_{ij}|^p\right)^{1/p},$$

then the upper bounds of Theorem 3.8 and Theorem 3.9 hold for $1 < p \leq 2$ and $p \geq 2$, instead of $1 < p \leq 4/3$ and $4 \leq p < \infty$, respectively. This is due to the fact that Hanner's inequalities are proven for $1 < p < \infty$ in ℓ_p-spaces [27]. △

2 Approximate (Ω, G)-decompositions and ranks

In the following, we will define the notions of approximate (Ω, G)-decompositions for both psd matrices and nonnegative tensors. We will also apply the results from Section 1 to obtain upper bounds for ranks of approximate (Ω, G)-decompositions.

The section is organized as follows. In Section 2.1 and Section 2.2, we define the approximate analogs of the (Ω, G)-ranks for psd matrices and nonnegative tensors, respectively. In Section 2.3, we introduce the notion of gauge functions, a relevant tool to obtain the upper bounds. Subsequently, we show upper bounds for general matrices (Section 2.4), psd matrices (Section 2.5), and separable states (Section 2.6).

As before we consider

$$\mathcal{M}_{d_0} \otimes \mathcal{M}_{d_1} \otimes \cdots \otimes \mathcal{M}_{d_n} \cong \mathcal{M}_{d_0 \cdots d_n}.$$

Further we fix a connected wsc Ω and a free group action G on Ω.

2.1 Approximate decompositions of psd matrices

In this part, we introduce the different notions of approximate (Ω, G)-ranks in the space $\mathcal{M}_{d_0} \otimes \cdots \otimes \mathcal{M}_{d_n}$. Given a matrix ρ, the notion of the approximate rank is the minimal rank of all matrices contained in the ε-ball of ρ with respect to the Schatten p-norm $\| \cdot \|_p$.

Definition 3.11. *Let $p \in [1, \infty)$ and $\varepsilon > 0$. Further let $M \in \mathcal{M}_{d_0} \otimes \cdots \otimes \mathcal{M}_{d_n}$ and $\rho \in \mathcal{M}^+_{d_0 \cdots d_n}$. We define*

$$\textit{rank}^{\varepsilon,p}_{(\Omega,G)}(M) := \min \left\{ \textit{rank}_{(\Omega,G)}(N) : \|M - N\|_p \leq \varepsilon, N \in \mathcal{M}_{d_0 \cdots d_n} \right\},$$

similarly

$$\textit{puri-rank}^{\varepsilon,p}_{(\Omega,G)}(\rho) := \min \left\{ \textit{puri-rank}_{(\Omega,G)}(\sigma) : \|\rho - \sigma\|_p \leq \varepsilon, \sigma \in \mathcal{M}_{d_0 \cdots d_n} \right\}$$

and

$$\textit{q-sqrt-rank}^{\varepsilon,p}_{(\Omega,G)}(\rho) := \min \left\{ \textit{q-sqrt-rank}_{(\Omega,G)}(\sigma) : \|\rho - \sigma\|_p \leq \varepsilon, \sigma \in \mathcal{M}_{d_0 \cdots d_n} \right\}.$$

If ρ is separable, we define

$$\textit{sep-rank}^{\varepsilon,p}_{(\Omega,G)}(\rho) := \min \left\{ \textit{sep-rank}_{(\Omega,G)}(\sigma) : \|\rho - \sigma\|_p \leq \varepsilon, \sigma \in \mathcal{M}_{d_0} \otimes \cdots \otimes \mathcal{M}_{d_n} \right\}.$$

Note that the exact purification rank and quantum square root rank are infinite if σ is not psd. Hence, an approximate purification and quantum square root decomposition is always attained by a psd matrix. This behavior is similar for the separable rank.

Further note that $\text{puri-rank}^{\varepsilon,p}_{(\Omega,G)}$ and $\text{q-sqrt-rank}^{\varepsilon,p}_{(\Omega,G)}$ can, in principle, also be defined for non-psd matrices and might be finite. This is due to the fact that in particular cases the ε-ball around a non-psd matrix can contain a psd matrix which has a finite purification rank. Similar $\text{sep-rank}^{\varepsilon,p}_{(\Omega,G)}$ can also be defined for non-separable matrices. Let us now go back again to Example 2.2 and Example 2.4, and explain the notions of approximate ranks on the cases therein.

A different notion in the study of approximate ranks is the notion of *border rank*, which is the minimal $k \in \mathbb{N}$ among all possible sequences $(\rho_n)_{n \in \mathbb{N}}$ consisting of rank-k matrices ρ_n converging to ρ [5]. As already pointed out, the upper bounds given in Theorem 3.8 diverge for $\varepsilon \to 0$. Hence the approach using the approximate Carathéodory Theorem fails when studying these notions of ranks.

In the following example, we will give some examples of approximate decompositions related to the relevant decompositions introduced and studied in Chapter 1 and Chapter 2.

Example 3.12. Let $M \in \mathcal{M}_{d_0\cdots d_n}$, $\varepsilon > 0$ and $p \in [1,\infty)$.

(i) An approximate Σ_n-decomposition (i.e. a decomposition corresponding to the n-simplex defined in Example 2.4) is given by a matrix $N \in \mathcal{M}_{d_0\cdots d_n}$ admitting a decomposition

$$N = \sum_{\alpha=1}^{r} N_\alpha^{[0]} \otimes \cdots \otimes N_\alpha^{[n]}$$

with $\|M - N\|_p \leq \varepsilon$. The approximate Σ_n-rank, $\mathrm{rank}_{\Sigma_n}^{\varepsilon,p}(\rho)$, is called *approximate tensor rank* of M and is the smallest integer r among all matrices N in the ε-ball of M and all Σ_n-decompositions.

(ii) Consider the line $\Omega = \Lambda_n$ of length n. The *approximate operator Schmidt rank*, $\mathrm{rank}_{\Lambda_n}^{\varepsilon,p}(M)$, is the minimal integer r among all $N \in \mathcal{M}_{d_0\cdots d_n}$ with $\|M - N\|_p \leq \varepsilon$ and all decompositions of the matrix product operator form with open boundary conditions.

(iii) For $n \geq 3$ consider the circle $\Omega = \Theta_n$ of length n together with the cyclic group $G = C_n$ whose elements are translations of the points on the line. Further let $M \in \mathcal{M}_{d^n}$ be C_n-invariant. In this example C_n-invariance corresponds to translational invariance of M. The *approximate translational invariant operator Schmidt rank* of M, $\mathrm{rank}_{(\Theta_n,C_n)}^{\varepsilon,p}(M)$, is the minimal integer r among all $N \in \mathcal{M}_{d^n}$ with $\|M - N\|_p \leq \varepsilon$ and all decompositions of N of the translational invariant matrix product operator form. △

2.2 Approximate decompositions of nonnegative tensors

In the following, we define the notions of approximate (Ω, G)-ranks similar to Section 2.1 using the exact (Ω, G)-decompositions defined in Section 2.3 of Chapter 2. The only difference in the definition of the approximate ranks is the norm used.

Motivated by the correspondence between psd matrices and nonnegative tensors given in Equation (2.3), we will use for $p \geq 1$ the ℓ_p-norm which is defined for $M \in \mathcal{K}_{n,d}$ as

$$\|M\|_{\ell_p} = \left(\sum_{i_0,\ldots,i_n} |m_{i_0,\ldots,i_n}|^p\right)^{1/p}.$$

For this reason, this norm is sometimes also called the entrywise p-norm. Since in the following definition of the approximate (Ω, G)-decompositions of nonnegative tensors the relevant norm is the ℓ_p-norm, we will indicate this fact by using ℓ_p instead of p.

Definition 3.13. *Let $p \in [1,\infty)$, $\varepsilon > 0$ and $M \in \mathcal{K}$. We define*

$$\mathit{rank}_{(\Omega,G)}^{\varepsilon,\ell_p}(M) := \min\left\{\mathit{rank}_{(\Omega,G)}(N) : \|M - N\|_{\ell_p} \leq \varepsilon, N \in \mathcal{K}\right\}.$$

If M is nonnegative, we similarly define

$$\mathit{nn\text{-}rank}_{(\Omega,G)}^{\varepsilon,\ell_p}(M) := \min\left\{\mathit{nn\text{-}rank}_{(\Omega,G)}(N) : \|M - N\|_{\ell_p} \leq \varepsilon, N \in \mathcal{K}\right\},$$

$$\mathit{psd\text{-}rank}_{(\Omega,G)}^{\varepsilon,\ell_p}(M) := \min\left\{\mathit{psd\text{-}rank}_{(\Omega,G)}(N) : \|M - N\|_{\ell_p} \leq \varepsilon, N \in \mathcal{K}\right\},$$

and

$$\textit{sqrt-rank}^{\varepsilon,\ell_p}_{(\Omega,G)}(M) := \min\left\{\textit{sqrt-rank}_{(\Omega,G)}(N) : \|M - N\|_{\ell_p} \leq \varepsilon, N \in \mathcal{K}\right\}.$$

Note that similar to the psd matrices, the latter three notions of ranks can also be defined for general elements in $\mathcal{K}$. In the following example, we want to relate the approximate tensor decompositions with the matrix factorizations introduced in Chapter 1.

Example 3.14. Let Λ_1 be the line of length one and $G = S_2$ the symmetric group acting on Λ_1. Further let $M \in \mathcal{K}_{2,d} \cong \mathcal{M}_d$ be nonnegative.

(i) An approximate Λ_1-decomposition of M is given by a tensor $N \in \mathcal{K}_{2,d}$ in the ε-ball of M and a decomposition

$$N = \sum_{\alpha=1}^{r} v_\alpha^{[0]} \otimes v_\alpha^{[1]}.$$

This decomposition corresponds to a minimal factorization of $N = A \cdot B^T$ (Definition 1.12) setting the column α of A and B to $v_\alpha^{[0]}$ and $v_\alpha^{[1]}$, respectively.

(ii) An approximate (Λ_1, G)-decomposition of a symmetric matrix M corresponds to a symmetric matrix $N \in \mathcal{K}_{2,d} \cong \mathcal{M}_d$ in the ε-ball of M and a decomposition.

$$N = \sum_{\alpha=1}^{r} v_\alpha \otimes v_\alpha.$$

This decomposition corresponds to the symmetric factorization $N = A \cdot A^T$ (Definition 1.16) associating the column α of A with v_α.

(iii) An approximate Λ_1-nonnegative decomposition corresponds to a nonnegative element $N \in \mathcal{K}_{2,d}$ in the ε-ball of M and a nonnegative factorization of N (Definition 1.13). Similarly, an approximate (Λ_1, G)-nonnegative decomposition corresponds to an approximate version of the completely positive factorization (Definition 1.16 (ii)).

(iv) An approximate Λ_1-psd-decomposition corresponds to a nonnegative element $N \in \mathcal{K}_{2,d}$ in the ε-ball of M and a psd-factorization (Definition 1.14)

$$N_{kl} = \sum_{\alpha_0,\alpha_1=1}^{r} \left(A_k^{[0]}\right)_{\alpha_0,\alpha_1} \cdot \left(A_l^{[1]}\right)_{\alpha_0,\alpha_1} = \operatorname{tr}\left(\left(A_k^{[0]}\right) \cdot \left(A_l^{[1]}\right)^T\right).$$

An approximate (Λ_1, G)-psd-decomposition corresponds to an approximate version of the cpsdt-factorization (Definition 1.16 (iii)) setting $A_k := A_k^{[0]} = A_k^{[1]}$.

△

In Example 3.27, we will illustrate the behavior of these ranks.

2.3 More norms for matrices

As we have already seen in the introduction of this chapter, the central idea which makes Carathéodory's Theorem and its approximate analog applicable is the fact that we can associate every convex cone - such as the set of psd matrices or the set of nonnegative tensors - with a convex hull generated infitiely many elementary tensors which define the summands in the different decompositions. In contrast to the exact Carathéodory Theorem, in the approximate version (Theorem 3.8), the generating set of the convex hull has to be bounded; in particular, the distance between all elements in the generating set has to be bounded. Hence, elements have to be normalized with a particular norm to associate them to an element of the convex hull generated by a bounded set.

In this section, we introduce and study the gauge function as a norm, which is convenient for fulfilling the properties mentioned above. In particular, we introduce a bounded convex set and reconstruct a norm whose unit ball is equivalent to the convex set. We start with some preparatory observations and definitions.

For the (non-scaled) Schatten norms, where $1 \leq p \leq q \leq \infty$, we have the following inequalities [38]:

$$\|M\|_q \leq \|M\|_p \leq \mathrm{rank}(M)^{\frac{1}{p}-\frac{1}{q}} \|M\|_q. \tag{3.2}$$

Further, we define for any $1 \leq p \leq \infty$ the generating set

$$P_p := \left\{ \pm\rho^{[0]} \otimes \cdots \otimes \rho^{[n]} \,:\, \rho^{[i]} \in \mathcal{M}^+_{d_i} \text{ for } i \in \{0,\ldots,n\} \text{ and } \|\bigotimes_{i=0}^{n} \rho^{[i]}\|_p \leq 1 \right\}$$

and consider the convex hull

$$B_p := \mathrm{conv}(P_p) \subseteq \mathrm{Her}_{d_0 \cdots d_n}$$

where $\mathrm{Her}_d \subseteq \mathcal{M}_d$ denotes the set of hermitian $d \times d$ matrices.

Note that for $p \leq q$, we have $B_p \subseteq B_q$, and already B_1 contains all separable states since all product matrices in the unit ball of the Schatten 1-norm are valid states. Each B_p is compact (i.e. closed and bounded with respect to the Schatten norm) and as a convex hull obviously convex. Furthermore, B_p is centrally symmetric (i.e. if $\rho \in B_p$ then $-\rho \in B_p$) and hence by convexity it contains the origin in its interior. We can thus understand it as the unit ball of a norm: For a set S in a real vector space V, the *gauge function* μ_S is defined by

$$\mu_S(v) := \inf\left\{ \lambda > 0 \mid \frac{1}{\lambda} v \in S \right\} \tag{3.3}$$

for $v \in V$. If S is compact, convex, centrally symmetric and has nonempty interior, the gauge function μ_S is in fact a norm (see for example Theorem 15.2 in Ref. [32]), and S is clearly its unit ball.

In the following, we denote the gauge function of B_p by μ_p for all $1 \leq p < \infty$. We now want to relate the gauge functions to a multipartite version of the robustness of entanglement [35].

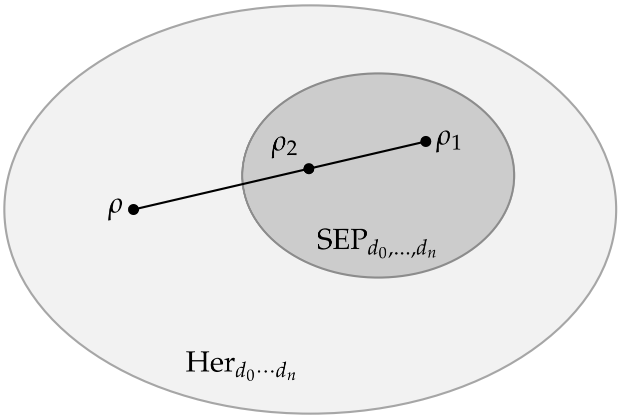


FIGURE 3.2: Visualization of the robustness of entanglement.

Definition 3.15. *Let $\rho \in \mathcal{M}_{d_0} \otimes \cdots \otimes \mathcal{M}_{d_n}$ be a state. The* robustness of entanglement of ρ *is defined as*

$$R(\rho) := \inf\Big\{\lambda \geq 1 : \rho = (1-\lambda)\cdot\rho_1 + \lambda\cdot\rho_2 \text{ and } \rho_i \text{ are separable states}\Big\}.$$

A visualization of this relation is shown in Figure 3.2.

Note that this definition varies from the original definition by a additive constant of 1. Obviously, $R(\rho) < \infty$ for all hermitian matrices ρ since all hermitian matrices can be written as linear combinations of elements contained in P_p.

Proposition 3.16. *For $\rho \in \mathrm{Her}_{d_0} \otimes \cdots \otimes \mathrm{Her}_{d_n} \cong \mathrm{Her}_{d_0\cdots d_n}$ and $1 \leq p \leq q < \infty$ we have:*

(i) $\|\rho\|_p \leq \mu_p(\rho)$.

(ii) $\mu_q(\rho) \leq \mu_p(\rho) \leq (d_0 \cdots d_n)^{1/p-1/q} \cdot \mu_q(\rho)$.

(iii) If ρ is a separable state, then $\mu_p(\rho) \leq 1$.

(iv) If ρ is a state, then $R(\rho) \leq \mu_1(\rho) \leq 2R(\rho)$.

(v) Define $\mu_{\sqrt{},p}(v) = \min\Big\{\mu_p(\sqrt{v})\Big\}$. If $v^2 = v$ then $\mu_{\sqrt{},p}(v) \leq \mu_p(v)$.

Proof. Since the unit ball of the Schatten p-norm contains P_p and is convex, it contains B_p. This implies that $\|\rho\|_p \leq \mu_p(\rho)$, which shows statement (i).

(ii) The first inequality is a consequence of the fact that we have for the corresponding unit balls $B_p \leq B_q$. The second inequality is a direct implication of relation (3.2).

(iii) is clear from the fact that $\rho \in B_1$ together with (ii).

For the first inequality in (iv) we express

$$\frac{1}{\mu_1(\rho)}\rho = \sum_i \lambda_i \xi_i$$

as a convex combination of elements $\xi_i \in P_1$, with all $\lambda_i > 0, \|\xi_i\| > 0$. From the minimality of $\mu_1(\rho)$ it follows that $\|\xi_i\|_1 = 1$ holds for all i. Sorting the positive and negative terms we obtain

$$\frac{1}{\mu_1(\rho)}\rho = r\sigma_1 - (1-r)\sigma_2$$

with separable states σ_i, where $0 \leq r \leq 1$ is the sum over those λ_i with ξ_i psd. Taking the trace on both sides shows $2\mu_1(\rho)r = 1 + \mu_1(\rho)$. Thus $\lambda := \mu_1(\rho)r$ yields $\rho = (1-\lambda)\sigma_2 + \lambda\sigma_1$ and thus $R(\rho) \leq \mu_1(\rho)$.

For the second inequality, we express $\rho = (1-\lambda)\rho_1 + \lambda\rho_2$ where ρ_i are separable states and $\lambda \geq 1$. Since $\mu_1(\rho_i) \leq 1$, the second inequality follows from the triangle inequality for μ_1.

(v) Immediate from the definition. □

2.4 Upper bounds for approximate ranks

We now prove an upper bound for the approximate (Ω, G)-rank which only depends on the gauge function value of the matrix and the approximation error ε.

Recall that a matrix $M \in \mathrm{Her}_{d_0} \otimes \cdots \otimes \mathrm{Her}_{d_n} \cong \mathrm{Her}_{d_0 \cdots d_n}$ is G-invariant if it is invariant with respect to the group action G which contains permutations of the local tensor product spaces.

Theorem 3.17. *Let $1 < p \leq 4/3$, $p = 2$ or $4 \leq p < \infty$ and $\varepsilon > 0$. Assume that $M \in \mathrm{Her}_{d_0} \otimes \cdots \otimes \mathrm{Her}_{d_n} \cong \mathrm{Her}_{d_0 \cdots d_n}$ is G-invariant. Then*

(a) $\mathrm{rank}^{\varepsilon,p}_{(\Omega,G)}(M) \leq \left\lceil C_p \cdot \left(\frac{2\mu_p(M)}{\varepsilon}\right)^{\frac{p}{p-1}} \right\rceil \cdot |G|$ *if* $1 < p \leq 4/3$

(b) $\mathrm{rank}^{\varepsilon,p}_{(\Omega,G)}(M) \leq \left\lceil D_p \cdot \left(\frac{2\mu_p(M)}{\varepsilon}\right)^{2} \right\rceil \cdot |G|$ *if* $p = 2$ *or* $4 \leq p < \infty$

The proof consists of two ingredients. First, we apply Theorem 3.9 with set P_p to construct a Ω-decomposition. Afterwards, we obtain a decomposition containing G-invariant elements by applying the construction in the proof of [12, Thm. 13].

Proof. First let $p = 2$ or $4 \leq p < \infty$. We prove the statement for the case that $\mu_p(M) = 1$; the general case follows by replacing M with $M/\mu_p(M)$. First we consider the case of trivial group action. By assumption $M \in \mathrm{conv}(P_p)$, and $\mathrm{diam}(P_p) = 2$. By Theorem 3.9 we find an M' such that $\|M - M'\|_p \leq \varepsilon$ and where M' is a convex combination of at most $\lceil 4D_p/\varepsilon^2 \rceil$ elements of P_p. This implies that

$$\mathrm{rank}_\Omega(M') \leq \mathrm{rank}_{\Sigma_n}(M') \leq \lceil 4D_p/\varepsilon^2 \rceil,$$

where the first inequality follows from [12, Prop. 36]. We now involve a group action and apply the construction of [12, Thm. 13] to M'. This procedure gives rise to an (Ω, G)-decomposition of an element $M'' = \frac{1}{|G|}\sum_g gM'$ with

$$\mathrm{rank}_{(\Omega,G)}(M'') \leq \lceil 4D_p/\varepsilon^2 \rceil \cdot |G|.$$

It remains to prove that M'' is contained in the ε-neighborhood of M:

$$\begin{aligned} \|M'' - M\|_p &= \left\| \frac{1}{|G|} \sum_g (gM' - gM) \right\|_p \\ &\leq \frac{1}{|G|} \sum_g \|gM' - gM\|_p \leq \varepsilon \end{aligned}$$

where we used the invariance $M = \frac{1}{|G|} \sum_g gM$, that the norm is unitarily invariant $\|gM' - gM\|_p = \|M' - M\|_p$, and that M' is in the ε-ball around M. The other case, $1 < p \leq 4/3$, is analogous. □

Remark 3.18. Note that Theorem 3.17 and all following theorems give upper bounds for particular convex combinations of the type

$$M_k = \frac{1}{k} \sum_{i=1}^{k} X_i,$$

where $X_i \in P_p$. This is due to the construction in Theorem 3.8. Hence, the sequence $\{X_i\}_{i=1}^k$ might contain a repetition of same elements which would decrease the rank of the decomposition, but our estimates do not exploit this fact. △

2.5 Upper bounds for approximate ranks of psd matrices

We now provide upper bounds for the approximate purification rank and the approximate quantum square root rank of psd matrices. More specifically, we provide an upper bound for the approximate quantum square root rank, which upper bounds the approximate purification rank due to Remark 2.9 (i).

Recall that Ω is a connected wsc and G a free group action on Ω.

Corollary 3.19. *Let $1 < p \leq 4/3$, $p = 2$ or $4 \leq p < \infty$ and $\varepsilon > 0$. Let $\rho \in \mathcal{M}^+_{d_0 \cdots d_n}$ be G-invariant. Then*

(a) $\text{puri-rank}^{\varepsilon,p}_{(\Omega,G)}(\rho) \leq \left\lceil C_p \cdot \left(\frac{2}{\sqrt{1+\varepsilon/\mu^2_{\sqrt{},p}(\rho)} - 1} \right)^{\frac{p}{p-1}} \right\rceil \cdot |G| \quad$ *if* $1 < p \leq 4/3$

(b) $\text{puri-rank}^{\varepsilon,p}_{(\Omega,G)}(\rho) \leq \left\lceil D_p \cdot \left(\frac{2}{\sqrt{1+\varepsilon/\mu^2_{\sqrt{},p}(\rho)} - 1} \right)^{2} \right\rceil \cdot |G| \quad$ *if* $p = 2$ *or* $4 \leq p < \infty$

and the same upper bounds hold for $\text{q-sqrt-rank}^{\varepsilon,p}_{(\Omega,G)}$.

Proof. We apply Theorem 3.17 to the square root of ρ (called M) which realises $\mu_{\sqrt{},p}(\rho)$. We have that $M/\mu_{\sqrt{},p}(\rho) \in \text{conv}(P)$ and thus $\|M\|_p \leq \mu_{\sqrt{},p}(\rho)$ (defined in Proposition 3.16 (v)). We now choose an element M' which is $\mu_{\sqrt{},p}(\rho) \cdot \delta$ - close to M

with $\delta = \sqrt{1+\varepsilon/\mu^2_{\sqrt{},p}(\rho)} - 1$. We define $\rho' = M'^2$ and compute

$$\begin{aligned} \|\rho - \rho'\|_p &= \|M^2 - M'^2\|_p = \|M - M'\|_p \cdot \|M + M'\|_p \\ &\leq \|M - M'\|_p \cdot (2\|M\|_p + \|M - M'\|_p) \\ &\leq \mu_{\sqrt{},p}(\rho)^2 \cdot \delta \cdot (2+\delta) \leq \varepsilon. \end{aligned}$$

Using that (see Theorem 2.13 (iv) and Definition 2.8 (ii))

$$\text{puri-rank}_{(\Omega,G)}(\rho') \leq \text{q-sqrt-rank}_{(\Omega,G)}(\rho') \leq \text{rank}_{(\Omega,G)}(M')$$

we obtain the result. □

Note that this result upper bounds the quantum square root rank, which might be arbitrarily larger than the purification rank (see Theorem 2.16 (b) for details).

2.6 Upper bounds for approximate ranks of separable states

An upper bound for separable states can be obtained without the use of a gauge function. This is possible because a separable state is in the convex hull of

$$\left\{ +\rho^{[0]} \otimes \cdots \otimes \rho^{[n]} : \rho^{[i]} \in \mathcal{M}^+_{d_i} \text{ and } \|\bigotimes_{i=0}^{n} \rho^{[i]}\|_p \leq 1 \right\} \subseteq P_p.$$

Using that $\text{sep-rank}_{(\Omega,G)}$ upper bounds $\text{rank}_{(\Omega,G)}$ and $\text{puri-rank}_{(\Omega,G)}$, we conclude that this upper bound holds for the other two ranks in the approximate case too. Recall that Ω is a connected wsc and G a free group action on Ω.

Proposition 3.20. *Let $1 < p \leq 4/3$, $p = 2$ or $4 \leq p < \infty$ and $\varepsilon > 0$. Further let $\rho \in \text{SEP}_{d_0,\ldots,d_n}$ be G-invariant. Then*

(a) $\text{sep-rank}^{\varepsilon,p}_{(\Omega,G)}(\rho) \leq \left\lceil C_p \cdot (2/\varepsilon)^{\frac{p}{p-1}} \right\rceil \cdot |G|$ *if* $1 < p \leq 4/3$

(b) $\text{sep-rank}^{\varepsilon,p}_{(\Omega,G)}(\rho) \leq \left\lceil D_p \cdot (2/\varepsilon)^2 \right\rceil \cdot |G|$ *if* $p = 2$ *or* $4 \leq p < \infty$

The same upper bound holds for $\text{rank}^{\varepsilon,p}_{(\Omega,G)}(\rho)$ *and* $\text{puri-rank}^{\varepsilon,p}_{(\Omega,G)}(\rho)$*, too.*

Proof. It is proven exactly as Theorem 3.17 by using that separable states are a convex combination of product states, which are a subset of P_p. The rest is an immediate consequence of Theorem 2.13 (i) and (ii). □

Note again that the separable rank may be arbitrarily larger than the rank and the purification rank, as we have already seen in Theorem 2.16 (c).

3 Disappearance of separations in the approximate case

We now turn to study relations between the different notions of ranks, especially their (lack of) separations in the approximate case.

Let $\mathcal{X}$ be an arbitrary set and $f, g\colon \mathcal{X} \to \mathbb{N}$ two functions. We say that there is a *separation between* f *and* g, and write $f \ll g$, if there exists a sequence in $\mathcal{X}$ along which f is bounded but g is not. This implies that the values of g cannot be bounded by a function of the values of f. In the exact case, there are several examples of separations between different ranks — for example on the set

$$\mathcal{X} := \bigcup_{d\in\mathbb{N}} \mathcal{M}^{+}_{d^{n+1}}$$

one has $\text{rank}_{(\Omega,G)} \ll \text{puri-rank}_{(\Omega,G)}$, for suitable choices of (Ω, G) [12, 14, 19]. In the following, we will show that many separations of ranks of psd matrices disappear (Section 3.1), and the same happens for nonnegative tensors (Section 3.2). This is an immediate consequence of the fact that the approximate ranks admit upper bounds independent of the dimension. We denote this behavior by $\not\ll$. In other words, we have $f \not\ll g$ if there exists a function $F : \mathbb{N} \to \mathbb{N}$ such that $g \leq F \circ f$ where $\circ$ denotes the composition of functions.

As in the previous section, we again fix a connected wsc Ω and a free group action G acting on Ω.

3.1 Positive semidefinite matrices

In the following, we show that the separation between several ranks of psd matrices vanishes in the approximate case. The strategy is simple: we use Theorem 3.17, Proposition 3.20 and Corollary 3.19 to upper bound the ranks independently of the matrix dimension. Since all ranks are bounded functions, it follows that many separations vanish.

For simplicity we assume that $d_i = d$ and hence consider the space

$$\mathcal{M}_d \otimes \cdots \otimes \mathcal{M}_d \cong \mathcal{M}_{d^{n+1}}$$

where the number of tensor product spaces $n+1$ keeps fixed. The result can be extended in a straightforward manner to the case that the local vector spaces of the tensor product have different dimensions.

Corollary 3.21. *Let* $\varepsilon > 0$, $K \in \mathbb{N}$ *and* $1 < p \leq 4/3$, $p = 2$ *or* $4 \leq p < \infty$. *We define the set*

$$\mathcal{X}_K := \{\rho \in \mathcal{M}^{+}_{d^{n+1}} : d \in \mathbb{N} \text{ and } \mu_{\sqrt{},p}(\rho) \leq K\}.$$

Then the following holds on $\mathcal{X}_K$*:*

$$\begin{array}{llcl} (i) & \textit{rank}^{\varepsilon,p}_{(\Omega,G)} & \not\ll & \textit{puri-rank}^{\varepsilon,p}_{(\Omega,G)} \\ (ii) & \textit{puri-rank}^{\varepsilon,p}_{(\Omega,G)} & \not\ll & \textit{q-sqrt-rank}^{\varepsilon,p}_{(\Omega,G)} \end{array}$$

Further, we define the set

$$\mathcal{X}_{sep} := \bigcup_{d\in\mathbb{N}} \mathrm{SEP}_{n,d}.$$

The following holds on $\mathcal{X}_{sep}$:

$$\begin{array}{lll} (iii) & rank^{\varepsilon,p}_{(\Omega,G)} & \not\ll \; sep\text{-}rank^{\varepsilon,p}_{(\Omega,G)} \\ (iv) & puri\text{-}rank^{\varepsilon,p}_{(\Omega,G)} & \not\ll \; sep\text{-}rank^{\varepsilon,p}_{(\Omega,G)} \end{array}$$

Note that (i) and (ii) imply that $\mathrm{rank}^{\varepsilon,p}_{(\Omega,G)} \not\ll \text{q-sqrt-rank}^{\varepsilon,p}_{(\Omega,G)}$ on $\mathcal{X}_K$.

Proof. (i)-(ii) For $\rho \in \mathcal{M}^+_{d^{n+1}}$ we have $\mathrm{rank}_{(\Omega,G)}(\rho) \le \text{puri-rank}_{(\Omega,G)}(\rho)^2$ [12, Prop. 29]. By the basic properties of $\text{q-sqrt-rank}_{(\Omega,G)}$ (see Remark 2.9 (i) for details) we obtain

$$\sqrt{\mathrm{rank}^{\varepsilon,p}_{(\Omega,G)}(\rho)} \le \text{puri-rank}^{\varepsilon,p}_{(\Omega,G)}(\rho) \le \text{q-sqrt-rank}^{\varepsilon,p}_{(\Omega,G)}(\rho).$$

Since Corollary 3.19 upper bounds $\text{q-sqrt-rank}^{\varepsilon,p}_{(\Omega,G)}$ by a constant which is independent of the dimension of $\rho \in \mathcal{X}_K$, all these ranks mentioned above are upper bounded on $\mathcal{X}_K$.

(iii)-(iv) Let $\rho \in \mathcal{X}_{\mathrm{sep}}$. Using again [12, Prop. 29] we have that

$$\begin{aligned} \mathrm{rank}^{\varepsilon,p}_{(\Omega,G)}(\rho) &\le \text{sep-rank}^{\varepsilon,p}_{(\Omega,G)}(\rho), \\ \text{puri-rank}^{\varepsilon,p}_{(\Omega,G)}(\rho) &\le \text{sep-rank}^{\varepsilon,p}_{(\Omega,G)}(\rho). \end{aligned}$$

By Proposition 3.20 $\text{sep-rank}_{(\Omega,G)}$ is upper bounded by a constant which is again independent of the dimension of ρ. □

Note that the previous result is not making any statement about the Schatten 1-norm, since the approximate Carathéodory Theorem is not applicable for this case. Using the results from Theorem 3.17, Proposition 3.20 and Equation (3.2), we now give an upper bound for the approximate rank and approximate separable rank in Schatten 1-norm, which is however dimension dependent.

Corollary 3.22. *Let $M \in \mathrm{Her}_{d^{n+1}}$ and $\rho \in SEP_{n,d}$ both be G-invariant. Then:*

$$(i) \quad rank^{\varepsilon,1}_{(\Omega,G)}(M) \le d^{n+1} \cdot rank^{\varepsilon,2}_{(\Omega,G)}(M) \le \left\lceil D_2 \cdot \left(\frac{2\mu_p(M)}{\varepsilon}\right)^2 \right\rceil \cdot |G| \cdot d^{n+1}$$

$$(ii) \quad sep\text{-}rank^{\varepsilon,1}_{(\Omega,G)}(\rho) \le d^{n+1} \cdot sep\text{-}rank^{\varepsilon,2}_{(\Omega,G)}(\rho) \le \lceil D_2 \cdot (2/\varepsilon)^2 \rceil \cdot |G| \cdot d^{n+1}$$

Proof. Let $A, A_k \in \mathcal{M}_{d^{n+1}}, \mathcal{S} \subseteq \mathcal{M}_{d^{n+1}}$ be chosen as in Theorem 3.8. Using Equation (3.2) and optimizing over all valid p-values gives the bound

$$\|A - A_k\|_1 \le \sqrt{d^{n+1}} \cdot \|A - A_k\|_2,$$

or equivalently

$$\|A - A_k\|_1 \le \varepsilon \quad \text{if} \quad k \ge d^{n+1} \cdot \left\lceil D_2 \cdot \left(\frac{\mathrm{diam}(S)}{\varepsilon}\right)^2 \right\rceil.$$

Applying this statement in the proofs of Theorem 3.17 and Proposition 3.20 shows the statement. □

Example 3.23. We now apply Corollary 3.22 to the running examples of Example 2.2 and Example 3.12.

(i) Consider the line $\Omega = \Lambda_n$ of size n and G the trivial group. Then for the approximate operator Schmidt rank in Schatten 1-norm of $\rho \in \mathcal{M}^+_{d^{n+1}}$ it holds that

$$\text{rank}^{\varepsilon,1}_{\Lambda_n}(\rho) \leq \left\lceil D_2 \cdot \left(\frac{2\mu_2(\rho)}{\varepsilon}\right)^2 \right\rceil \cdot d^{n+1}.$$

If $\rho \in \text{SEP}_{n,d}$, then

$$\text{sep-rank}^{\varepsilon,1}_{\Lambda_n}(\rho) \leq \lceil D_2 \cdot (2/\varepsilon)^2 \rceil \cdot d^{n+1}.$$

Note that in the exact case rank_{Λ_n} is also bounded by a constant times d^{n+1} [14, Prop. 49]. In contrast, the upper bound of $\text{sep-rank}_{\Lambda_n}$ in the exact case is linear in $d^{2(n+1)}$. Hence, for fixed ε and sufficiently large dimension d this yields a better upper bound in the approximate case.

(ii) Consider the circle of n elements, $\Omega = \Theta_n$, together with the cyclic group C_n. Further let $\rho \in \mathcal{M}^+_{d^n}$ be C_n-invariant. Since $|C_n| = n$ we obtain

$$\text{rank}^{\varepsilon,1}_{(\Theta_n,C_n)}(\rho) \leq \left\lceil D_2 \cdot \left(\frac{2\mu_2(\rho)}{\varepsilon}\right)^2 \right\rceil \cdot n \cdot d^n$$

and

$$\text{sep-rank}^{\varepsilon,1}_{(\Theta_n,C_n)}(\rho) \leq \lceil D_2 \cdot (2/\varepsilon)^2 \rceil \cdot n \cdot d^n.$$

△

3.2 Nonnegative tensors

In the following, we use the correspondence between nonnegative tensors and diagonal psd matrices given in Equation (2.3) to show that the separation between $\text{nn-rank}_{(\Omega,G)}$, $\text{rank}_{(\Omega,G)}$ and $\text{psd-rank}_{(\Omega,G)}$ of nonnegative tensors vanishes in the approximate case (Corollary 3.25).

As in Section 2.3, we consider the space $\mathcal{K}_{n,d} := \bigotimes_{i=0}^n \mathbb{C}^d$ equipped with the ℓ_p-norm. In contrast to psd matrices, which are equipped with the Schatten p-norm, the disappearance of separations for nonnegative tensors equipped with the ℓ_p-norm can be proven for all $1 < p < \infty$. Recall that Ω is a connected wsc and a G a free group action.

We start with a preliminary lemma, which is a direct consequence of Corollary 2.14.

Lemma 3.24. *Let $M \in \mathcal{K}_{n,d}$ and $p \in [1, \infty)$. Then the following holds:*

$$\begin{array}{lrcl}
(i) & \text{rank}^{\varepsilon,\ell_p}_{(\Omega,G)}(M) & \leq & \text{nn-rank}^{\varepsilon,\ell_p}_{(\Omega,G)}(M) \\
(ii) & \text{psd-rank}^{\varepsilon,\ell_p}_{(\Omega,G)}(M) & \leq & \text{nn-rank}^{\varepsilon,\ell_p}_{(\Omega,G)}(M) \\
(iii) & \text{rank}^{\varepsilon,\ell_p}_{(\Omega,G)}(M) & \leq & \text{psd-rank}^{\varepsilon,\ell_p}_{(\Omega,G)}(M)^2 \\
(iv) & \text{psd-rank}^{\varepsilon,\ell_p}_{(\Omega,G)}(M) & \leq & \text{sqrt-rank}^{\varepsilon,\ell_p}_{(\Omega,G)}(M)
\end{array}$$

Proof. We only show the first statement. The others are similar.

(i) For $M' \in \mathcal{K}_{n,d}$ we denote the corresponding diagonal matrix $\sigma' \in \mathcal{M}_d$. By definition of σ and σ' we have that $||M - M'||_{\ell_p} = \|\sigma - \sigma'\|_p$. Hence, using Corollary 2.14 it follows immediately

$$\begin{aligned} \operatorname{rank}^{\varepsilon,\ell_p}_{(\Omega,G)}(M) &= \min\left\{\operatorname{rank}_{(\Omega,G)}(M') : \|M - M'\|_{\ell_p} \leq \varepsilon, M' \in \mathcal{K}_{n,d}\right\} \\ &\leq \min\left\{\text{nn-rank}_{(\Omega,G)}(M') : \|M - M'\|_{\ell_p} \leq \varepsilon, M' \in \mathcal{K}_{n,d}\right\} \\ &= \text{nn-rank}^{\varepsilon,\ell_p}_{(\Omega,G)}(M). \end{aligned}$$

□

Note that it is not clear whether a similar result to Theorem 2.12 holds for the approximate ranks. This is because the ε-ball of diagonal matrices also contains non-diagonal matrices with a possibly smaller rank than the diagonal matrices.

Corollary 3.25. *Let $\varepsilon > 0$, $n, K \in \mathbb{N}$ and $1 < p < \infty$. We define the set*

$$\mathcal{Y}_K := \{M \in \mathcal{K}_{n,d} : d \in \mathbb{N}, M \text{ nonnegative and } \|M\|_{\ell_p} \leq K\}.$$

Then the following holds on $\mathcal{Y}_K$:

$$\begin{aligned} &(i) \quad \text{rank}^{\varepsilon,\ell_p}_{(\Omega,G)} \not\ll \text{psd-rank}^{\varepsilon,\ell_p}_{(\Omega,G)} \\ &(ii) \quad \text{psd-rank}^{\varepsilon,\ell_p}_{(\Omega,G)} \not\ll \text{nn-rank}^{\varepsilon,\ell_p}_{(\Omega,G)} \end{aligned}$$

Further, we define the set

$$\mathcal{Y}_{\sqrt{},K} := \{M \in \mathcal{K}_{n,d} : d \in \mathbb{N}, M \text{ nonnegative and } \mu_{\sqrt{},p}(\sigma) \leq K\}$$

where σ denotes the corresponding diagonal matrix of M. Then the following holds on $\mathcal{Y}_{\sqrt{},K}$:

$$\begin{aligned} &(iii) \quad \text{psd-rank}^{\varepsilon,\ell_p}_{(\Omega,G)} \not\ll \text{sqrt-rank}^{\varepsilon,\ell_p}_{(\Omega,G)} \\ &(iv) \quad \text{rank}^{\varepsilon,\ell_p}_{(\Omega,G)} \not\ll \text{sqrt-rank}^{\varepsilon,\ell_p}_{(\Omega,G)} \end{aligned}$$

Note that (i) and (ii) imply that $\operatorname{rank}^{\varepsilon,\ell_p}_{(\Omega,G)} \not\ll \text{nn-rank}^{\varepsilon,\ell_p}_{(\Omega,G)}$ on $\mathcal{Y}_K$.

Proof. (i)-(ii) Let $M \in \mathcal{Y}_K$. By Proposition 3.24, we have that $\operatorname{rank}^{\varepsilon,\ell_p}_{(\Omega,G)}(M)$ and $\text{psd-rank}^{\varepsilon,\ell_p}_{(\Omega,G)}(M)$ are upper bounded by $\text{nn-rank}^{\varepsilon,\ell_p}_{(\Omega,G)}(M)$. To obtain an upper bound of $\text{nn-rank}^{\varepsilon,\ell_p}_{(\Omega,G)}(M)$ we restrict to $p = 2$, as the other cases are analogous. We set $M' = M/\|M\|_{\ell_2}$ and $\varepsilon' = \varepsilon/K$. Then the corresponding diagonal matrix σ' is a separable state. Note that any ε'-approximation of M' is immediately an ε-approximation of M. Using

$$P_2' := \left\{\rho^{[0]} \otimes \cdots \otimes \rho^{[n]} : \rho^{[i]} \in \mathcal{M}_d^+ \text{ diagonal and } \|\bigotimes_{i=0}^{n} \rho_i\|_2 \leq 1\right\} \subseteq P_2$$

instead of P_2 in the proof of Theorem 3.17, we obtain that

$$\text{nn-rank}^{\varepsilon,\ell_2}_{(\Omega,G)}(M) \leq \lceil D_2 \cdot (2K/\varepsilon)^2 \rceil \cdot |G|$$

which is independent of the ambient dimension.

Since the Schatten p-norm on the space of diagonal matrices is equivalent to the ℓ_p-norm on the corresponding tensor space, we can apply both Theorem 3.9 and Theorem 3.17 for $1 < p < \infty$ (see Remark 3.10 for details).

(iii)-(iv) Let $M \in \mathcal{Y}_{\sqrt{},K}$. Again, by Proposition 3.24, $\text{psd-rank}^{\varepsilon,p}_{(\Omega,G)}(M)$ is upper bounded by $\text{sqrt-rank}^{\varepsilon,p}_{(\Omega,G)}(M)$. Applying Corollary 3.19 to the corresponding diagonal matrix σ gives an upper bound for $\text{q-sqrt-rank}^{\varepsilon,p}_{(\Omega,G)}(\sigma)$. This is again independent of the ambient dimension by the choice of M. □

Remark 3.26. Similarly to psd matrices (Corollary 3.22) we can upper bound the approximate ranks for nonnegative tensors in the ℓ_1-norm. The correspondence between diagonal psd matrices and nonnegative matrices yields for a nonnegative and G-invariant $M \in \mathcal{K}_{n,d}$ the inequality

$$\text{rank}^{\varepsilon,\ell_1}_{(\Omega,G)}(M) \leq d^{n+1} \cdot \text{rank}^{\varepsilon,\ell_2}_{(\Omega,G)}(M).$$

This is the same d-dependence as the one of the exact Carathéodory Theorem [37]. Similar results appear for all other ranks of nonnegative tensors. △

Example 3.27. We now give some concrete examples of decompositions whose separations between ranks disappear in the approximate case. From now on we consider the space $\mathcal{K}_{1,d} = \mathbb{C}^d \otimes \mathbb{C}^d \cong \mathcal{M}_d$ and Λ_1-decompositions, in particular, decompositions of the form

$$M = \sum_{\alpha=1}^{r} v_\alpha^{[0]} \otimes v_\alpha^{[1]},$$

where $v_\alpha^{[i]} \in \mathbb{C}^d$.

Recall that for M nonnegative, $\text{nn-rank}_{\Lambda_1}(M)$ is the smallest integer r such that there exists a decomposition of M with $v_\alpha^{[i]} \in \mathbb{R}_+^d$ for all $\alpha = 1, \dots, r$ (i.e. vectors with nonnegative entries). Furthermore, $\text{psd-rank}_{\Lambda_1}(M)$ is the smallest integer r such that there exist $A_j^{[i]} \in \mathcal{M}_r^+$ for $j = 1, \dots, d$ and $i = 0, 1$ which generate a psd-factorization (Definition 1.14)

$$M_{kl} = \sum_{\alpha_0,\alpha_1=1}^{r} \left(A_k^{[0]}\right)_{\alpha_0,\alpha_1} \cdot \left(A_l^{[1]}\right)_{\alpha_0,\alpha_1} = \text{tr}\left(\left(A_k^{[0]}\right) \cdot \left(A_l^{[1]}\right)^T\right).$$

Assume $1 < p \leq 4/3$, $p = 2$ or $4 \leq p < \infty$ and $\varepsilon > 0$ fixed. The separations are based on examples mentioned in [19].

(i) Consider the normalized Euclidean distance matrix

$$U_d = M_d / \|M_d\|_{\ell_p} \in \mathcal{M}_d,$$

where M_d is defined as

$$M_d := \left((i-j)^2\right)_{i,j=1}^{d}.$$

It was shown that $\text{sqrt-rank}_{\Lambda_1}(M_d) = \text{psd-rank}_{\Lambda_1}(M_d) = 2$ and

$$\text{nn-rank}_{\Lambda_1}(M_d) \geq \log_2(d).$$

Obviously the same statement is true for U_d. Hence, we obtain

$$\text{sqrt-rank}_{\Lambda_1} \ll \text{nn-rank}_{\Lambda_1} \quad \text{and} \quad \text{psd-rank}_{\Lambda_1} \ll \text{nn-rank}_{\Lambda_1}.$$

Since $\|U_d\|_{\ell_p} = 1$, Corollary 3.25 shows that the separations for this example vanish, i.e. $\text{nn-rank}^{\varepsilon,\ell_p}_{\Lambda_1}(U_d)$ can be upper bounded independently of d.

(ii) Let S_d be the slack matrix of a d-gon in the plane. We define the normalized slack matrix as $V_d := S_d/\|S_d\|_{\ell_p}$. It was shown that $\text{rank}_{\Lambda_1}(S_d) = 3$ and $\text{psd-rank}_{\Lambda_1}(S_d)$ diverges if d goes to infinity [21]. Obviously, the same holds for V_d. Hence

$$\text{rank}_{\Lambda_1} \ll \text{psd-rank}_{\Lambda_1}.$$

Since $\|V_d\|_{\ell_p} = 1$, Corollary 3.25 shows that the separation for this example vanishes, i.e. $\text{psd-rank}^{\varepsilon,\ell_p}_{\Lambda_1}(V_d)$ can be upper bounded independently of d. △

Note that this discussion can also be extended to Λ_n-decompositions with $n \geq 1$. One application fulfilling the normalization condition is, for example, the interpretation of a tensor $M \in \mathcal{K}_{n,d}$ as a probability mass function $P(X_0, \ldots, X_n)$ over $n+1$ discrete random variables taking values in $\{1, \ldots, d\}$,

$$m_{i_0,\ldots,i_n} := P(X_0 = i_0, \ldots, X_n = i_n).$$

In this case M is nonnegative and bounded with $\|M\|_{\ell_p} \leq \|M\|_{\ell_1} = 1$. The nonnegative Λ_n-decomposition corresponds to a *hidden Markov model* [20].

4 Computation of approximate decompositions

4.1 The gradient method for Schatten p- and ℓ_p-norms

The sequence in Theorem 3.8, which approximates a matrix in the given convex hull, can be computed by a deterministic algorithm presented in [26] for uniformly smooth Banach spaces. In the following, we give an explicit description of this algorithm for Schatten p-classes (i.e. spaces equipped with the Schatten p-norm). Note first that for $1 < p < \infty$ the Schatten p-norm is everywhere differentiable except for 0. Let $X, Y \in \mathcal{M}_d \setminus \{0\}$ be two arbitrary matrices. Further, let

$$X = U \cdot \Sigma \cdot V^\dagger$$

be a singular value decomposition, in particular, $U, V \in \mathcal{M}_{d,r}$ are isometries where $r \leq d$ and

$$\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_r) \in \mathcal{M}_r$$

is a diagonal matrix with positive entries. Further, we denote the ith column of U and V by u_i and v_i, respectively. Then, the directional derivative of $\|\cdot\|_p$ at X in direction

Y can be evaluated (see Ref. [36] for details) by

$$\mathcal{D}_Y \| \cdot \|_p \Big|_X = \frac{1}{\|X\|_p^{p-1}} \cdot \sum_{i=1}^{r} \sigma_i^{p-1} \cdot u_i^\dagger \cdot Y \cdot v_i. \tag{3.4}$$

Since $X \cdot v_i = \sigma_i \cdot u_i$, it further holds that $\mathcal{D}_X \| \cdot \|_p \Big|_X = \|X\|_p$. This property allows us to formulate the algorithm to compute the sequence $\{X_i\}_{i=1}^{\infty} \subseteq \mathcal{S}$ of Theorem 3.8.

Algorithm 3.28. The sequence $\{X_i\}_{i=1}^{\infty} \subseteq \mathcal{S}$ which generates the approximation $A_k = \frac{1}{k} \sum_{i=1}^{k} X_k$ of $A \in \text{conv}(\mathcal{S})$ satisfying Theorem 3.8 can be constructed in the following way:

(i) X_1 is an arbitrary point in S.

(ii) For the constructed sequence $\{X_1, \dots, X_k\} \subseteq \mathcal{S}$, $k \geq 1$ we choose $X_{k+1} \in \mathcal{S}$ such that for $Y = X_{k+1} - A$ the following holds

$$\mathcal{D}_Y \| \cdot \|_p \Big|_{A_k - A} \leq 0. \tag{3.5}$$

△

Note that as long as $A \in \text{conv}(\mathcal{S})$, there always exists an $X_{k+1} \in \mathcal{S}$ such that the inequality is satisfied. Moreover, Algorithm 3.28 does not further constrain the choice of X_{k+1}, and hence the upper bound given in Theorem 3.8 is satisfied for any sequence constructed with this algorithm.

To apply this algorithm to the ℓ_p-norm, we only have to replace Equation (3.4) with

$$\mathcal{D}_Y \| \cdot \|_{\ell_p} \Big|_X = \frac{1}{\|X\|_{\ell_p}^{p-1}} \cdot \sum_{i,j=1}^{d} y_{ij} \cdot x_{ij} \cdot |x_{ij}|^{p-2} \tag{3.6}$$

where x_{ij} and y_{ij} the entries of X and Y respectively at position (i, j).

4.2 Application to nonnegative decompositions

We now apply the above algorithm to a nonnegative Λ_1-decomposition on the space $\mathcal{K}_{1,d}$. We set

$$\mathcal{S} = \left\{ e_i \otimes e_j : i, j \in \{1, \dots, d\} \right\}$$

where e_i the standard basis-vectors in $\mathbb{C}^d$. Assume that $A \in \mathcal{K}_{1,d}$ is nonnegative and $\|A\|_{\ell_1} = 1$. Obviously, we have that $A \in \text{conv}(\mathcal{S})$ and the corresponding convex combination is a valid nonnegative Λ_1-decomposition of A.

Note that in step (ii) of the Algorithm 3.28, there is, in general, not a unique choice that satisfies Equation (3.5). Hence, we distinguish between a standard and a greedy method of choice.

Method 1. Define an order on $\mathcal{S}$ and choose the smallest element which satisfies Equation (3.5).

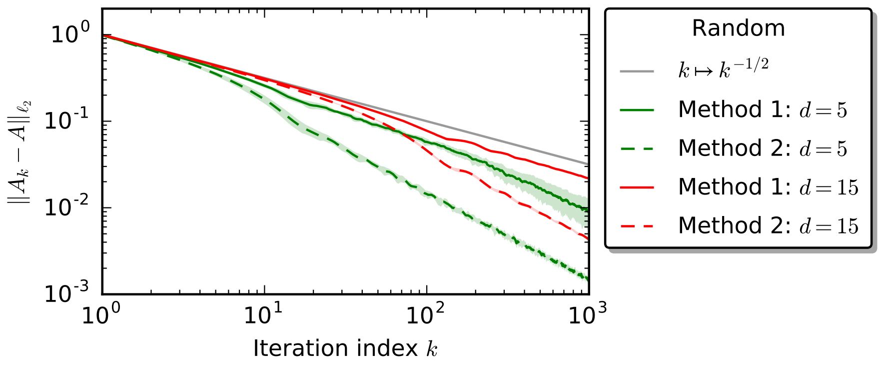


FIGURE 3.3: Application of Algorithm 3.28 to random matrices A at two different dimensions $d = 5$ (green) and $d = 15$ (red). The entries are independently uniformly distributed on $[0, 1]$ with the constraint that $\|A\|_{\ell_1} = 1$. The x-axis shows the iteration index and the y-axis shows the error measured in the ℓ_2-norm. Gray shows the function $k \mapsto k^{-1/2}$ as an orientation for the theoretical convergence rate (up to a constant). Method 1 is plotted as a continuous line, and Method 2 as dashed line. The sampling size for the random matrices is 20, and the plots show the mean value and the standard deviation.

Method 2. Choose the element in $\mathcal{S}$ which attains the smallest value on the left hand side of Equation (3.5).

In the following numerical examples we use for Method 1 the lexicographic ordering

$$(i, j) \preceq (i', j') :\Longleftrightarrow i < i' \text{ or } (i = i' \text{ and } j \leq j').$$

Figure 3.3 shows the application for random matrices with uniformly independently distributed entries normalized to 1. Both methods show a $k^{-1/2}$ convergence for small k and a transition to a faster convergence rate depending on the method. As expected, the greedy type Method 2 converges faster than Method 1. Concerning the choice of the ordering in Method 1, there would be no difference to other orderings since the entries are uniformly independently distributed. The numerical experiments also indicate that the iteration index k where the transition of faster convergence appears grows with increasing dimension d of the matrices.

Figure 3.4 shows the application of the algorithm to random rank-1 matrices. The results are qualitatively similar to the case of random matrices, which have almost surely full rank. This is due to the fact that random rank-1 matrices are almost surely a linear combination of all d^2 elements of S. Hence the algorithm cannot distinguish between random matrices and random rank-1 matrices.

Since the algorithm cannot distinguish matrices with different ranks, Figure 3.5, which shows the application to the Euclidean distance matrix M_d normalized to 1, also shows a similar convergence rate in comparison to Figure 3.3 and Figure 3.4. Note that the fluctuations of the convergence are a natural consequence of the fact that for every iteration step k, the prefactor of the linear combinations $1/k$ is fixed. Since the graphs in Figure 3.3 and Figure 3.4 show an average convergence rate the fluctuations do not appear therein.

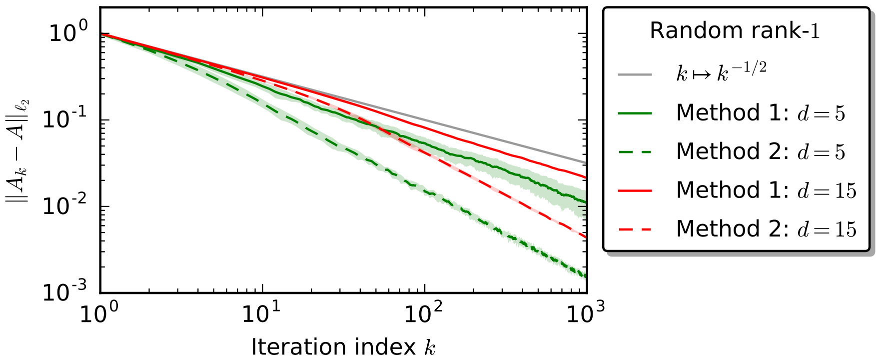


FIGURE 3.4: For the description of the graph we refer to the caption of Figure 3.3. $A = a \cdot b^T$ is a nonnegative random rank-1 matrix with $a, b \in [0, 1]^d$ uniformly distributed and A normalized to 1, i.e. $\|A\|_{\ell_1} = 1$. The sampling size is 20, and the plots show the mean value and the standard deviation.

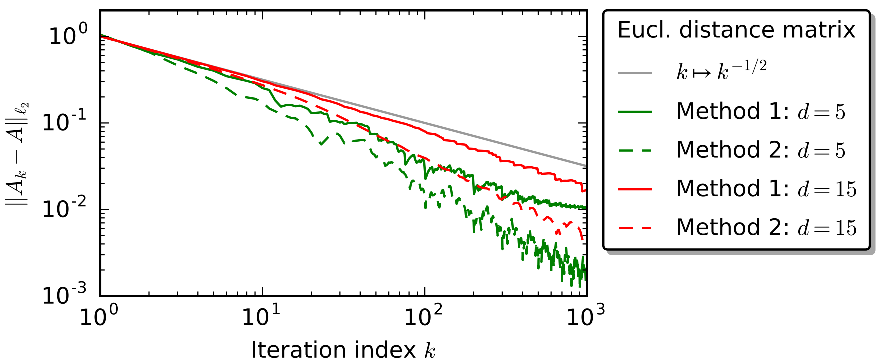


FIGURE 3.5: For the description of the different lines we refer to the caption of Figure 3.3. A is the normalized Euclidean distance matrix $M_d / \|M_d\|_{\ell_1}$.

Chapter 4

Conclusions and Outlook

In this thesis, we have introduced different notions of exact and approximate tensor decompositions in the tensor product of matrix spaces and the tensor product of spaces $\mathbb{C}^d$, respectively. After introducing a particular type of tensor network decompositions, namely the matrix product density operator in Chapter 1, we have used the framework of (Ω, G)-decompositions presented in [12] to define the different notions of approximate (Ω, G)-ranks in Chapter 2. The weighted simplicial complex Ω determines the arrangement of the indices in the sum of elementary tensor factors, and the group action G contains the permutations of tensor product indices which leave the elementary tensors in the (Ω, G)-decomposition invariant. The approximate rank, introduced in Chapter 3, describes the minimal rank among all elements within an ε-ball measured globally with the Schatten p-norm or the ℓ_p-norm.

In Chapter 1, we have shown that decompositions of diagonal positive semidefinite matrices in the bipartite regime correspond to factorizations of nonnegative matrices (Theorem 1.18). This theorem has been generalized to correspondences between arbitrary (Ω, G)-decompositions of diagonal psd matrices and nonnegative tensors in Chapter 2 (Theorem 2.12).

We have further studied in Chapter 1 an application of matrix factorization in the context of correlation complexities (Theorem 1.19 and Theorem 1.22). With this characterization, we have shown that there exists a separation between the random correlation complexity and the quantum correlation complexity (Section 3.3).

In Chapter 2, we have given relations between all defined notions of ranks for psd matrices (Theorem 2.13) and nonnegative tensors (Corollary 2.14). Further, we have studied different separations between all notions of ranks (Theorem 2.16).

In Chapter 3, we have shown that a matrix contained in a convex hull can always be approximated by a dimension-independent number of generators of the convex hull for several Schatten p-norms (Theorem 3.8) and ℓ_p-norms (Remark 3.10). Using these results we have proven the existence of upper bounds for the approximate (Ω, G)-rank (Theorem 3.17), the approximate (Ω, G)-purification rank for psd matrices (Corollary 3.19) and the approximate (Ω, G)-separable rank (Proposition 3.20) for separable states, which are (up to a gauge function defined in Equation (3.3) and the cardinality of the group action G) dimension independent. Using these upper bounds we have shown that many separations between exact (Ω, G)-ranks disappear in the approximate case for psd matrices (Corollary 3.21) and for nonnegative tensors (Corollary 3.25). Finally, we have presented a procedure (Algorithm 3.28) to compute such approximations, attaining the bounds of Theorem 3.8.

The presented results might have several applications in various fields. As we have seen in Chapter 1, there exists a separation between the random and the quantum correlation complexity. Due to the disappearance of the separations in the approximate case also the separation between the correlation complexities would disappear. Another field of application would be the field of tensor network decompositions of mixed states in quantum many-body physics. Although the operator Schmidt decomposition (Definition 1.1) is the most efficient decomposition it is not applicable due to the lack of a local certificate of positivity. Contrary, the purification form (Definition 2.8), which has a local certificate of positivity, can be arbitrarily more costly than the former decomposition. That is, there is a separation between the ranks of the decompositions. The application of the results studied and presented in Chapter 3 might lead to an efficient approximate representation of every mixed state in the local purification form. For this application it would be important to study effective methods to compute upper bounds of our approximate ranks, in particular the gauge function $\mu_{\surd,p}$ for the purification rank. The presented results might also have several implications in the field of tensor decompositions. This is due to the fact that although for matrices, the low-rank approximations are well studied, for tensors of higher order, a best low-rank approximation might not exist [5,16].

An interesting open question is whether one can use the tensor product structure to obtain better upper bounds, especially for the $p = 1$ case. So far, we did not make use of this. Further, it would be interesting to do a similar study measuring the distance of approximation locally. Reference [9] provides local notions of approximations studied in the case of matrix product states.

Bibliography

[1] K. Ball, E. A. Carlen, and E. H. Lieb. Sharp uniform convexity and smoothness inequalities for trace norms. *Inventiones Mathematicae*, 115(1):463–482, 1994.

[2] W. Beckner. Inequalities in Fourier Analysis. *Annals of Mathematics*, 102(1):159–182, 1975.

[3] A. Berman, M. Dür, and N. Shaked-Monderer. Open problems in the theory of completely positive and copositive matrices. *Electronic Journal of Linear Algebra*, 29(1):46–58, 2015.

[4] B. Bollobás. *Modern graph theory*. Springer, 1998.

[5] W. Bruzda, S. Friedland, and K. Życzkowski. Tensor rank and entanglement of pure quantum states. 2019.

[6] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete. Matrix product density operators: Renormalization fixed points and boundary theories. *Annals of Physics*, 378:100–149, 2017.

[7] J. E. Cohen and U. G. Rothblum. Nonnegative ranks, decompositions, and factorizations of nonnegative matrices. *Linear Algebra and Its Applications*, 190(C):149–168, 1993.

[8] P. Comon, G. Golub, L. H. Lim, and B. Mourrain. Symmetric tensors and symmetric tensor rank. *SIAM Journal on Matrix Analysis and Applications*, 30(3):1254–1279, 2008.

[9] A. M. Dalzell and F. G. S. L. Brandão. Locally accurate MPS approximations for ground states of one-dimensional gapped local Hamiltonians. *Quantum*, 3:187, 2019.

[10] R. J. M. Dawson. Homology of weighted simplicial complexes. *Cahiers Topologie Geom. Differentielle Categ.*, 31:229, 1990.

[11] G. De las Cuevas, T. S. Cubitt, J. I. Cirac, M. M. Wolf, and D. Pérez-García. Fundamental limitations in the purifications of tensor networks. *Journal of Mathematical Physics*, 57(7):1–8, 2016.

[12] G. De las Cuevas, M. Hoogsteder Riera, and T. Netzer. Tensor decompositions on simplicial complexes with invariance. 2019.

[13] G. De las Cuevas, A. Klingler, and T. Netzer. Approximate tensor decompositions: disappearance of many separations. 2020.

[14] G. De las Cuevas and T. Netzer. Mixed states in one spatial dimension: decompositions and correspondence with nonnegative matrices. *Journal of Mathematical Physics*, 61(4):1–37, 2020.

[15] G. De las Cuevas, N. Schuch, D. Pérez-García, and J. I. Cirac. Purifications of multipartite states: Limitations and constructive methods. *New Journal of Physics*, 15, 2013.

[16] V. De Silva and L. H. Lim. Tensor rank and the ill-posedness of the best low-rank approximation problem. *SIAM Journal on Matrix Analysis and Applications*, 30(3):1084–1127, 2008.

[17] C. Eckart and G. Young. The approximation of one matrix by another of lower rank. *Psychometrika*, 1(3):211–218, 1936.

[18] J. Eisert, M. Cramer, and M. B. Plenio. Colloquium: Area laws for the entanglement entropy. *Reviews of Modern Physics*, 82(1):277–306, 2010.

[19] H. Fawzi, J. Gouveia, P. A. Parrilo, R. Z. Robinson, and R. R. Thomas. Positive semidefinite rank. *Mathematical Programming*, 153(1):133–177, 2015.

[20] I. Glasser, R. Sweke, N. Pancotti, J. Eisert, and J. I. Cirac. Expressive power of tensor-network factorizations for probabilistic modeling, with applications from hidden Markov models to quantum machine learning. (NeurIPS), 2019.

[21] J. Gouveia, P. A. Parrilo, and R. R. Thomas. Lifts of convex sets and cone factorizations. *Mathematics of Operations Research*, 38(2):248–264, 2013.

[22] S. Gribling, D. de Laat, and M. Laurent. *Lower Bounds on Matrix Factorization Ranks via Noncommutative Polynomial Optimization*, volume 19. Springer US, 2019.

[23] M. B. Hastings. An area law for one-dimensional quantum systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08024–P08024, 2007.

[24] N. J. Higham. *Functions of matrices : theory and computation*. Society for Industrial and Applied Mathematics, 2008.

[25] R. A. Horn and C. R. Johnson. *Matrix Analysis*. 1985.

[26] G. Ivanov. Approximate Carathéodory's Theorem in Uniformly Smooth Banach Spaces. *Discrete and Computational Geometry*, pages 1–6, 2019.

[27] J. Lindenstrauss. On the modulus of smoothness and divergent series in Banach spaces. *The Michigan Mathematical Journal*, 10(3):241–252, 1963.

[28] R. E. Megginson. *An introduction to Banach space theory*, volume 183. 1998.

[29] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information 10th Anniversary Edition*. 2010.

[30] R. Orús. A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Annals of Physics*, 349:117–158, 2014.

[31] S.-J. Ran, E. Tirrito, C. Peng, X. Chen, L. Tagliacozzo, G. Su, and M. Lewenstein. *Tensor Network Contractions*, volume 964. 2020.

[32] R. T. Rockafellar. *Convex Analysis*. Princeton Univ. Press, Princeton, N. J, no. 28 edition, 1970.

[33] F. Verstraete and J. I. Cirac. Matrix product states represent ground states faithfully. *Physical Review B - Condensed Matter and Materials Physics*, 73(9):1–8, 2006.

[34] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac. Matrix product density operators: Simulation of finite-temperature and dissipative systems. *Physical Review Letters*, 93(20):12–15, 2004.

[35] G. Vidal and R. Tarrach. Robustness of entanglement. *Physical Review A - Atomic, Molecular, and Optical Physics*, 59(1):141–155, 1999.

[36] G. A. Watson. Characterization of the subdifferential of some matrix norms. *Linear Algebra and Its Applications*, 170(C):33–45, 1992.

[37] R. Webster. *Convexity*. Oxford University Press, 1994.

[38] M. Wolf. *Quantum Channels and Operations*. unpublished lecture notes, 2012.