Mémoire

Auteur : Louvet, Jonathan
Promoteur(s) : Bastin, Thierry
Faculté : Faculté des Sciences
Diplôme : Master en sciences physiques, à finalité approfondie
Année académique : 2021-2022
URI/URL : http://hdl.handle.net/2268.2/14799

Faculty of Science
Department of physics

Semidefinite optimization for the separability problem

Jonathan Louvet

A dissertation submitted for the degree of Master in Physics

Supervisor
Prof. Thierry Bastin

Reading Committee
Prof. Eric Bousquet
Prof. Geoffroy Lumay
Prof. John Martin
Prof. Pierre Mathonet

Academic Year 2021-2022

Acknowledgements

I would first like to express my very great appreciation to Professor Thierry Bastin, my supervisor, for its valuable guidance in the writing of this manuscript. I highly am grateful for the time he generously invested throughout the year in this work.

I would like to thank Professors Pierre Mathonet, John Martin, Eric Bousquet, and Geoffroy Lumay for being part of my reading committee. I hope they will find this thesis interesting.

I would like to thank my family and friends for their support throughout my years of study and the realization of this manuscript.

Contents

Acknowledgements **i**

Introduction **1**

1 The separability problem **4**
1.1 Quantum mechanics background 4
1.1.1 Hilbert spaces 4
1.1.2 Operator spaces 6
1.1.3 Symmetric states 9
1.1.4 Density operators 10
1.1.5 The separability problem 12
1.2 Bloch representation of states 13
1.2.1 Qubit case 13
1.2.2 Qudit case 14
1.3 Tensorial representation of N-qudit states 15

2 Entanglement and the moment problem **17**
2.1 Algebraic preliminaries 17
2.1.1 Polynomials 17
2.1.2 Matrices 20
2.2 Truncated moment sequence of states 22
2.2.1 Moments and truncated moment sequences 22
2.2.2 N-Qubit case 23
2.2.3 N-qudit case 29
2.3 The $\mathcal{A}$K-truncated moment problem 33
2.3.1 Moment matrices and localizing matrices 33
2.3.2 Shifted moment sequence and localizing matrix 36
2.3.3 The truncated moment problem 38

3 Semidefinite optimization for the separability problem **44**
3.1 Semidefinite optimization for the truncated moment problem 44
3.1.1 Optimization problems 44
3.1.2 Semidefinite programming 48
3.1.3 Moment relaxation for polynomial optimization 50
3.1.4 Semidefinite algorithm for the K-tms problem 52
3.2 Semidefinite optimization for the separability problem 54
3.2.1 Separability problem algorithm 54

3.2.2 Results 56

A Sets, basis, and truncated moment sequence **62**

B Extracting globally optimal solutions **63**

Introduction

In the 1930's, John von Neumann developed a formalism that describes how non-relativistic physical systems behave at the atomic scale and below. This field, known as quantum mechanics, has been ever growing since then. In 1935, Einstein, Podolsky, Rosen [1], and Schrödinger [2] described a "spooky" quantum phenomenon called quantum entanglement. It describes how multipartite quantum states may not always be written as a product of the individual states. In other words, the knowledge of the common system does not infer the knowledge of its individual subsystems. In 1964, Bell tried to quantify this correlation between quantum systems and described how it is impossible for this feature to be simulated in a classical formalism [3]. Quantum entanglement has since been considered as the most distinguishable feature that separates quantum mechanics from classical mechanics. More than just a subject of philosophical discussions, it is a resource, and the key ingredient in applications that cannot be carried out, or very inefficiently, with classical resources, e.g., quantum teleportation [4], and quantum cryptography [5], which, combined with the idea of quantum computation, gave birth to a field called quantum information [6]. Since the discovery of quantum entanglement and its importance for applications, the theoretical description of quantum entanglement has been fast-growing [7, 8, 9]. It revolves around the characterisation and detection of entanglement, that is, if a state is entangled or not, its quantification, and manipulation. The problem of determining whether a quantum state is entangled or separable is called the *separability problem* and is the problem of interest in this work.

Although the characterization and detection of multipartite entanglement remains an open question, the separability problem has been solved for any systems made of 2 qubits or one qubit and one qutrit [10], and for any pure states [11]. For systems made of arbitrary dimensional qudits, the number of variables increases exponentially with the number of subsystems, and the problem becomes disheartening. Restricting the problem to quantum systems whose states are invariant under the permutation of their constituent makes the problem more approachable. Such states are called symmetric states. A criterion on the separability of N symmetric mixed qubits has been found in 2014 [12]. In 2017, a solution for the separability problem was described [13] by mapping it onto a problem in probability theory called the *moment problem*.

The moment problem has been extensively studied in the literature [14]. In probability theory, a probability distribution tells how likely it is for a particular event to happen. There exist many tools that describe the shape of a probability distribution, e.g., its mean. They are called the *moments* of the probability distribution. The moment problem is the inverse problem: given a sequence of moments, the moment problem asks whether there exists a probability distribution (a non-negative measure) that satisfies the given moments. If it exists, the measure is called a representing measure. In the multivariate case, when

all moments are given, the moment problem was solved in 1991 [15]. When the number of moments given is truncated, i.e., finite, the problem is called a *truncated moment problem*. In 2005 Curto and Fialkow presented a necessary and sufficient condition for a truncated moment sequence to admit a representing measure [16]. In 2012, a *semidefinite optimization* algorithm that determines if such a representing measure does exist for a given truncated moment sequence was presented [17], and generalized in 2014 [18].

Optimization problems are widely used in science. Nature tends to optimize: physical systems naturally tend to evolve to a state of minimum energy. Optimization problems consist in the minimization (or maximization) of a function. They can be classified depending on the nature of the function to minimize, and its constraints. Semidefinite optimization, also referred as semidefinite programming, is a convex optimization problem, *i.e.*, a problem whose solution is unique, of a linear function, and where the matrix whose elements are the variables is constrained to be positive semidefinite. Since 1990s, semidefinite programming has been widely used in optimization. It is considered among the most powerful tools in theory and practice. They are commonly utilized in variety of fields, such as approximation algorithms, graph theory, geometry, quantum information and computation [19] [20] [21]. Indeed, the semidefinite and convexity property appears naturally in quantum information. Applications include quantum error correction [22], quantum state discrimination [23], and many others [24].

The use of semidefinite programming for the separability problem was already proposed in a variety of publications. An algorithm presented in [25] detects entanglement but never stops if the state is separable. Contrariwise, the algorithm presented in [26] identifies if the state is separable and never stops if it is entangled. The algorithm presented in [13] provides a certificate of separability and entanglement, and gives a decomposition into product states if the state is separable. It applies to arbitrary quantum states with an arbitrary number of constituents, and arbitrary symmetries between the subparts.

The aim of this work is to give a comprehensive description on how one can map the separability problem onto a moment problem, how to solve a moment problem, and how to implement an algorithm that detects separability and entanglement using semidefinite optimization. More specifically, this work aims to present the following equivalences:

A separable state is a convex combination of product states.

$\Leftrightarrow$ The global expectation values of basis operators can be written as a convex combination of the product of local expectation values of the individual basis operators.

$\Leftrightarrow$ There exists a representing atomic probability measure whose moment sequence of order 1 is given by the local expectation values of the individual basis operators.

$\Leftrightarrow$ There exists a flat extension of the above moment sequence such that its moment matrix and localizing moment matrices are positive semidefinite.

$\Leftrightarrow$ A semidefinite optimization whose variables are the moments of the flat extension above is feasible.

Chapter 1 presents the necessary background in quantum mechanics used throughout this work. The first section presents the basic notions of Hilbert spaces, state vectors, operator spaces, operators and the Generalized Gell-Mann operators, including the equivalence between two operators. Next, symmetric states, density operators, pure and mixed states are exposed, followed by the separability problem in terms of state vectors and density operators. The second section presents the Bloch representation of states for qubits and qudits, followed by the tensorial representation of states, and the equivalence between separability in terms of product states, and in terms of product of local expectation values of the basis operators.

Chapter 2 presents how one can map the separability problem onto a truncated moment problem, and the necessary and sufficient condition to solve a truncated moment problem. The first section presents the basic algebraic notions of monomials, polynomials, rank and flat extension of a matrix. The second section introduces the notions of moments and truncated moment sequences. It is followed by a presentation of how a separable state is equivalent to the existence of a probability measure whose first order moments are given by the local expectation values of the basis operators for qubits, qudits, for general and symmetric states. The rest of the second chapter then presents the notions of moment matrices, localizing matrices, and a description of the truncated moment problem. A necessary and sufficient condition to solve a truncated moment problem is presented, which will lead to a necessary and sufficient condition for the separability of arbitrary states.

Chapter 3 presents a semidefinite optimization algorithm to solve a truncated moment problem, and thus the separability problem. The first section presents the basic notions of optimization problems, convex programming, and linear programming, followed by a description of semidefinite programming. The dual theory of linear programming is then presented, followed by an introduction to the concept of moment relaxation for polynomial optimization, which shows that determining if a representing measure exist amounts to determine if a linear program is feasible. A description of a semidefinite algorithm to solve the truncated moment problem is then presented, which thus solves the separability problem. The second section of chapter 3 then presents results of our implementation of the algorithm.

Chapter 1

The separability problem

The aim of this first chapter is to present the separability problem in quantum mechanics. The first section of this chapter presents a brief mathematical background of quantum mechanics and describes the basic concepts of state spaces, quantum states, symmetric states, linear operators, as well as density operators. The separability problem is then presented for state vectors and density operators. The second section presents the Bloch representation of states, the tensorial representation of states, followed by the equivalence between a convex combination of product states and a convex combination of products of the local expectation values of basis operators. The content of this chapter comes from various sources [6, 13, 27, 28, 29, 30, 31], where more insights and precisions of the concepts presented can be found.

1.1 Quantum mechanics background

1.1.1 Hilbert spaces

The mathematical foundations of quantum mechanics are based on a Hilbert space formalism developed in the 1930s by John von Neumann [32]. A $\mathbb{C}$-Hilbert space, or simply Hilbert space, denoted as $\mathcal{H}$, is a complex vector space with a defined inner product $\langle\cdot|\cdot\rangle$, and a metric induced by the norm defined as $\|\cdot\| = \sqrt{\langle\cdot|\cdot\rangle}$. The elements of $\mathcal{H}$ are complex vectors denoted as $|\cdot\rangle$ and called *ket* vectors. In quantum mechanics, a quantum system, e.g., a particle, is associated to a Hilbert space $\mathcal{H}$ called the *state space*. In this context, a quantum system is completely described by a normalized *state vector* $|\psi\rangle \in \mathcal{H}$, *i.e.*, $\|\psi\|^2 = \langle\psi|\psi\rangle = 1$.

An orthonormal basis of a Hilbert space $\mathcal{H}$ of dimension d is a set

$$\mathcal{B}^d = \{|u_i\rangle \in \mathcal{H}, i \in \{0, \ldots, d-1\} : \langle u_i|u_j\rangle = \delta_{ij}, \forall i, j \in \{0, \ldots, d-1\}\}. \tag{1.1}$$

When the basis vectors $|i\rangle \equiv |u_i\rangle$ are numbered from 0 to $d-1$ the basis

$$\mathcal{B}^d = \{|i\rangle, i \in \{0, \ldots, d-1\}\} \tag{1.2}$$

is referred as a *computational basis*. Any vector $|\psi\rangle \in \mathcal{H}$ can be written as a linear

combination of the elements of a computational basis $\mathcal{B}^d$ as

$$|\psi\rangle = \sum_{i=0}^{d-1} c_i |i\rangle , \tag{1.3}$$

where the $c_i = \langle i|\psi\rangle$ are the coefficients of the decomposition of $|\psi\rangle$ in $\mathcal{B}^d$. It follows that $\sum_{i=0}^{d-1} \langle i|\psi\rangle |i\rangle = \sum_{i=0}^{d-1} |i\rangle \langle i|\psi\rangle = |\psi\rangle$, which leads to the *completeness relation* $\sum_{i=0}^{d_1} |i\rangle \langle i| = \hat{\mathbb{1}}$ where $\hat{\mathbb{1}}$ is the identity operator in $\mathcal{H}$. For a given basis, the coefficients c_i completely characterize the vector $|\psi\rangle$.

A *d-dimensional qudit*, or simply a *qudit*, is a physical system whose state space is $\mathbb{C}^d$ up to an isomorphism. For $d = 2$, the system is called a *qubit*, for $d = 3$, it is called a *qutrit*. For any qudit written in a computational basis as in (1.3), the normalization condition reads

$$\|\psi\|^2 = \langle\psi|\psi\rangle = \sum_{i=0}^{d-1} |c_i|^2 = 1. \tag{1.4}$$

Multipartite systems

Consider a quantum system made of 2 subsystems, called a *bipartite system*. Consider the case of 2 qudits. Let $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(2)}$ be the state space of the first and second subsystem respectively, and let $\{|i_1\rangle , i_1 \in \{0, \ldots, d-1\}\}$, and $\{|i_2\rangle , i_2 \in \{0, \ldots, d-1\}\}$ be a computational basis of $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(2)}$ respectively. The Hilbert state space $\mathcal{H}$ of the system of the two qudits is the *tensor product* of the two subsystems $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(2)}$, that is $\mathcal{H} = \mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)}$. The computational basis of $\mathcal{H}$ of dimension d^2 is given by

$$\mathcal{B}^{d^2} = \{|i_1 i_2\rangle \equiv |i_1\rangle \otimes |i_2\rangle , i_1, i_2 \in \{0, \ldots, d-1\}\} .$$

For instance, for a two-qubit system, given a computational basis $\{|0\rangle , |1\rangle\}$ of $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(2)}$, the computational basis of the state space $\mathcal{H} = \mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)}$ is

$$\mathcal{B}^4 = \{|00\rangle , |01\rangle , |10\rangle , |11\rangle\},$$

and any state vector $|\psi\rangle \in \mathcal{H}$ can be written as

$$|\psi\rangle = \sum_{i_1, i_2=0}^{1} c_{i_1 i_2} |i_1 i_2\rangle = c_{00} |00\rangle + c_{01} |01\rangle + c_{10} |10\rangle + c_{11} |11\rangle , \tag{1.5}$$

such that $\sum_{i_1, i_2=0}^{1} |c_{i_1 i_2}|^2 = |c_{00}|^2 + |c_{01}|^2 + |c_{10}|^2 + |c_{11}|^2 = 1$.

When the system is made of N subsystems, the Hilbert state space is the tensor product of all the subsystem state spaces $\mathcal{H}^{(i)}$:

$$\mathcal{H} = \bigotimes_{i=1}^{N} \mathcal{H}^{(i)} = \mathcal{H}^{(1)} \otimes \cdots \otimes \mathcal{H}^{(N)}. \tag{1.6}$$

This is called a *multipartite* system made of N subsystems. Consider the case of N qudits. Let $\{|i_j\rangle , i_j \in \{0, \ldots, d-1\}\}$, $\forall j = 1, \ldots, N$ be a computational basis of $\mathcal{H}^{(j)}$ respectively. The computational basis $\mathcal{B}^{d^N}$ of $\mathcal{H}$ is given by

$$\mathcal{B}^{d^N} = \{|i_1 \ldots i_N\rangle \equiv |i_1\rangle \otimes \cdots \otimes |i_N\rangle , i_1, \ldots, i_N \in \{0, \ldots, d-1\}\}, \tag{1.7}$$

be the basis of $\mathcal{H}$ of dimension d^N. Any vector $|\psi\rangle \in \mathcal{H}$ can be written in this basis as

$$|\psi\rangle = \sum_{i_1,\ldots,i_N=0}^{d-1} c_{i_1\ldots i_N} |i_1 \ldots i_N\rangle . \tag{1.8}$$

For a given basis, the state $|\psi\rangle$ of the system is entirely described by its coefficients $c_{i_1\ldots i_N}$. For any multipartite system made of N d-dimensional qudits written as in (1.8), the normalization constraint reads

$$\||\psi\|^2 = \sum_{i_1,\ldots,i_2=0}^{d-1} |c_{i_1\ldots i_N}|^2 = 1. \tag{1.9}$$

1.1.2 Operator spaces

A linear operator $\hat{A}$ defined on a finite dimensional Hilbert space $\mathcal{H}$ is an internal linear map acting on the elements $|\psi\rangle$ of $\mathcal{H}$ as $\hat{A} : \mathcal{H} \to \mathcal{H} : |\psi\rangle \to \hat{A}|\psi\rangle , \forall |\psi\rangle \in \mathcal{H}$. One can show that all linear operators defined on a finite-dimensional Hilbert space are bounded. Throughout, $\mathcal{L}(\mathcal{H})$ will denote the complex vector space of all linear operators acting on the Hilbert space $\mathcal{H}$, and $\mathcal{L}^+(\mathcal{H}) \subset \mathcal{L}(\mathcal{H})$ will denote the real vector subspace of all hermitian operators. The operator space $\mathcal{L}(\mathcal{H})$ can be endowed with the scalar product $\left\langle \hat{A} \middle| \hat{B} \right\rangle \equiv \mathrm{Tr}\left(\hat{A}^\dagger \hat{B}\right)$, which simplifies to $\mathrm{Tr}(\hat{A}\hat{B})$, $\forall \hat{A}, \hat{B} \in \mathcal{L}^+(\mathcal{H})$. To every observable in classical mechanics denoted by A, there correspond a linear hermitian operator $\hat{A}$ defined on a Hilbert state space $\mathcal{H}$ whose eigenvalues a_n are associated to eigenstates $|a_n\rangle \in \mathcal{H}$ which form an orthogonal basis of $\mathcal{H}$: $\hat{A}|a_n\rangle = a_n |a_n\rangle$. The matrix that represents the operators $\hat{A}$ with entries $\langle a_n|\hat{A}|a_m\rangle$ is hermitian, that is, $A^\dagger = A$, where $\hat{A}^\dagger$ is the hermitian conjugate of $\hat{A}$.

Consider a d-dimensional system of state space $\mathcal{H}$. In $\mathcal{L}^+(\mathcal{H})$, a basis of operators is given by $\hat{\lambda}_0 = \hat{\mathbb{1}}$ and $\hat{\lambda}_i$ $(i = 1, \ldots, d^2 - 1)$, where the operators $\hat{\lambda}_i$ are the $d^2 - 1$ traceless hermitian generator of the special unitary group $SU(d)$, *i.e.*, $\mathrm{Tr}(\hat{\lambda}_i) = 0$, $\forall i = 1, \ldots, d^2 - 1$. They can be obtained by the *generalized Gell-Mann matrices* (GGM operators) defined as [33] :

- $\frac{d(d-1)}{2}$ symmetric GGM

$$\hat{\lambda}_s^{jk} = |j\rangle\langle k| + |k\rangle\langle j| , 1 \leq j < k \leq d, \tag{1.10}$$

- $\frac{d(d-1)}{2}$ antisymmetric GGM

$$\hat{\lambda}_a^{jk} = -i|j\rangle\langle k| + i|k\rangle\langle j| , 1 \leq j < k \leq d, \tag{1.11}$$

- $d - 1$ diagonal GGM

$$\hat{\lambda}^l = \sqrt{\frac{2}{l(l+1)}} \left(\sum_{j=1}^{l} |j\rangle\langle j| - l\,|l+1\rangle\langle l+1| \right) , 1 \leq l \leq d-1. \tag{1.12}$$

For $d = 2$, one gets 3 generators which correspond to the Pauli matrices

$$\lambda_s^{12} = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \lambda_a^{12} = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \lambda^1 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.13}$$

For $d = 3$, one gets 8 generators $\hat{\lambda}_i$ as follows .

- 3 GGM symmetric matrices

$$\lambda_s^{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_s^{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \lambda_s^{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{1.14}$$

- 3 GGM antisymmetric matrices

$$\lambda_a^{12} = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_a^{13} = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \qquad \lambda_a^{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \tag{1.15}$$

- 2 GGM diagonal matrices

$$\lambda^1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda^2 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \tag{1.16}$$

One has

$$\mathrm{Tr}(\hat{\lambda}_i \hat{\lambda}_j) = \alpha_i \delta_{ij}, \tag{1.17}$$

$\forall i, j = 0, \ldots, d^2 - 1$, with $\alpha_0 = d$ and $\alpha_i = 2$ for $i \neq 0$. Hence the operator basis $\{\hat{\lambda}_i, i = 0, \ldots, d^2 - 1\}$ is orthogonal.

Any operator $\hat{A} \in \mathcal{L}^+(\mathcal{H})$ can be expanded in the orthogonal basis $\{\hat{\lambda}_i\}$ according to

$$\hat{A} = \sum_{i=0}^{d^2-1} \frac{a_i}{\alpha_i} \hat{\lambda}_i \tag{1.18}$$

with

$$a_i = \mathrm{Tr}(\hat{A} \hat{\lambda}_i) \in \mathbb{R}. \tag{1.19}$$

More explicitly, this yields

$$\hat{A} = \frac{1}{d} \mathrm{Tr}(\hat{A}) \hat{\mathbb{1}} + \frac{1}{2} \sum_{i=1}^{d^2-1} a_i \hat{\lambda}_i. \tag{1.20}$$

If $\hat{A}$ is traceless, it thus follows that

$$\hat{A} = \frac{1}{2} \sum_{i=1}^{d^2-1} \mathrm{Tr}(\hat{A} \hat{\lambda}_i) \hat{\lambda}_i. \tag{1.21}$$

The basis expansion (1.18) yields an interesting equality criterion for two operators. For any $\hat{A}, \hat{B} \in \mathcal{L}^+(\mathcal{H})$, one gets

$$\hat{A} = \hat{B} \tag{1.22}$$

$$\Leftrightarrow \qquad \mathrm{Tr}(\hat{A}\hat{\lambda}_i) = \mathrm{Tr}(\hat{B}\hat{\lambda}_i), \quad \forall i = 0, \ldots, d^2 - 1. \tag{1.23}$$

For an N-multipartite system of state space $\mathcal{H} = \bigotimes_{j=1}^{N} \mathcal{H}^{(j)}$, with $\mathcal{H}^{(j)} \simeq \mathbb{C}^{d_j}$, it follows

$$\mathcal{L}^+(\mathcal{H}) = \bigotimes_{j=1}^{N} \mathcal{L}^+(\mathcal{H}^{(j)}) \tag{1.24}$$

and a basis of operators in $\mathcal{L}^+(\mathcal{H})$ is given by

$$\{\hat{\Lambda}_{i_1 \ldots i_N} = \bigotimes_{j=1}^{N} \hat{\lambda}_{i_j}^{(j)}, i_j \in \{0, \ldots, d_j^2 - 1\}, \forall j = 1, \ldots, N\}. \tag{1.25}$$

It forms an orthogonal basis :

$$\begin{aligned}
\mathrm{Tr}\left(\hat{\Lambda}_{i_1 \ldots i_N} \hat{\Lambda}_{i'_1 \ldots i'_N}\right) &= \mathrm{Tr}\left(\bigotimes_{j=1}^{N} \hat{\lambda}_{i_j}^{(j)} \hat{\lambda}_{i'_j}^{(j)}\right) \\
&= \prod_{j=1}^{N} \mathrm{Tr}\left(\hat{\lambda}_{i_j}^{(j)} \hat{\lambda}_{i'_j}^{(j)}\right) \\
&= \prod_{j=1}^{N} \alpha_{i_j} \delta_{i_j i'_j} \\
&= \alpha_{(i_1, \ldots, i_N)} \delta_{(i_1, \ldots, i_N)(i'_1, \ldots, i'_N)}
\end{aligned}$$

with

$$\alpha_{(i_1, \ldots, i_N)} = \prod_{j=1}^{N} \alpha_{i_j}, \tag{1.26}$$

and with $\alpha_{i_j} = d_j$ if $i_j = 0$ and 2 otherwise.

Any operator $\hat{A} \in \mathcal{L}^+(\mathcal{H})$ can be expanded according to

$$\hat{A} = \sum_{i_1, \ldots, i_N} \frac{a_{i_1 \ldots i_N}}{\alpha_{(i_1, \ldots, i_N)}} \hat{\Lambda}_{i_1, \ldots, i_N} \tag{1.27}$$

with $a_{i_1 \ldots i_N} = \mathrm{Tr}\left(\hat{A}\hat{\Lambda}_{i_1 \ldots i_N}\right) \in \mathbb{R}$. More explicitly, this yields

$$\hat{A} = \frac{1}{\prod_{j=1}^{N} d_j} \mathrm{Tr}(\hat{A}) \hat{\mathbb{1}} + \sum_{\substack{i_1, \ldots, i_N \\ (i_1, \ldots, i_N) \neq (0, \ldots, 0)}} \frac{a_{(i_1 \ldots i_N)}}{\alpha_{(i_1 \ldots i_N)}} \hat{\Lambda}_{(i_1, \ldots, i_N)}. \tag{1.28}$$

Two operators $\hat{A}$ and $\hat{B} \in \mathcal{L}^+(\mathcal{H})$ are identical if and only if

$$\mathrm{Tr}\left(\hat{A}\hat{\Lambda}_{i_1 \ldots i_N}\right) = \mathrm{Tr}\left(\hat{B}\hat{\Lambda}_{i_1 \ldots i_N}\right), \quad \forall i_1, \ldots, i_N. \tag{1.29}$$

1.1.3 Symmetric states

Consider a two qubit system whose state $|\psi\rangle \in \mathcal{H}$ is written as (1.5), and let us *permute* the two qubits. The operation that permutes the particle 1 and 2 is called the *permutation operator* and is denoted by $\hat{\Pi}_{12}$: $\hat{\Pi}_{12}|i_1\rangle \otimes |i_2\rangle = |i_2\rangle \otimes |i_1\rangle$. The state $|\psi'\rangle$ of the system after the permutation is

$$|\psi'\rangle = \hat{\Pi}_{12}|\psi\rangle = \sum_{i_1,i_2=0}^{1} c_{i_1 i_2}|i_2\rangle \otimes |i_1\rangle = c_{00}|00\rangle + c_{01}|10\rangle + c_{10}|01\rangle + c_{11}|11\rangle . \quad (1.30)$$

If the state is unchanged after the permutation, *i.e.*, $|\psi'\rangle = |\psi\rangle$, the state $|\psi\rangle$ is said to be *symmetric*, which, in this case, amounts to state that $|\psi\rangle$ is symmetric if $c_{01} = c_{10} \equiv c_S$, that is

$$|\psi_S\rangle = c_{00}|00\rangle + c_S(|01\rangle + |10\rangle) + c_{11}|11\rangle .$$

The state space $\mathcal{H}_S$ of all the symmetric states is a subspace of $\mathcal{H}$ and called the *symmetric subspace*. A basis of $\mathcal{H}_S$ in this case is given by $\mathcal{B}_S^3 = \{|00\rangle , \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), |11\rangle\}$ where $\frac{1}{\sqrt{2}}$ is the normalization constant satisfying equation (1.4). If $|\psi'\rangle = -|\psi\rangle$, the state is said to be *antisymmetric*, *i.e.*, in our case, $|\psi\rangle$ is antisymmetric if $c_{01} = -c_{10} \equiv c_A, c_{00} = -c_{00} = 0$, and $c_{11} = -c_{11} = 0$:

$$|\psi_A\rangle = c_A(|01\rangle - |10\rangle).$$

The state space $\mathcal{H}_A$ of all the antisymmetric states is a subspace of $\mathcal{H}$ and called the *antisymmetric subspace*. A basis of $\mathcal{H}_S$ in this case is given by $\mathcal{B}_A^1 = \{\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)\}$. An interesting basis for $\mathcal{H}$ is the basis made of the different elements of $\mathcal{B}_S^3$ and $\mathcal{B}_A^1$, that is,

$$\mathcal{B}^4 = \{|00\rangle , \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), |11\rangle , \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)\}.$$

In this basis, any state $|\psi\rangle \in \mathcal{H}$ can be written as

$$|\psi\rangle = c_{00}|00\rangle + c_S\frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) + c_{11}|11\rangle + c_A\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle).$$

If $c_A = 0$, $|\psi\rangle$ is symmetric, and if c_A is the only coefficient $\neq 0$, $|\psi\rangle$ is antisymmetric.

For a multipartite systems $\mathcal{H} = \bigotimes_{i=1}^{N} \mathcal{H}^{(i)}$ of N qudits, the permutation operator denoted as $\hat{\Pi}_{ij}$ is the operation that permutes the states of the particle i and j, that is, the subsystem i and j. Permutation operators are hermitian ($\hat{\Pi}_{ij}^\dagger = \hat{\Pi}_{ji} = \hat{\Pi}_{ij}$) which implies that their eigenvalues are real, and unitary ($\hat{\Pi}_{ij}\hat{\Pi}_{ij}^\dagger = \left(\hat{\Pi}_{ij}\right)^2 = \hat{\mathbb{1}}$) which means that their eigenvalues are ± 1. If $\hat{\Pi}_{ij}|\psi\rangle = |\psi\rangle$, *i.e.*, the eigenstate $|\psi\rangle$ is associated to the eigenvalue 1, for all $i, j \in \{1, \ldots, N\}$, that is for all the possible permutations between particles, the state is a symmetric state. Conversely, if $\hat{\Pi}_{ij}|\psi\rangle = -|\psi\rangle$, the eigenstate $|\psi\rangle$ is associated to the eigenvalue -1 for all $i, j \in \{1, \ldots, N\}$, then $|\psi\rangle$ is antisymmetric.

Note that quantum states can be neither symmetric nor antisymmetric. For $d = 2$, a basis of $\mathcal{H}_S$ is given by the states defined as

$$\left|D_N^{(k)}\right\rangle = \mathcal{N}\sum_{\pi}\left|\underbrace{0\ldots0}_{k}\underbrace{1\ldots1}_{N-k}\right\rangle, \quad (1.31)$$

with $N-k$ excitations, $k \in \{0,\ldots,N\}$, and where $\mathcal{N}$ is a normalization constant, and where the sum runs over all the possible permutations π of the subsystems. These states are called the *Dicke states*, and can be generalized for any dimension d [34]. One can *project* any state $|\psi\rangle \in \mathcal{H}$ written in the computational basis onto the symmetric subspace $\mathcal{H}_S$ by using the *projection operator* $\hat{P}_S$ made of the different Dicke states. Projection operators are hermitian operators with the property $\hat{P}^2 = \hat{P}$. For $N=2$ and $d=2$, the projection operator $\hat{P}_S$ is

$$\hat{P}_S = |00\rangle\langle 00| + \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10|) + |11\rangle\langle 11|,$$

and its matrix representation in the computational basis is

$$\hat{P}_S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

1.1.4 Density operators

The mathematical formalism of quantum mechanics presented in the previous sections describes quantum system by using state vectors $|\psi\rangle \in \mathcal{H}$, which can be written in a computational basis as $\sum_{i=0}^{d-1} c_i |i\rangle$ in $\mathcal{H}$ such that $\langle\psi|\psi\rangle = 1$. Another formulation can be done using the operator defined as

$$\hat{\rho} = |\psi\rangle\langle\psi| = \left(\sum_{i=0}^{d-1} c_i |i\rangle\right)\left(\sum_{j=0}^{d-1} c_j^* \langle j|\right) = \sum_{i,j=0}^{d-1} c_i c_j^* |i\rangle\langle j|. \tag{1.32}$$

This operator is called the *density operator*. Although it is mathematically equivalent to the description of a quantum system using state vectors $|\psi\rangle$, they are more convenient to describe certain systems in quantum mechanics, as shown in the next sections. Density operators $\hat{\rho}$ have an hermitian matrix representation called the *density matrix* ρ. One can observe that $\hat{\rho} = |\psi\rangle\langle\psi|$ is the projection operator on the state $|\psi\rangle$: $\hat{\rho}|\phi\rangle = \langle\psi|\phi\rangle|\psi\rangle, \forall |\phi\rangle \in \mathcal{H}$. It thus has the property $\hat{\rho}^2 = \hat{\rho}$. For a qubit system, the density operator reads

$$\hat{\rho} = \sum_{i,j=0}^{1} c_i c_j^* |i\rangle\langle j|,$$

and its density matrix is in the computational basis is

$$\rho = \begin{pmatrix} |c_0|^2 & c_0 c_1^* \\ c_1 c_0^* & |c_1|^2 \end{pmatrix}, \tag{1.33}$$

for $|c_0|^2 + |c_1|^2 = 1$, which can then also be written as $\mathrm{Tr}(\hat{\rho}) = 1$, *i.e.*, the sum of its diagonals elements, its *trace*, has to be unity : $\sum_i \hat{\rho}_{ii} = 1$. Any density operator representing a state in quantum mechanics has to satisfy this condition. Density operators also have the property to be *positive semidefinite*, that is, they are hermitian and all their eigenvalues are ≥ 0. The set of all positive semidefinite operators acting on $\mathcal{H}$ is denoted as $\mathcal{P}(\mathcal{H})$, and $\mathcal{P}(\mathcal{H}) \subset \mathcal{L}^+(\mathcal{H})$.

For a multipartite system made of N qudits, any density operator can be written in the computational basis as

$$\hat{\rho} = \sum_{\substack{i_1,\ldots,i_N=0 \\ j_1,\ldots,j_N=0}}^{d-1} c_{i_1\ldots i_N} c^*_{j_1\ldots j_N} \left|i_1 \ldots i_N\right\rangle \left\langle j_1 \ldots j_N\right| , \tag{1.34}$$

such that $\mathrm{Tr}(\hat{\rho}) = 1$. The mathematical formalism described in the previous sections for state vectors can be rewritten in terms of density operators $\hat{\rho} \in \mathcal{P}(\mathcal{H})$. For instance, one can show that the average value $\langle\psi|\hat{A}|\psi\rangle$ of an operator $\hat{A}$ for a quantum system in the state $|\psi\rangle$ can be rewritten as $\mathrm{Tr}(\hat{A}\hat{\rho}) = \mathrm{Tr}(\hat{\rho}\hat{A})$. Indeed,

$$\begin{aligned}
\left\langle \hat{A} \right\rangle_\psi = \langle\psi|\hat{A}|\psi\rangle = \langle\psi|\hat{\mathbb{1}}\hat{A}\hat{\mathbb{1}}|\psi\rangle &= \sum_{i,j} \langle\psi|i\rangle \ \langle i|\hat{A}|j\rangle \ \langle j|\psi\rangle \\
&= \sum_{i,j} \langle i|\hat{A}|j\rangle \ \langle j|\psi\rangle \ \langle\psi|i\rangle \\
&= \sum_{i,j} \langle i|\hat{A}|j\rangle \ \langle j|\,\hat{\rho}\,|i\rangle \\
&= \sum_{i,j} \hat{A}_{ij}\hat{\rho}_{ji} = \sum_i (\hat{A}\hat{\rho})_{ii} \\
&= \mathrm{Tr}(\hat{A}\hat{\rho}) = \left\langle \hat{A} \right\rangle_{\hat{\rho}} .
\end{aligned} \tag{1.35}$$

Pure and mixed states

The density operator $\hat{\rho}$ is convenient to describe a quantum system whose state is not entirely known in the sense that it is known to have probability p_1 to be in the state $|\psi_1\rangle$, p_2 to be in the state $|\psi_2\rangle, \ldots, p_k$ to be in the state $|\psi_k\rangle$. These states are said to be in a *probability mixture* of states, and described by the operator

$$\hat{\rho} = \sum_k p_k \left|\psi_k\right\rangle \left\langle\psi_k\right| , \tag{1.36}$$

where $\sum_k p_k = 1$. These states are called *mixed states*. When the mixture is made of only one state, $\hat{\rho} = |\psi_1\rangle\langle\psi_1|$, $\hat{\rho}$ is called a *pure state*. Pure states are a particular case of mixed states. The description of pure states in the density operator formalism or in the state vector formalism are equivalent.

Similarly to pure states, mixed states $\hat{\rho}$ are represented by a positive semidefinite matrix and have the property $\mathrm{Tr}(\hat{\rho}) = 1$. However, they do not have the same form as projection operators, which means that for mixed states, $\hat{\rho}^2 \neq \hat{\rho}$. Thus a convenient way to make the distinction between mixed and pure states is the number $\mathrm{Tr}(\hat{\rho}^2)$ called the *purity*. Indeed, if $\hat{\rho}$ is a pure state, then $\mathrm{Tr}(\hat{\rho}^2) = \mathrm{Tr}(\hat{\rho}) = 1$. If $\hat{\rho}$ is a mixed state, one can easily show that $\mathrm{Tr}(\hat{\rho}^2) < 1$. For a multipartite system $\mathcal{H} = \bigotimes_{j=1}^N \mathcal{H}^{(j)}$ made of N qudits, any mixed state $\hat{\rho} \in \mathcal{P}(\mathcal{H})$ can be written in the computational basis $\mathcal{B}^{d^N} = \{|i_1 \ldots i_N\rangle , i_j \in \{0, \ldots, d-1\}, \forall j\}$ as

$$\hat{\rho} = \sum_k p_k \left(\sum_{\substack{i_1,\ldots,i_N=0 \\ j_1,\ldots,j_N=0}}^{d-1} c_{k;i_1\ldots i_N} c^*_{k;j_1\ldots j_N} \left|i_1 \ldots i_N\right\rangle \left\langle j_1 \ldots j_N\right| \right) \tag{1.37}$$

such that $\mathrm{Tr}(\hat{\rho}) = 1$. The average value of an operator $\hat{A}$ for a mixed state $\hat{\rho}$ is

$$\mathrm{Tr}(\hat{A}\hat{\rho}) = \mathrm{Tr}(\sum_k p_k \hat{A} \left|\psi_k\right\rangle \left\langle\psi_k\right|) = \sum_k p_k \,\mathrm{Tr}(\hat{A} \left|\psi_k\right\rangle \left\langle\psi_k\right|) = \sum_k p_k \left\langle A\right\rangle_{\psi_k}. \tag{1.38}$$

1.1.5 The separability problem

Entanglement for pure states

Consider a multipartite system $\mathcal{H} = \bigotimes_{j=1}^{N} \mathcal{H}^{(j)}$ made of N qudit subsystems $\mathcal{H}^{(j)}$. The computational basis of $\mathcal{H}$ is

$$\mathcal{B}^{d^N} = \{\left|i_1 \dots i_N\right\rangle = \left|i_1\right\rangle \otimes \cdots \otimes \left|i_N\right\rangle, i \in \{0, \dots, d-1\}\},$$

for $\left|i_j\right\rangle \in \mathcal{H}^{(j)}, j \in \{1, \dots, N\}$. Any state $\left|\psi\right\rangle \in \mathcal{H}$ can be expanded in this basis as

$$\left|\psi\right\rangle = \sum_{i_1,\dots,i_N=0}^{d-1} c_{i_1 \dots i_N} \left|i_1 \dots i_N\right\rangle.$$

If $\left|\psi\right\rangle$ can be written as

$$\left|\psi\right\rangle = \left|\psi_1\right\rangle \otimes \cdots \otimes \left|\psi_N\right\rangle, \tag{1.39}$$

with $\left|\psi_j\right\rangle \in \mathcal{H}^{(j)}$, $\forall j = 1, \dots, N$, the state $\left|\psi\right\rangle$ is said to be *separable*. Otherwise, the state is said to be *entangled*. If $\left|\psi\right\rangle$ is separable, then

$$\left|\psi\right\rangle = \bigotimes_{j=1}^{N} \left|\psi_j\right\rangle \tag{1.40}$$

$$= \bigotimes_{j=1}^{N} \left(\sum_{i_j=0}^{d-1} c_{i_j}^{(j)} \left|i_j\right\rangle \right), \tag{1.41}$$

and there exist $d \cdot N$ coefficients $c_{i_j}^{(j)}$ for $i_j \in \{0, \dots, d-1\}$, $j = 1, \dots, N$ such that the d^N coefficients $c_{i_1 \dots i_N}$ read

$$c_{i_1 \dots i_N} = \prod_{j=1}^{N} c_{i_j}^{(j)}, \tag{1.42}$$

for all $i_1, \dots, i_N$. The problem of determining if $\left|\psi\right\rangle$ is separable or not is called the *separability problem*.

Entanglement for mixed states

Consider an N-qudit system. If $\hat{\rho}$ can be written as a convex combination of product states $\hat{\rho}_k = \hat{\rho}_k^{(1)} \otimes \cdots \otimes \hat{\rho}_k^{(N)}$, that is

$$\hat{\rho} = \sum_k w_k \left(\hat{\rho}_k^{(1)} \otimes \cdots \otimes \hat{\rho}_k^{(N)} \right) \tag{1.43}$$

for $w_k \in [0, 1]$, and $\sum_k w_k = 1$, then the state is separable. Otherwise, it is entangled.

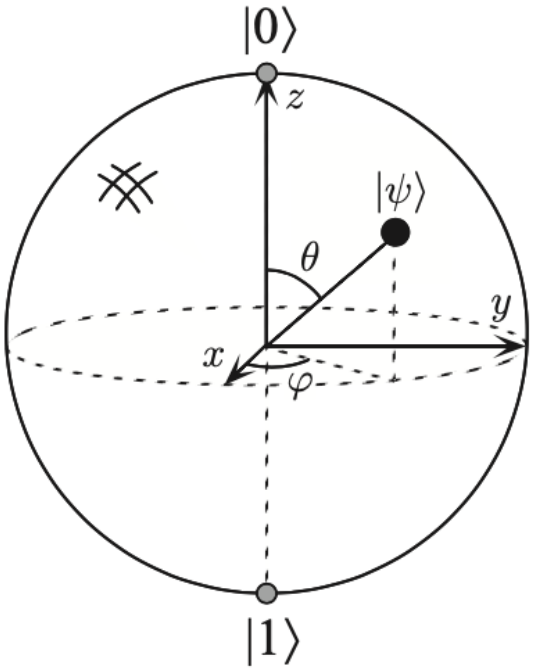

Figure 1.1: Bloch representation of a qubit.

1.2 Bloch representation of states

1.2.1 Qubit case

A convenient geometric picture to represent a qubit state $|\psi\rangle = c_0 |0\rangle + c_1 |1\rangle$ is called the *Bloch representation*. Rewriting $c_0 = e^{i\gamma_0} \cos\left(\frac{\theta}{2}\right)$ and $c_1 = e^{i\gamma_1} \sin\left(\frac{\theta}{2}\right)$ with $\theta \in [0, \pi], \gamma_0, \gamma_1 \in [0, 2\pi[$, for $\gamma_0, \gamma_1, \theta$ real numbers. In this context, $|\psi\rangle$ can be written as

$$\begin{aligned} |\psi\rangle &= e^{i\gamma_0} \cos\left(\frac{\theta}{2}\right) |0\rangle + e^{i\gamma_1} \sin\left(\frac{\theta}{2}\right) |1\rangle \\ &= e^{i\gamma_0} \left(\cos\left(\frac{\theta}{2}\right) |0\rangle + e^{i\phi} \sin\left(\frac{\theta}{2}\right) |1\rangle \right), \end{aligned}$$

where $\phi = \gamma_1 - \gamma_0$. Since quantum states are indistinguishable up to a global phase factor, it follows

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right) |0\rangle + e^{i\phi} \sin\left(\frac{\theta}{2}\right) |1\rangle . \tag{1.44}$$

ϕ and θ define a point on the unit three-dimensional sphere as illustrated on Figure 1.1.

This representation can be generalized for mixed states. Since $\mathrm{Tr}(\hat{\rho}) = 1$, any qubit density operator $\hat{\rho} \in \mathcal{P}(\mathcal{H})$ can be written as [see Eq. (1.20)]

$$\begin{aligned} \hat{\rho} &= \frac{1}{2}\hat{\mathbb{1}} + \frac{1}{2}(\mathbf{b} \cdot \hat{\sigma}) \\ &= \frac{1}{2}\hat{\mathbb{1}} + \frac{1}{2}(b_x \hat{\sigma}_x + b_y \hat{\sigma}_y + b_z \hat{\sigma}_z) \end{aligned}$$

where $\hat{\sigma} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$, with $\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z$ the Pauli operators. They correspond to the GGM operators for $d = 2$, and where $\mathbf{b} = (b_x, b_y, b_z)$, with $b_\alpha = \mathrm{Tr}\,(\hat{\rho}\hat{\sigma}_\alpha) \in \mathbb{R}$ $(\forall \alpha = x, y, z)$. The vector $\mathbf{b}$ is called the *Bloch vector* and entirely describes the state $\hat{\rho}$. In this context, the density matrix ρ in the computational basis reads

$$\rho = \frac{1}{2}\left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b_x \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + b_y \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + b_z \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right]. \tag{1.45}$$

As described before, density matrices are positive semidefinite, *i.e.*, their eigenvalues a are non negative. This leads to a constraint on the Bloch vector $\mathbf{b} = (b_x, b_y, b_z)$. Indeed, from the characteristic polynomial $\det(\hat{\rho} - a\mathbb{1})$, one has

$$\begin{aligned} a &= \frac{1 \pm \sqrt{1 - (1 - b_x^2 - b_y^2 - b_z^2)}}{2} \\ &= \frac{1 \pm \sqrt{|\mathbf{b}^2|}}{2} \\ &= \frac{1 \pm |\mathbf{b}|}{2} \geq 0 \end{aligned}$$

which leads to the constraint $|\mathbf{b}| \leq 1$. If $\hat{\rho}$ is a mixed state, one can show that the purity $\mathrm{Tr}(\hat{\rho}^2) \leq 1$ leads to the same constraint $|\mathbf{b}| \leq 1$, where strict equality corresponds to a pure state.

In summary, the Bloch representation of a qubit state is a vector $\mathbf{b} \in \mathbb{R}^3$ with $|\mathbf{b}| = 1$ for pure states, i.e, on the surface of the Bloch sphere, and with $|\mathbf{b}| \leq 1$ for mixed states, *i.e.*, inside the Bloch sphere.

1.2.2 Qudit case

The Bloch representation can be generalized for qudits. Any qudit state $\hat{\rho} \in \mathcal{P}(\mathcal{H})$ can be written as [see Eq. (1.20)]

$$\hat{\rho} = \frac{1}{d}\hat{\mathbb{1}} + \frac{1}{2}\left(\mathbf{b} \cdot \dot{\lambda}\right) = \frac{1}{d}\hat{\mathbb{1}} + \frac{1}{2}\left(\sum_{i=1}^{d^2-1} b_i \hat{\lambda}_i\right), \tag{1.46}$$

where $\dot{\lambda} = (\hat{\lambda}_1, \ldots, \hat{\lambda}_{d^2-1})$, with $\hat{\lambda}_i, (i = 1, \ldots, d^2 - 1)$ the GGM operators, and where $\mathbf{b} = (b_1, \ldots, b_{d^2-1})$, with $b_i = \mathrm{Tr}(\hat{\rho}\hat{A}_i) \in \mathbb{R}$. The vector $\mathbf{b}$ is called the Bloch vector and entirely describes the state $\hat{\rho}$.

Since $\mathrm{Tr}(\hat{\rho}^2) \leq 1$, the purity for any mixed qudit states reads

$$\begin{aligned} \mathrm{Tr}(\hat{\rho}^2) &= \mathrm{Tr}\left(\left(\frac{1}{d}\hat{\mathbb{1}} + \frac{1}{2}\sum_{i=1}^{d^2-1} b_i \hat{\lambda}_i\right)\left(\frac{1}{d}\hat{\mathbb{1}} + \frac{1}{2}\sum_{i=1}^{d^2-1} b_i \hat{\lambda}_i\right)\right) \\ &= \underbrace{\frac{1}{d^2}\underbrace{\mathrm{Tr}(\hat{\mathbb{1}})}_{d}}_{1/d} + \underbrace{\frac{1}{d}\left(\sum_{i=1}^{d^2-1} b_i \underbrace{\mathrm{Tr}(\hat{\lambda}_i)}_{0}\right)}_{0} + \underbrace{\frac{1}{4}\left(\sum_{i,j=1}^{d^2-1} b_i b_j \underbrace{\mathrm{Tr}(\hat{\lambda}_i\hat{\lambda}_j)}_{2\delta_{ij}}\right)}_{|\mathbf{b}|^2/2} \\ &= \frac{1}{d} + \frac{|\mathbf{b}|^2}{2} \leq 1, \end{aligned}$$

which leads to the constraint

$$|\mathbf{b}| \leq \sqrt{\frac{2(d-1)}{d}}. \tag{1.47}$$

To derive the constraints for the positivity of $\hat{\rho}$, one can express the coefficients a_i of the characteristic polynomial using the Faddeev-LeVerrier algorithm [27]

$$a_{d-m} = -\frac{1}{m}\sum_{k=1}^{m}(-1)^k a_{d-m+k}\,\mathrm{Tr}(\hat{\rho}^k), \tag{1.48}$$

for $m \in \{1,\ldots,d\}$ with $a_d = 1$ and $a_{d-1} = \mathrm{Tr}(\hat{\rho}) = 1$ the normalization constraint. Using Descartes sign rule, $\hat{\rho}$ is positive if the coefficients a_i $i \in \{d-2,\ldots,0\}$ are non-negative. For a qutrit, they read

$$\begin{aligned} a_1 &= \frac{1}{2}\left(1 - \mathrm{Tr}(\hat{\rho}^2)\right) \geq 0, \\ a_0 &= \frac{1}{6}\left(2\,\mathrm{Tr}(\hat{\rho}^3) - 3\,\mathrm{Tr}(\hat{\rho}^2) + 1\right) \geq 0. \end{aligned} \tag{1.49}$$

One can observe that the constraint $a_1 \geq 0$ is the purity. These constraint can also be rewritten in terms of the GGM operators $\hat{\lambda}_i$.

In summary, the Bloch representation of a qudit state is a vector $\mathbf{b} \in \mathbb{R}^{d^2-1}$ with $|\mathbf{b}| = \sqrt{\frac{2(d-1)}{d}}$ for pure states, and with $|\mathbf{b}| \leq \sqrt{\frac{2(d-1)}{d}}$ for mixed states.

Equation (1.46) can also be written in the form

$$\hat{\rho} = \mathcal{N}\sum_{i=0}^{d^2-1} X_i \hat{\lambda}_i \tag{1.50}$$

with $\mathcal{N}$ a normalization constant chosen so as to have $X_0 = 1$, *i.e.*, $\mathcal{N} = 1/d$. It follows that

$$X_i = \frac{d}{2}b_i = \frac{d}{\alpha_i}b_i, \quad \forall i = 1,\ldots,d^2-1. \tag{1.51}$$

The numbers X_i $(i = 0,\ldots,d^2-1)$ are called the *real coordinates* of $\hat{\rho}$.

1.3 Tensorial representation of N-qudit states

Consider a N-qudit state $\hat{\rho}$ acting on $\mathcal{H} = \bigotimes_{j=1}^{N}\mathcal{H}^{(j)}$ with $\mathcal{H}^{(j)} \simeq \mathbb{C}^{d_j}, \forall j$. Any state $\hat{\rho} \in \mathcal{L}^+(\mathcal{H})$ can be written as [see Eq. (1.28)]

$$\hat{\rho} = \frac{1}{\prod_{j=1}^{N} d_j}\hat{\mathbb{1}}^{\otimes N} + \frac{1}{2^N}\left(\sum_{\substack{i_1,\ldots,i_N=0 \\ (i_1,\ldots,i_N)\neq(0,\ldots,0)}} b_{i_1\ldots i_N}\hat{\Lambda}_{i_1\ldots i_N}\right), \tag{1.52}$$

where

$$b_{i_1\ldots i_N} = \frac{2^N}{\alpha_{(i_1,\ldots,i_N)}}\,\mathrm{Tr}\left(\hat{A}\hat{\Lambda}_{i_1\ldots i_N}\right) \tag{1.53}$$

with $\alpha_{(i_1,\ldots,i_N)}$ as defined in Eq. (1.26).

Equation (1.52) can also be written in the form

$$\hat{\rho} = \mathcal{N} \sum_{i_1,\ldots,i_N} X_{i_1\ldots i_N} \hat{\Lambda}_{i_1\ldots i_N} \tag{1.54}$$

with $\mathcal{N}$ a normalization constant chosen so that $X_{0,\ldots 0} = 1$, *i.e.*,

$$\mathcal{N} = \frac{1}{\prod_{j=1}^{N} d_j}. \tag{1.55}$$

This implies

$$X_{i_1,\ldots i_N} = \frac{1}{2^N} \left(\prod_{j=1}^{N} d_j \right) b_{i_1,\ldots i_N} = \left(\prod_{j=1}^{N} \frac{d_j}{\alpha_{i_j}} \right) \mathrm{Tr}(\hat{A}\hat{\Lambda}_{i_1\ldots i_N}), \tag{1.56}$$

for all $i_1, \ldots, i_N : (i_1, \ldots, i_N) \neq (0, \ldots, 0)$. Equation (1.54) is the so-called *tensorial representation* of multipartite N-qudit states.

The tensorial representation allows one to reformulate the separability problem. A state $\hat{\rho} \in \mathcal{P}(\mathcal{H})$ is separable if it can be written in the form

$$\hat{\rho} = \sum_k w_k \left(\hat{\rho}_k^{(1)} \otimes \cdots \otimes \hat{\rho}_k^{(N)} \right) \tag{1.57}$$

for $w_k \in [0, 1]$, and $\sum_k w_k = 1$. Equivalently [see Eq. (1.29)], a state $\hat{\rho}$ is separable if and only if

$$\begin{aligned}
& \mathrm{Tr}\left(\hat{\rho}\hat{\Lambda}_{i_1\ldots i_N}\right) = \mathrm{Tr}\left[\left(\sum_k w_k \hat{\rho}_k^{(1)} \otimes \cdots \otimes \hat{\rho}_k^{(N)} \right) \hat{\Lambda}_{(i_1,\ldots,i_N)} \right], \quad \forall i_1, \ldots i_N \\
\Leftrightarrow \quad & \mathrm{Tr}\left(\hat{\rho}\hat{\Lambda}_{i_1\ldots i_N}\right) = \sum_k w_k \,\mathrm{Tr}\left(\bigotimes_{j=1}^{N} \hat{\rho}_k^{(j)} \lambda_{i_j}^{(j)} \right) \\
& \qquad\qquad = \sum_k w_k \prod_{j=1}^{N} \mathrm{Tr}\left(\hat{\rho}_k^{(j)} \lambda_{i_j}^{(j)}\right), \quad \forall i_1, \ldots, i_N \\
\Leftrightarrow \quad & \left\langle \hat{\Lambda}_{i_1\ldots i_N} \right\rangle_{\hat{\rho}} = \sum_k w_k \prod_{j=1}^{N} \left\langle \hat{\lambda}_{i_j}^{(j)} \right\rangle_{\hat{\rho}_k^{(j)}}, \quad \forall i_1, \ldots, i_N
\end{aligned} \tag{1.58}$$

In other words, $\hat{\rho}$ is separable if the expectation values of all operators $\hat{\Lambda}_{i_1\ldots i_N}$ can be written as a convex combination of products of the local expectation values. Since

$$\left\langle \hat{\lambda}_{i_j}^{(j)} \right\rangle_{\hat{\rho}_k^{(j)}} = \mathrm{Tr}(\hat{\rho}_k^{(j)} \hat{\lambda}_{i_j}^{(j)}) = b_{k;i_j}^{(j)}, \tag{1.59}$$

where $b_{k;0}^{(j)} = 1$, $\mathbf{b}_k^{(j)} = (b_{k;1}^{(j)}, \ldots, b_{k;d^2-1}^{(j)})$ is the Bloch vector of the jth-qubit state $\hat{\rho}_k^{(j)}$, one can write for a separable N-qudit state

$$\left\langle \hat{\Lambda}_{i_1\ldots i_N} \right\rangle_{\hat{\rho}} = \sum_k w_k \prod_{j=1}^{N} b_{k;i_j}^{(j)}. \tag{1.60}$$

or

$$X_{i_1\ldots i_N} = \sum_k w_k \prod_{j=1}^{N} X_{k;i_j} \tag{1.61}$$

with $X_{k;i_j}$ the real coordinates of the state $\hat{\rho}_k^{(j)}$.

Chapter 2

Entanglement and the moment problem

The aim of this chapter is to present the moment problem, and how one can map the separability problem to a truncated moment problem. The first section describes a basic algebraic background on monomials, polynomials, and matrices. The second section presents the concepts of moments, moment sequences, and moment sequences for states. The third section then introduce the notions of moment matrices, shifted moment sequences, and localizing matrices, followed by a presentation of the moment problem and the necessary and sufficient conditions to solve a truncated moment problem. The last section of this chapter presents how one can map the separability problem onto a moment problem, and describes a necessary and sufficient condition for a quantum state to be separable. The content of this chapter comes from various references (mainly [13, 14, 27, 35, 36, 37]).

2.1 Algebraic preliminaries

2.1.1 Polynomials

Monomials

For $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, a *monomial* $m_\alpha(\mathbf{x})$ in the n-tuple $\mathbf{x} \equiv (x_1, \ldots, x_n) \in \mathbb{R}^n$ is a function $\mathbb{R}^n \to \mathbb{R} : \mathbf{x} \to m_\alpha(\mathbf{x})$ defined as

$$m_\alpha(\mathbf{x}) = \mathbf{x}^\alpha \equiv \prod_{i=1}^{n} x_i^{\alpha_i} = x_1^{\alpha_1} \ldots x_n^{\alpha_n}, \tag{2.1}$$

where the integer α_i indicates the degree of x_i in the monomial $\mathbf{x}^\alpha$. The monomial $\mathbf{x}^\alpha$ for which $\alpha = (0, \ldots, 0)$ is $\mathbf{x}^\alpha = 1$. The set of all monomials in n variables is $\mathbb{T}^n \equiv \{\mathbf{x}^\alpha | \alpha \in \mathbb{N}^n\}$. The degree of a monomial is defined as

$$\deg(\mathbf{x}^\alpha) = |\alpha| \equiv \sum_{i=1}^{n} \alpha_i. \tag{2.2}$$

This means that for a given degree $|\alpha|$, there are different n-tuples α, each corresponding to different monomials, for which $|\alpha|$ is the same. If the degree of the monomial in n

variables is of maximum d, that is $|\alpha| \leq d$, there are

$$\binom{n+d}{d} = \frac{(d+1)\dots(d+n)}{n!} \tag{2.3}$$

distinct n-tuples $\alpha = (\alpha_1, \dots \alpha_n)$, each corresponding to distinct monomials $\mathbf{x}^\alpha$ of degree $\leq d$. The set of all these n-tuple forms the set

$$\mathbb{N}^n_d \equiv \{\alpha \in \mathbb{N}^n, |\alpha| = \sum_{i=1}^{n} \alpha_i \leq d\}, \tag{2.4}$$

and the set

$$\mathbb{T}^n_d \equiv \{\mathbf{x}^\alpha | \alpha \in \mathbb{N}^n_d\} \tag{2.5}$$

is the set of all monomial in n variables of degree $\leq d$.

Polynomials

Let $\mathbf{x}^\alpha$ be a monomial in n variables. A real-valued *polynomial* is a *finite* linear combination of monomials :

$$p(\mathbf{x}) = \sum_\alpha p_\alpha \mathbf{x}^\alpha, \tag{2.6}$$

with $p_\alpha \in \mathbb{R}, \forall \alpha$. The set of all real-valued polynomials $p(\mathbf{x})$ forms a real vector space $\mathbb{R}[\mathbf{x}] \equiv \mathbb{R}[x_1, \dots, x_n]$, where $\mathbf{x} \in \mathbb{R}^n$ stands for the n-tuple $(x_1, \dots, x_n)$. For $p_\alpha \neq 0$, $p_\alpha \mathbf{x}^\alpha$ is called a *term of* $p(\mathbf{x})$.

The degree of $p(\mathbf{x})$ is defined as

$$\deg(p(\mathbf{x})) \equiv \max_\alpha \{\deg(\mathbf{x}^\alpha)\}. \tag{2.7}$$

The set of all polynomials of degree $\leq d$ is a vector subspace $\mathbb{R}[\mathbf{x}]_d$ of $\mathbb{R}[\mathbf{x}]$.

The *monomial basis* $\mathcal{B}^n$ is the set made of all monomials $\mathbf{x}^\alpha$ sorted by degree, and within each degree in a lexicographic order,

$$\mathcal{B}^n \equiv \left((1), (x_1, x_2, \dots, x_n), (x_1^2, x_1x_2, x_1x_3, \dots, x_{n-1}x_n, x_n^2), \dots\right). \tag{2.8}$$

It forms a basis of $\mathbb{R}[\mathbf{x}]$. The monomial basis $\mathcal{B}^n_d$ is the set made of all $\binom{n+d}{d}$ monomials in n variables of maximum degree d sorted as

$$\mathcal{B}^n_d \equiv \left((1), (x_1, x_2, \dots, x_n), (x_1^2, x_1x_2, \dots, x_1x_n, x_2x_3, \dots, x_n^2), \dots, (x_1^d, \dots, x_n^d)\right), \tag{2.9}$$

and its dimension is $\binom{n+d}{d}$.

In this context, any polynomial $p(\mathbf{x})$ of degree $\leq d$ can be represented as $p = (p_\alpha)_{\alpha \in \mathbb{N}^n_d}$ which denotes its sequence of coefficients p_α in the monomial basis of dimension $\binom{n+d}{d}$.
A polynomial $p(\mathbf{x})$ is a *sum of square of polynomials* if it can be written as

$$p(\mathbf{x}) = \sum_{j=1}^{m} u_j^2(\mathbf{x}) \tag{2.10}$$

for some polynomials $u_1(\mathbf{x}), \dots, u_m(\mathbf{x})$. One can show that any non-negative quadratic polynomial is a sum of square [14].

Exemple 2.1. Let $n = 3$ with $d = 2$.

- $\mathbb{R}[\mathbf{x}] = \mathbb{R}[x_1, x_2, x_3]$,
- There are $\binom{3+2}{2} = \frac{(2+1)(2+2)(2+3)}{3!} = 10$ distinct triplet $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{N}^3_2$ $= \{(0,0,0), (1,0,0), (0,1,0), (0,0,1), (2,0,0), (1,1,0), (0,1,1), (1,0,1), (0,2,0), (0,0,2)\}$,
- $\mathbf{x}^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha} \in \mathbb{T}^3_2 = \{\mathbf{x}^\alpha | \alpha \in \mathbb{N}^3_2\} = \{1, x_1, x_2, x_3, x_1^2, x_1x_2, x_1x_3, x_2^2, x_2x_3, x_3^2\}$.
- $\mathcal{B}^3_2 = (1, x_1, x_2, x_3, x_1^2, x_1x_2, x_1x_3, x_2^2, x_2x_3, x_3^2)$.

Product of monomials and polynomials

The product of two monomials of degree d $\mathbf{x}^\alpha$ and $\mathbf{x}^\beta$, $\alpha, \beta \in \mathbb{N}^n_d$, is written as

$$\mathbf{x}^\alpha * \mathbf{x}^\beta = \mathbf{x}^{\alpha+\beta}, \tag{2.11}$$

or simply $\mathbf{x}^\alpha \mathbf{x}^\beta$, for all $\alpha, \beta \in \mathbb{N}^n_d$. The product of two polynomials of degree d

$$p(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}^n_d} p_\alpha \mathbf{x}^\alpha,$$

and

$$g(\mathbf{x}) = \sum_{\beta \in \mathbb{N}^n_d} g_\beta \mathbf{x}^\beta,$$

is

$$\begin{aligned}(p * g)(\mathbf{x}) &= p(\mathbf{x})g(\mathbf{x}) \\ &= \sum_{\alpha \in \mathbb{N}^n_d} p_\alpha \mathbf{x}^\alpha \sum_{\beta \in \mathbb{N}^n_d} g_\beta \mathbf{x}^\beta \\ &= \sum_{\alpha,\beta \in \mathbb{N}^n_d} p_\alpha g_\beta \mathbf{x}^\alpha \mathbf{x}^\beta \\ &= \sum_{\gamma \in \mathbb{N}^n_{l_{pg}}} (p * g)_\gamma \mathbf{x}^{\alpha+\beta} \end{aligned} \tag{2.12}$$

for all $\alpha, \beta \in \mathbb{N}^n_d$ and $\gamma \in \mathbb{N}^n_{l_{pg}}$ where $l_{pg} \equiv \deg((p * g)(\mathbf{x}))$ with

$$\deg((p * g)(\mathbf{x})) = \max_{\alpha+\beta} \{\deg(\mathbf{x}^{\alpha+\beta})\},$$

and where $(p * g)_\gamma$ is the sequence of coefficients representing the polynomial $(p * g)(\mathbf{x})$ in the monomial basis $\mathcal{B}^n_{l_{pg}}$.

Exemple 2.2. Consider the polynomials

$$p(\mathbf{x}) = 1 - 3x_2 + 7x_3$$

and

$$g(\mathbf{x}) = 5 + 2x_1 - x_3$$

both of degree $d = 1$. $p(\mathbf{x})$ is represented in $\mathcal{B}_1^3$ by the sequence

$$p = (p_\alpha)_{\alpha \in \mathbb{N}_1^3} = (1, 0, -3, 7),$$

and $g(\mathbf{x})$ s represented in $\mathcal{B}_1^3$ by the sequence

$$g = (g_\beta)_{\beta \in \mathbb{N}_1^3} = (5, 2, 0, -3).$$

The product $(p * g)(\mathbf{x})$ of $p(\mathbf{x})$ and $g(\mathbf{x})$ is then

$$\begin{aligned}(p * g)(\mathbf{x}) &= \sum_{\gamma \in \mathbb{N}^n_{l_{pg}}} (p * g)_\gamma \mathbf{x}^{\alpha+\beta} \\ &= 5 - 15x_2 + 35x_3 + 2x_1 - 6x_1x_2 + 14x_1x_3 - x_3 + 3x_2x_3 - 7x_3^2 \\ &= 5 + 2x_1 - 15x_2 + 34x_3 - 6x_1x_2 + 14x_1x_3 + 3x_2x_3 - 7x_3^2\end{aligned}$$

with

$$\deg(p * g)(\mathbf{x}) = \max_{\alpha+\beta}\{\deg(\mathbf{x}^{\alpha+\beta})\} = 2,$$

and $\gamma \in \mathbb{N}_2^3$. It's sequence of coefficients $(p * g)_\gamma$ in the monomial basis $\mathcal{B}_2^3$ is

$$(5, 2, -15, 34, 0, -6, 14, 0, 3, -7).$$

2.1.2 Matrices

Transpose, symmetry, and trace.

Let $M_n(\mathbb{R})$ be the set of all real $n \times n$ matrices. Consider $M = (m_{ij}) \in M_n(\mathbb{R})$. Throughout, M^T denotes the transpose matrix of M. A matrix is *symmetric* if $M = M^T$. The set of all symmetric matrices in $M_n(\mathbb{R})$ is denoted $\mathcal{S}^n$. The trace of an $n \times n$ matrix is the sum of its diagonal entries:

$$\mathrm{Tr}(M) = \sum_{i=1}^{n} m_{ii}. \tag{2.13}$$

For two matrices $A = (a_{ij})$ and $B = (b_{ij})$, the map $\langle \cdot | \cdot \rangle : M_n(\mathbb{R}) \times M_n(\mathbb{R}) \to \mathbb{R} : (A, B) \to \langle A, B \rangle$ where

$$\langle A, B \rangle = \mathrm{Tr}\left(A^T B\right) = \sum_{i,j=1}^{n} a_{ij} b_{ij} \tag{2.14}$$

defines a scalar product on $M_n(\mathbb{R})$.

Rank

An $m \times n$ matrix $A = (a_{ij})$ can be seen as a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n : \mathbf{x} \to A\mathbf{x}$ for $\mathbf{x} \in \mathbb{R}^m$. The n columns

$$C_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \ldots, C_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \tag{2.15}$$

define the vector subspace $\mathcal{C}(A)$ in $\mathbb{R}^m$ spanned, or generated, by the vectors $(C1, \ldots, C_m)$:

$$\mathcal{C}(A) \equiv \mathcal{L}(C_1, \ldots, C_n) \tag{2.16}$$

that is, made of all the linear combination of the vectors $(C_1, \ldots, C_n)$. $\mathcal{C}(A)$ is called the *column space* of A. Similarly, the m rows of A,

$$\begin{aligned} R_1 &= \begin{pmatrix} a_{11}, \ldots, a_{1n} \end{pmatrix} \\ &\vdots \\ R_m &= \begin{pmatrix} a_{m1}, \ldots, a_{mn} \end{pmatrix} \end{aligned}$$

generate the vector subspace $\mathcal{R}(A) \in \mathbb{R}^n$

$$\mathcal{R}(A) \equiv \mathcal{L}(R_1, \ldots, R_m) \tag{2.17}$$

called the *row space* of A. One can also show that $\dim(\mathcal{R}(A)) = \dim(\mathcal{C}(A)) \equiv r$, which indicates that $\dim(\mathcal{R}(A)) = \dim(\mathcal{C}(A))$ is an integer that *characterises* A. This number is called the *rank* $r \equiv \text{rank}(A)$ of the matrix A, that is the dimension of its column space and row space. Clearly, $\text{rank}(A) \leq \min(m, n)$, and A is said to have *maximal* rank if $\text{rank}(A) = \min(m, n)$. For a square matrix $M = (m_{ij}) \in \mathbb{R}^{n \times n}, \text{rank}(M) \leq \min(n, n) = n$ and M has maximal rank when $\text{rank}\, M = n$.

Positive semidefinite matrices

A symmetric matrix $M \in M_n(\mathbb{R})$ is *positive semidefinite* and denoted $M \succeq 0$ if

$$\mathbf{x}^T M \mathbf{x} \geq 0, \forall \mathbf{x} \neq 0 \in \mathbb{R}^n. \tag{2.18}$$

There are several equivalent characterization: $M \succeq 0$ if and only if any of the 3 following equivalent properties holds.

- $\exists V \in M_n(\mathbb{R})$ such that $M = VV^T$. This decomposition is sometimes known as a *Gram decomposition of M*. V can be chosen in $\mathbb{R}^{n \times r}$ where $r = \text{rank}(M)$.
- $M = (\mathbf{v}_i^T \mathbf{v}_j)_{i,j=1}^n$ for some vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$. The $\mathbf{v}_i$'s may be chosen in $\mathbb{R}^r$, where $r = \text{rank}(M)$.
- All eigenvalues of M are non-negative.

Flat extension of matrices

Let $M \in \mathcal{S}^n$ with block form

$$M = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \tag{2.19}$$

where $A = A^T, C = C^T$. M is said to be a *flat extension* of A if

$$\text{rank}(M) = \text{rank}(A), \tag{2.20}$$

or equivalently if

$$B = AW \tag{2.21}$$

and

$$C = B^T W = W^T AW \tag{2.22}$$

for some matrix W. If M is a flat extension of A then

$$M \succeq 0 \Leftrightarrow A \succeq 0. \tag{2.23}$$

2.2 Truncated moment sequence of states

2.2.1 Moments and truncated moment sequences

Any function $pr(\mathbf{x}) \in [0,1]$ defined $\forall \mathbf{x} \in \mathbb{R}^n$ such that $\int_{\mathbb{R}^n} pr(\mathbf{x})d\mathbf{x} = 1$ can define a function $\mu : K \in \mathbb{R}^n \to [0,1] : \mu(K) \to \int_K pr(\mathbf{x})d\mathbf{x}$, and $pr(\mathbf{x})d\mathbf{x}$ will be referred as a *probability measure supported on* K, or simply a measure on $\mathbb{R}^n$, and denoted $d\mu(\mathbf{x})$.

Moments of order n - univariate case

Consider the univariate case. For a given probability measure $d\mu(x)$ supported on $\mathbb{R}$, its moment of order n is the real quantity defined as

$$y_n \equiv \int_{\mathbb{R}} x^n d\mu(x) \tag{2.24}$$

for $x \in \mathbb{R}$. The moments of a probability measure give informations on the measure. The moment of order 0 of $d\mu(x)$ is

$$y_0 = \int_{\mathbb{R}} x^0 d\mu(x) = \int_{\mathbb{R}} 1 d\mu(x) = 1,$$

since $d\mu(x)$ is a probability measure. It is called the *volume of the measure*. The moment of order 1 of $d\mu(x)$ is

$$y_1 = \int_{\mathbb{R}} x d\mu(x),$$

and is called the *mean* of the measure. The moment of order 2 of $d\mu(x)$ is $y_2 = \int_{\mathbb{R}} x^2 d\mu(x)$, and

$$\int_{\mathbb{R}} x^2 d\mu(x) - \left(\int_{\mathbb{R}} x d\mu(x) \right)^2$$

is called the *variance* of the measure.

Moment of order α - multivariate case

Similarly as in the univariate case, one can define the moment of a measure $d\mu(\mathbf{x})$ supported on $\mathbb{R}^n$ for the multivariate case. The quantity

$$y_\alpha \equiv \int_{\mathbb{R}^n} \mathbf{x}^\alpha d\mu(\mathbf{x}) = \int_{\mathbb{R}^n} x_1^{\alpha_1} \dots x_n^{\alpha_n} d\mu(\mathbf{x}). \tag{2.25}$$

with $\alpha \in \mathbb{N}^n$ and $\mathbf{x}^\alpha \in \mathbb{T}^n$, is called the *moment of order* α of $d\mu(\mathbf{x})$. As an example, for $n = 2$, the moment of order $\alpha = (1, 3)$ of $d\mu(\mathbf{x})$ supported on $\mathbb{R}^2$ is

$$y_{13} = \int_{\mathbb{R}^2} x_1^1 x_2^3 d\mu(\mathbf{x}).$$

Truncated moment sequence

The *moment sequence* y is the (infinite) sequence of numbers $y_\alpha, \forall \alpha \in \mathbb{N}^n$. A *truncated moment sequence* y of order d (tms) is the finite sequence of numbers $y_\alpha, \forall \alpha \in \mathbb{N}^n_d$, that is, the sequence is of all the moments y_α up to order d. In this case, one can also define the column vector $\mathbf{y} = y^T \in \mathbb{R}^{t(d)}$, $t(k) = \binom{n+k}{k}, \forall k \geq 0$.

2.2.2 N-Qubit case

Truncated moment sequence for qubit states

Let us associate to each $\mathcal{H}^{(i)}$ an $\mathbb{R}^3$-variable $\mathbf{x}^{(i)} \equiv (x_j^{(i)})$, $j \in \{1, 2, 3\}$. Any monomial in these variables can be written as

$$(\mathbf{x}^{(i)})^{\alpha^{(i)}} \equiv \prod_{j=1}^{3} (x_j^{(i)})^{\alpha_j^{(i)}}, \tag{2.26}$$

where $\alpha^{(i)} = (\alpha_1^{(i)}, \alpha_2^{(i)}, \alpha_3^{(i)})$.

For a given state $\hat{\rho}^{(i)}$ of the qubit i, one can define the tms $y^{(i)}$ of order 1 $y^{(i)} \equiv (y_{000}^{(i)}, y_{100}^{(i)}, y_{010}^{(i)}, y_{001}^{(i)})$ whose elements are given by the Bloch vector elements of $\hat{\rho^{(i)}}$, that is $y^{(i)} = (b_0^{(i)}, b_1^{(i)}, b_2^{(i)}, b_3^{(i)})$. Explicitly, they read

$$\begin{aligned}
y_{000}^{(i)} &= \int_{K^{(i)}} (\mathbf{x}^{(i)})^{000} d\mu(\mathbf{x}^{(i)}) = \int_{K^{(i)}} d\mu(\mathbf{x}^{(i)}) = b_0^{(i)} = 1, \\
y_{100}^{(i)} &= \int_{K^{(i)}} (\mathbf{x}^{(i)})^{100} d\mu(\mathbf{x}^{(i)}) = \int_{K^{(i)}} x_1^{(i)} d\mu(\mathbf{x}^{(i)}) = b_1^{(i)}, \\
y_{010}^{(i)} &= \int_{K^{(i)}} (\mathbf{x}^{(i)})^{010} d\mu(\mathbf{x}^{(i)}) = \int_{K^{(i)}} x_2^{(i)} d\mu(\mathbf{x}^{(i)}) = b_2^{(i)}, \\
y_{001}^{(i)} &= \int_{K^{(i)}} (\mathbf{x}^{(i)})^{001} d\mu(\mathbf{x}^{(i)}) = \int_{K^{(i)}} x_3^{(i)} d\mu(\mathbf{x}^{(i)}) = b_3^{(i)}
\end{aligned}$$

One can observe that there is a unique correspondence between $\alpha^{(i)} \in \mathbb{N}_1^3$ (the tuples $\alpha^{(i)}$ contain 1 at most once) and $\mu_i, \forall i \in \{1, \dots, N\}$, that is, $\alpha^{(i)}$ is the index such that

$$(\mathbf{x}^{(i)})^{\alpha^{(i)}} = \prod_{\mu=1}^{3} (x_\mu^{(i)})^{\alpha_\mu^{(i)}}, \tag{2.27}$$

since each Bloch vector element $b_\mu^{(i)}$ appears at most once in the expansion

$$\hat{\rho}^{(i)} = \frac{1}{2}\sum_{\mu=0}^{3} b_\mu^{(i)}\hat{\sigma}_\mu^{(i)}.$$

Any state $\hat{\rho}^{(i)} \in \mathcal{P}(\mathcal{H}^{(i)})$ can then be expanded as

$$\hat{\rho}^{(i)} = \frac{1}{2}\left(\sum_{\mu=0}^{3} y_\mu^{(i)}\hat{\sigma}_\mu^{(i)}\right) \tag{2.28}$$

One can construct a 3-dimensional vector $\mathbf{y}^{(i)} \in \mathbb{R}^3$ made of the different moments of the moment sequence $y^{(i)} = (y^{(i)}_{\alpha^{(i)}})_{\alpha^{(i)}\in\mathbb{N}_1^3}$ except $y^{(i)}_{0\ldots0} = 1$,

$$\begin{aligned}\mathbf{y}^{(i)} &= \left(y_1^{(i)}, y_2^{(i)}, y_3^{(i)}\right)\\ &= \left(y_{100}^{(i)}, y_{010}^{(i)}, y_{001}^{(i)}\right)\\ &= \left(b_1^{(i)}, b_2^{(i)}, b_3^{(i)}\right).\end{aligned}$$

Entanglement and the moment problem

Recall that for a separated state $\hat{\rho}$, one has

$$\langle\hat{\sigma}_{\mu_1\ldots\mu_N}\rangle_{\hat{\rho}} = \sum_k w_k b^{(1)}_{k;\mu_1}\ldots b^{(N)}_{k;\mu_N} = \sum_k w_k \prod_{i=1}^{N} b^{(i)}_{k;\mu_i}, \tag{2.29}$$

which can then be rewritten as

$$\langle\hat{\sigma}_{\mu_1\ldots\mu_N}\rangle_{\hat{\rho}} = \sum_k w_k y^{(1)}_{k;\mu_1}\ldots y^{(N)}_{k;\mu_N} = \sum_k w_k \prod_{i=1}^{N} y^{(i)}_{k;\mu_i}. \tag{2.30}$$

Every density operator $\hat{\rho}^{(i)}$ is positive, that is

$$\hat{\rho}^{(i)} = \sum_{\mu_i=0}^{d^2-1} y_{\mu_i}^{(i)}\hat{\sigma}_{\mu_i}^{(i)} \geq 0 \tag{2.31}$$

for all $i \in \{1,\ldots,N\}$. As presented in the previous chapter, it amounts to constraint the coefficients of the characteristic polynomial of each $\hat{\rho}^{(i)}$, given by the Faddeev-LeVerrier algorithm, to be non-negative. One can observe from equation (1.48) that each of these coefficients is a linear combination of traces of power of $\hat{\rho}^{(i)}$, which in turns, from equation (2.31), means that the coefficients are functions of the moments $y_{\mu_i}^{(i)}$, $\mu_i \in \{0,1,2,3\}$, which themselves are functional in the variables $x_j^{(i)}$, $j \in \{1,2,3\}$. These polynomial inequalities define the compact subsets $K^{(i)} \subset \mathbb{R}^{d^2-1}$ to which each vector $\mathbf{y}_k^{(i)} = \left(y_{k;1}^{(i)}, y_{k;2}^{(i)}, y_{k;3}^{(i)}\right)$ is restricted. Positivity on any separable mixed state $\hat{\rho} \in \mathcal{P}(\mathcal{H})$ amounts then to restrict the n-dimensional vectors $\mathbf{y}_k \equiv \left(\mathbf{y}_k^{(1)},\ldots,\mathbf{y}_k^{(N)}\right) \in \mathbb{R}^n$ on the compact $K \subset \mathbb{R}^n$ where $K = K^{(1)} \times \cdots \times K^{(N)}$, with $n = N \cdot 3$.

If equation (2.30) holds, that is for any separable state $\hat{\rho}$, it can always be rewritten in an integral form [38]

$$\langle \hat{\sigma}_{\mu_1 \dots \mu_N} \rangle_{\hat{\rho}} = \int_K (\mathbf{x}^{(1)})^{\alpha^{(1)}} \dots (\mathbf{x}^{(N)})^{\alpha^{(N)}} d\mu(\mathbf{x}) \tag{2.32}$$

$$= \int_K x_{\mu_1}^{(1)} \dots x_{\mu_N}^{(N)} d\mu(\mathbf{x}) \tag{2.33}$$

with $d\mu(\mathbf{x})$ an atomic probability measure supported on K defined by

$$d\mu(\mathbf{x}) = \sum_k w_k \delta(\mathbf{x} - \mathbf{y}_k) d\mathbf{x}, \tag{2.34}$$

and $x_0^{(i)} = 1$, $\mathbf{x} \in \mathbb{R}^n$,

$$\mathbf{x} = (x_1, \dots, x_n) = \left(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N)}\right) \tag{2.35}$$

$(x_j^{(i)}) \in \mathbb{R}^3$ for $j \in \{1, 2, 3\}$ the vector of n variables associated to $\mathcal{H}$. Conversely, one can show that equation (2.33) can be reduced to a finite sum as in equation (2.30) such that the measure $d\mu(\mathbf{x})$ can be written as an atomic measure. This is called *Carathéodory's theorem* [39]

Let us denote $\langle \hat{\sigma}_{\mu_1 \dots \mu_N} \rangle$ as y_α where the tuple α is such that

$$\alpha = \left((\alpha_1^{(1)}, \alpha_2^{(1)}, \alpha_3^3), \dots, (\alpha_1^{(N)}, \alpha_2^{(N)}, \alpha_3^{(N)})\right)$$

where each $\alpha_j^{(j)} \in \mathbb{N}_1^3$,*i.e.*, the N-tuple α is such that each tuple $\alpha_i^{(j)}$ contains 1 at most once. Equation (2.33) can be rewritten as

$$y_\alpha = \int_K \mathbf{x}^\alpha d\mu(\mathbf{x}) \tag{2.36}$$

where α can only takes the value as described above, that is $\alpha \in \mathcal{A} \subset \mathbb{N}_N^3$. A state is then separable if and only if its real coordinates can be written such that (2.36) is met. For (2.36) to holds, it requires the existence of an atomic measure as in (2.34) . In other words, the problem of determining if $\hat{\rho} \in \mathcal{L}^+(\mathcal{H})$ is separable is equivalent to the existence of a probability measure such that (2.36) holds, that is, whose first moments are given by the real coordinates of $\hat{\rho}$, *i.e.*, if (2.32) holds for all $y_\alpha, \alpha \in \mathcal{A}$, for $\mathcal{A} \subset \mathbb{N}_N^3$. The existence of an arbitrary measure such that equation (2.33) holds is equivalent to the existence of an atomic measure of the form (2.34). If there exists a representative measure, then the state can be written as (2.30), and is then separable. The problem of determining whether there exists a probability measure such that (2.36) holds, $\alpha \in \mathcal{A}$ is called an *$\mathcal{A}K$-truncated moment problem.*

Exemple 2.3. Let us consider a two qubit system. The state space $\mathcal{H}$ is the tensor product of the state spaces $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(2)}$ of the individual qubits. A basis of $\mathcal{L}(\mathcal{H}^{(1)})$ is given by $\{\hat{\sigma}_{\mu_1}, \mu_1 = 0, 1, 2, 3\} = \{\hat{\sigma}_0^{(1)} = \hat{\mathbb{1}}, \hat{\sigma}_1^{(1)}, \hat{\sigma}_2^{(1)}, \hat{\sigma}_3^{(1)}\}$ where $\hat{\sigma}_{\mu_i}$ for $\mu_i = 1, 2, 3$ are the Pauli operators. Similarly, a basis of $\mathcal{L}(\mathcal{H}^{(2)})$ is $\{\hat{\sigma}_0^{(2)} = \hat{\mathbb{1}}, \hat{\sigma}_1^{(2)}, \hat{\sigma}_2^{(2)}, \hat{\sigma}_3^{(2)}\}$. Any mixed

states $\hat{\rho}^{(1)} \in \mathcal{L}^+(\mathcal{H}^{(1)})$ and $\hat{\rho}^{(2)} \in \mathcal{L}^+(\mathcal{H}^{(2)})$ can be written in their respective basis as

$$\hat{\rho}^{(1)} = \frac{1}{2}\left(\sum_{\mu_1=0}^{3} b_{\mu_1}^{(1)} \hat{\sigma}_{\mu_1}^{(1)}\right),$$
$$\hat{\rho}^{(2)} = \frac{1}{2}\left(\sum_{\mu_2=0}^{3} b_{\mu_2}^{(2)} \hat{\sigma}_{\mu_2}^{(2)}\right),$$

where $b_{\mu_1}^{(1)} = \mathrm{Tr}(\hat{\rho}^{(1)}\hat{\sigma}_{\mu_1}^{(1)})$ and $b_{\mu_2}^{(2)} = \mathrm{Tr}(\hat{\rho}_k^{(2)}\hat{\sigma}_{\mu_2}^{(2)})$. Let us now associate to $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(2)}$ a set of 3 variables $\mathbf{x}^{(1)} = (x_1^{(1)}, x_2^{(1)}, x_3^{(1)})$ and $\mathbf{x}^{(2)} = (x_1^{(2)}, x_2^{(2)}, x_3^{(2)})$ respectively. For a given mixed state $\hat{\rho}^{(1)} \in \mathcal{L}^+(\mathcal{H}^{(1)})$ and $\hat{\rho}^{(2)} \in \mathcal{L}^+(\mathcal{H}^{(2)})$, one can define the truncated moment sequence $y^{(1)} \equiv (y_{\alpha^{(1)}}^{(1)})_{\alpha^{(1)} \in \mathbb{N}_1^3}$ and $y^{(2)} \equiv (y_{\alpha^{(2)}}^{(2)})_{\alpha^{(2)} \in \mathbb{N}_1^3}$ of each pure state $\hat{\rho}^{(1)}$ and $\hat{\rho}^{(2)}$ respectively. The moments of the sequences are

$$y_{\alpha^{(1)}}^{(1)} = \int_{K^{(1)}} (\mathbf{x}^{(1)})^{\alpha^{(1)}} d\mu(\mathbf{x}^{(1)}) = b_{\mu_1},$$
$$y_{\alpha^{(2)}}^{(2)} = \int_{K^{(2)}} (\mathbf{x}^{(2)})^{\alpha^{(2)}} d\mu(\mathbf{x}^{(2)}) = b_{\mu_2},$$

for $\mu_1 \in \{0,1,2,3\}$, $\alpha^{(1)} = (\alpha_1^{(1)}, \alpha_2^{(1)}, \alpha_3^{(1)})$, $\mu_2 \in \{0,1,2,3\}$, $\alpha^{(2)} = (\alpha_1^{(2)}, \alpha_2^{(2)}, \alpha_3^{(2)})$, where

$$(\mathbf{x}^{(1)})^{\alpha^{(1)}} = (x_1^{(1)})^{\alpha_1^{(1)}} (x_2^{(1)})^{\alpha_2^{(1)}} (x_3^{(1)})^{\alpha_3^{(1)}},$$
$$(\mathbf{x}^{(2)})^{\alpha^{(2)}} = (x_1^{(2)})^{\alpha_1^{(2)}} (x_2^{(2)})^{\alpha_2^{(2)}} (x_3^{(2)})^{\alpha_3^{(1)}},$$

for $d\mu(\mathbf{x}^{(1)})$ and $d\mu(\mathbf{x}^{(2)})$ two probability measure supported on the semi-algebraic sets $K^{(1)}$ and $K^{(2)}$ defined by the polynomial inequalities derived from the positivity constraint of $\hat{\rho}^{(1)}$ and $\hat{\rho}^{(2)}$,

$$\hat{\rho}^{(1)} = \frac{1}{2}\left(\sum_{\mu_1=0}^{3} y_{\mu_1}^{(1)} \hat{\sigma}_{\mu_1}^{(1)}\right) \geq 0 \tag{2.37}$$
$$\hat{\rho}^{(2)} = \frac{1}{2}\left(\sum_{\mu_2=0}^{3} y_{\mu_2}^{(2)} \hat{\sigma}_{\mu_2}^{(2)}\right) \geq 0 \tag{2.38}$$

The positivity constraints can be formulated by using the coefficients of the characteristic polynomial of $\hat{\rho}^{(1)}$ and $\hat{\rho}^{(2)}$, imposing them to be non-negative. They can be obtained using the Faddeev-Leverier algorithm

$$g_{d-m}(\mathbf{x}) = -\frac{1}{m}\sum_{k=1}^{m} (-1)^k g_{d-m+k}(\mathbf{x})\,\mathrm{Tr}((\hat{\rho}^{(i)})^k),$$

for $m = 1, \ldots, d$ with $g_d(\mathbf{x}) = 1$ and $g_{d-1}(\mathbf{x}) = \mathrm{Tr}((\hat{\rho}^{(i)})^k) = 1$. For $\hat{\rho}^{(1)}$, $g_2^{(1)}(\mathbf{x}^{(1)}) = 1$ and

$$g_1^{(1)}(\mathbf{x}^{(1)}) = \mathrm{Tr}((\hat{\rho}^{(1)})) = (x_1^{(1)})^2 + (x_2^{(1)})^2 + (x_3^{(1)})^2 = 1,$$

they read

$$g_0^{(1)}(\mathbf{x}^{(1)}) = (x_1^{(1)})^2 + (x_2^{(1)})^2 + (x_3^{(1)})^2 \leq 1. \tag{2.39}$$

Similarly, for $\hat{\rho}^{(2)}$

$$g_0^{(2)}(\mathbf{x}^{(2)}) = (x_1^{(2)})^2 + (x_2^{(2)})^2 + (x_3^{(2)})^2 \leq 1. \tag{2.40}$$

They define the subsets $K^{(1)}$ and $K^{(2)}$ respectively. The moments of the two moment sequences for $\alpha^{(1)}, \alpha^{(2)} \in \mathbb{N}_1^3$ are

$$\begin{aligned}
y_{000}^{(1)} &= \int_{K^{(1)}} (\mathbf{x}^{(1)})^{000} d\mu(\mathbf{x}^{(1)}) = \int_{K^{(1)}} 1 d\mu(\mathbf{x}^{(1)}) = 1 = b_0^{(1)} \\
y_{100}^{(1)} &= \int_{K^{(1)}} (\mathbf{x}^{(1)})^{100} d\mu(\mathbf{x}^{(1)}) = \int_{K^{(1)}} x_1^{(1)} d\mu(\mathbf{x}^{(1)}) = b_1^{(1)}, \\
y_{010}^{(1)} &= \int_{K^{(1)}} (\mathbf{x}^{(1)})^{010} d\mu(\mathbf{x}^{(1)}) = \int_{K^{(1)}} x_2^{(1)} d\mu(\mathbf{x}^{(1)}) = b_2^{(1)}, \\
y_{001}^{(1)} &= \int_{K^{(1)}} (\mathbf{x}^{(1)})^{001} d\mu(\mathbf{x}^{(1)}) = \int_{K^{(1)}} x_3^{(1)} d\mu(\mathbf{x}^{(1)}) = b_3^{(1)},
\end{aligned}$$

and

$$\begin{aligned}
y_{000}^{(2)} &= \int_{K^{(2)}} (\mathbf{x}^{(2)})^{000} d\mu(\mathbf{x}^{(2)}) = \int_{K^{(2)}} 1 d\mu(\mathbf{x}^{(2)}) = 1 = b_0^{(2)} \\
y_{100}^{(2)} &= \int_{K^{(2)}} (\mathbf{x}^{(12)})^{100} d\mu(\mathbf{x}^{(2)}) = \int_{K^{(2)}} x_1^{(2)} d\mu(\mathbf{x}^{(2)}) = b_1^{(2)}, \\
y_{010}^{(2)} &= \int_{K^{(2)}} (\mathbf{x}^{(12)})^{010} d\mu(\mathbf{x}^{(2)}) = \int_{K^{(2)}} x_2^{(2)} d\mu(\mathbf{x}^{(2)}) = b_2^{(2)}, \\
y_{001}^{(2)} &= \int_{K^{(2)}} (\mathbf{x}^{(12)})^{001} d\mu(\mathbf{x}^{(2)}) = \int_{K^{(2)}} x_3^{(2)} d\mu(\mathbf{x}^{(2)}) = b_3^{(2)},
\end{aligned}$$

The basis of $\mathcal{L}^+(\mathcal{H})$ is $\{\hat{\sigma}_{\mu_1\mu_2} = \hat{\sigma}_{\mu_1}^{(1)} \otimes \hat{\sigma}_{\mu_2}^{(2)}, \mu_1, \mu_2 \in \{0, \ldots, 3\}\}$. For any separable state $\hat{\rho} \in \mathcal{P}(\mathcal{H})$, one can write

$$\langle \hat{\sigma}_{\mu_1\mu_2} \rangle_{\hat{\rho}} = \sum_k w_k \left(y_{k;\mu_1}^{(1)} y_{k;\mu_2}^{(2)} \right), \tag{2.41}$$

which can be rewritten as

$$\begin{aligned}
y_\alpha = \langle \hat{\sigma}_{\mu_1\mu_2} \rangle_{\hat{\rho}} &= \int_K (\mathbf{x}^{(1)})^{\alpha^{(1)}} (\mathbf{x}^{(2)})^{\alpha^{(2)}} d\mu(\mathbf{x}) \\
&= \int_K x_{\mu_1}^{(1)} x_{\mu_2}^{(2)} d\mu(\mathbf{x}) \\
&= \int_K \mathbf{x}^\alpha d\mu(\mathbf{x}),
\end{aligned} \tag{2.42}$$

with $x_0^{(i)} = 1$, $\mathbf{x} \in \mathbb{R}^6$,

$$\begin{aligned}
\mathbf{x} &= \left(\mathbf{x}^{(1)}, \mathbf{x}^{(2)} \right) \\
&= \left((x_1^{(1)}, x_2^{(1)}, x_3^{(1)}), (x_1^{(2)}, x_2^{(2)}, x_3^{(2)}) \right) \\
&= (x_1, x_2, x_3, x_4, x_5, x_6),
\end{aligned}$$

and $d\mu(\mathbf{x})$ a probability measure supported on $K = K^{(1)} \times K^{(2)}$, that is,

$$\begin{aligned} K &= \{\mathbf{x} \in \mathbb{R}^6 | g_0^{(1)}(\mathbf{x}^{(1)}) = 1 - (x_1^{(1)})^2 - (x_2^{(1)})^2 - (x_3^{(1)})^2 \geq 0, \\ &\quad g_0^{(2)}(\mathbf{x}^{(2)}) = 1 - (x_1^{(2)})^2 - (x_2^{(2)})^2 - (x_3^{(2)})^2 \geq 0\} \\ &= \{\mathbf{x} \in \mathbb{R}^6 | g_0^{(1)}(\mathbf{x}) = 1 - x_1^2 - x_2^2 - x_3^2 \geq 0, \\ &\quad g_0^{(2)}(\mathbf{x}) = 1 - x_4^2 - x_5^2 - x_6^2 \geq 0\}, \end{aligned}$$

and $d\mu(\mathbf{x})$ defined by

$$d\mu(\mathbf{x}) = \sum_k w_k \delta(\mathbf{x} - \mathbf{y}_k) d\mathbf{x}. \tag{2.43}$$

The problem of of determining if $\hat{\rho} \in \mathcal{L}^+(\mathcal{H})$ is separable is equivalent to an $\mathcal{A}K$-tms problem, that is, to determine if there exists a representative measure $d\mu(\mathbf{x})$ such that (2.42) holds for all $y_\alpha, \alpha \in \mathcal{A}$, for $\mathcal{A} \subset \mathbb{N}_2^3$ such that

$$\begin{aligned} \alpha &= (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6) \\ &= \left((\alpha_1^{(1)}, \alpha_2^{(1)}, \alpha_3^{(1)},), (\alpha_1^{(2)}, \alpha_2^{(2)}, \alpha_3^{(2)})\right) \end{aligned}$$

where each $\alpha_j^{(i)} \in \mathbb{N}_1^3$, *i.e.*, the 2-tuple α is such that each tuple $\alpha_i^{(j)}$ contains 1 at most once.

Explicitly, $\hat{\rho}$ is separable if and only if there exists a representing measure $d\mu(\mathbf{x})$ supported on K such that

$$\begin{aligned} \langle \hat{\sigma}_{10} \rangle_{\hat{\rho}} &= y_{100000} = \int_K x_1 d\mu(\mathbf{x}) = \int_K x_1^{(1)} d\mu(\mathbf{x}) = b_1^{(1)} \\ \langle \hat{\sigma}_{20} \rangle_{\hat{\rho}} &= y_{010000} = \int_K x_2 d\mu(\mathbf{x}) = \int_K x_2^{(1)} d\mu(\mathbf{x}) = b_2^{(1)} \\ \langle \hat{\sigma}_{30} \rangle_{\hat{\rho}} &= y_{001000} = \int_K x_3 d\mu(\mathbf{x}) = \int_K x_3^{(1)} d\mu(\mathbf{x}) = b_3^{(1)} \\ \langle \hat{\sigma}_{01} \rangle_{\hat{\rho}} &= y_{000100} = \int_K x_4 d\mu(\mathbf{x}) = \int_K x_1^{(2)} d\mu(\mathbf{x}) = b_1^{(2)} \\ \langle \hat{\sigma}_{02} \rangle_{\hat{\rho}} &= y_{000010} = \int_K x_5 d\mu(\mathbf{x}) = \int_K x_2^{(2)} d\mu(\mathbf{x}) = b_2^{(2)} \\ \langle \hat{\sigma}_{03} \rangle_{\hat{\rho}} &= y_{000001} = \int_K x_6 d\mu(\mathbf{x}) = \int_K x_3^{(2)} d\mu(\mathbf{x}) = b_3^{(2)} \\ \langle \hat{\sigma}_{11} \rangle_{\hat{\rho}} &= y_{100100} = \int_K x_1 x_4 d\mu(\mathbf{x}) = \int_K x_1^{(1)} x_1^{(2)} d\mu(\mathbf{x}) = b_1^{(1)} b_1^{(2)} \\ \langle \hat{\sigma}_{12} \rangle_{\hat{\rho}} &= y_{100010} = \int_K x_1 x_5 d\mu(\mathbf{x}) = \int_K x_1^{(1)} x_2^{(2)} d\mu(\mathbf{x}) = b_1^{(1)} b_2^{(2)} \\ \langle \hat{\sigma}_{13} \rangle_{\hat{\rho}} &= y_{100001} = \int_K x_1 x_6 d\mu(\mathbf{x}) = \int_K x_1^{(1)} x_3^{(2)} d\mu(\mathbf{x}) = b_1^{(1)} b_3^{(2)} \\ \langle \hat{\sigma}_{21} \rangle_{\hat{\rho}} &= y_{010100} = \int_K x_2 x_4 d\mu(\mathbf{x}) = \int_K x_2^{(1)} x_1^{(2)} d\mu(\mathbf{x}) = b_2^{(1)} b_1^{(2)} \\ \langle \hat{\sigma}_{22} \rangle_{\hat{\rho}} &= y_{010010} = \int_K x_2 x_5 d\mu(\mathbf{x}) = \int_K x_2^{(1)} x_2^{(2)} d\mu(\mathbf{x}) = b_2^{(1)} b_2^{(2)} \\ \langle \hat{\sigma}_{23} \rangle_{\hat{\rho}} &= y_{010001} = \int_K x_2 x_6 d\mu(\mathbf{x}) = \int_K x_2^{(1)} x_3^{(2)} d\mu(\mathbf{x}) = b_2^{(1)} b_3^{(2)} \\ \langle \hat{\sigma}_{31} \rangle_{\hat{\rho}} &= y_{001100} = \int_K x_3 x_4 d\mu(\mathbf{x}) = \int_K x_3^{(1)} x_1^{(2)} d\mu(\mathbf{x}) = b_3^{(1)} b_1^{(2)} \end{aligned}$$

$$\langle\hat{\sigma}_{32}\rangle_{\hat{\rho}} = y_{001010} = \int_K x_3 x_5 d\mu(\mathbf{x}) = \int_K x_3^{(1)} x_2^{(2)} d\mu(\mathbf{x}) = b_3^{(1)} b_2^{(2)}$$
$$\langle\hat{\sigma}_{33}\rangle_{\hat{\rho}} = y_{001001} = \int_K x_3 x_6 d\mu(\mathbf{x}) = \int_K x_3^{(1)} x_3^{(2)} d\mu(\mathbf{x}) = b_3^{(1)} b_3^{(2)}$$

If such a measure exists, *i.e.*, such that the integrals above holds, then

$$\langle\hat{\sigma}_{\mu_1\mu_2}\rangle_{\hat{\rho}} = \sum_k w_k \left(y_{k;\mu_1}^{(1)} y_{k;\mu_2}^{(2)}\right), \tag{2.44}$$

which means that the state is separable.

2.2.3 N-qudit case

Truncated moment sequence for qudit states

Let us associate to each $\mathcal{H}^{(i)}$ an $\mathbb{R}^{t^{(i)}}$-variable $\mathbf{x}^{(i)} \equiv (x_j^{(i)})$, $j \in \{1, \ldots, t^{(i)}\}$. Any monomial in these variables can be written as

$$(\mathbf{x}^{(i)})^{\alpha^{(i)}} \equiv \prod_{j=1}^{t^{(i)}} (x_j^{(i)})^{\alpha_j^{(i)}}, \tag{2.45}$$

where $\alpha^{(i)} = (\alpha_1^{(i)}, \ldots, \alpha_{t^{(i)}}^{(i)})$. For a given state $\hat{\rho}^{(i)}$ of the qudit i, one can define the tms $y^{(i)}$ of order 1, *i.e.*, $y^{(i)} \equiv (y_{\alpha^{(i)}}^{(i)})_{\alpha^{(i)} \in \mathbb{N}_1^{t^{(i)}}}$ whose elements are given by the real coordinates $X_\mu^{(i)}$ of $\hat{\rho}^{(i)}$. Explicitly,

$$y_{\alpha^{(i)}}^{(i)} = \int_{K^{(i)}} (\mathbf{x}^{(i)})^{\alpha^{(i)}} d\mu(\mathbf{x}^{(i)}) \tag{2.46}$$

There is a unique correspondence between $\alpha^{(i)} \in \mathbb{N}_1^{t^{(i)}}$ (the tuples $\alpha^{(i)}$ contain 1 at most once) and $\mu_i, \forall i \in \{1, \ldots, N\}$, that is, $\alpha^{(i)}$ is the index such that

$$(\mathbf{x}^{(i)})^{\alpha^{(i)}} = \prod_{\mu=1}^{t^{(i)}} (x_\mu^{(i)})^{\alpha_\mu^{(i)}}. \tag{2.47}$$

since each real coordinate $X_\mu^{(i)}$ appears once and alone in the expansion

$$\hat{\rho}^{(i)} = \mathcal{N}^{(i)} \sum_{\mu=0}^{t^{(i)}} X_\mu^{(i)} \hat{\lambda}_\mu^{(i)}.$$

Any state $\hat{\rho}^{(i)} \in \mathcal{P}(\mathcal{H}^{(i)})$ can then be expanded as

$$\hat{\rho}^{(i)} = \mathcal{N}^{(i)} \left(\sum_{\mu=0}^{t^{(i)}} y_\mu^{(i)} \hat{\lambda}_\mu^{(i)} \right) \tag{2.48}$$

One can construct a $t^{(i)}$-dimensional vector $\mathbf{y}^{(i)} \in \mathbb{R}^{t^{(i)}}$ made of the different moments of the moment sequence $y^{(i)} = (y_{\alpha^{(i)}}^{(i)})_{\alpha^{(i)} \in \mathbb{N}_1^{t^{(i)}}}$ except $y_{0\ldots0}^{(i)} = 1$, that is

$$\begin{aligned} \mathbf{y}^{(i)} &= \left(y_1^{(i)}, \ldots, y_{t^{(i)}}^{(i)}\right) \\ &= \left(X_1^{(i)}, \ldots, X_{t^{(i)}}^{(i)}\right). \end{aligned}$$

Entanglement and the moment problem

Recall that for a separable state $\hat{\rho}$, one has

$$\left\langle \hat{\Lambda}_{\mu_1 \dots \mu_N} \right\rangle_{\hat{\rho}} = \sum_k w_k \left(X^{(1)}_{k;\mu_1} \dots X^{(N)}_{k;\mu_N} \right) = \sum_k w_k \left(\prod_{i=1}^{N} X^{(i)}_{k;\mu_i} \right). \tag{2.49}$$

which can then be rewritten as

$$\left\langle \hat{\Lambda}_{\mu_1 \dots \mu_N} \right\rangle_{\hat{\rho}} = \sum_k w_k \left(y^{(1)}_{k;\mu_1} \dots y^{(N)}_{k;\mu_N} \right) = \sum_k w_k \left(\prod_{i=1}^{N} y^{(i)}_{k;\mu_i} \right). \tag{2.50}$$

Every density operator $\hat{\rho}^{(i)}$ is positive, that is

$$\hat{\rho}^{(i)} = \sum_{\mu_i=0}^{t^{(i)}} y^{(i)}_{\mu_i} \hat{\Lambda}^{(i)}_{\mu_i} \geq 0 \tag{2.51}$$

for all $i \in \{1, \dots, N\}$. Similarly as presented for the qubit case, the polynomial inequalities given by the Faddeev-Leverrier algorithm define the compact subsets $K^{(i)} \subset \mathbb{R}^{t^{(i)}}$ to which each vector $\mathbf{y}^{(i)}_k = \left(y^{(i)}_{k;1}, \dots, y^{(i)}_{k;t^{(i)}} \right)$ is restricted. Positivity on any separable state $\hat{\rho} \in \mathcal{L}^+(\mathcal{H})$ amounts then to restrict the n-dimensional vectors $\mathbf{y}_k \equiv \left(\mathbf{y}^{(1)}_k, \dots, \mathbf{y}^{(N)}_k \right) \in \mathbb{R}^n$ on the compact $K \subset \mathbb{R}^n$ where $K = K^{(1)} \times \dots \times K^{(N)}$, with $n = \sum_{i=1}^{N} t^{(i)}$. As described for the qubit case, if equation (2.50) is met, that is for any separable state $\hat{\rho} \in \mathcal{L}^+(\mathcal{H})$, it can be rewritten in an integral form [13]

$$\left\langle \hat{\Lambda}_{\mu_1 \dots \mu_N} \right\rangle_{\hat{\rho}} = \int_K (\mathbf{x}^{(1)})^{\alpha^{(1)}} \dots (\mathbf{x}^{(N)})^{\alpha^{(N)}} d\mu(\mathbf{x}) \tag{2.52}$$

$$= \int_K x^{(1)}_{\mu_1} \dots x^{(N)}_{\mu_N} d\mu(\mathbf{x}) \tag{2.53}$$

with $d\mu(\mathbf{x})$ a probability measure supported on K defined by

$$d\mu(\mathbf{x}) = \sum_k w_k \delta(\mathbf{x} - \mathbf{y}_k) d\mathbf{x}, \tag{2.54}$$

and $x^{(i)}_0 = 1$, $\mathbf{x} \in \mathbb{R}^n$,

$$\mathbf{x} = (x_1, \dots, x_n) = \left(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N)} \right) \tag{2.55}$$

$(x^{(i)}_j) \in \mathbb{R}^{t^{(i)}}$ for $j \in \{1, \dots, t^{(i)}\}$ the vector of n variables associated to the product space $\mathcal{H}$. Conversely, one can show that equation (2.53) can be reduced to a finite sum as in equation (2.50) such that the measure $d\mu(\mathbf{x})$ can be written as an atomic measure [39].

Let us denote $\left\langle \hat{\Lambda}_{\mu_1 \dots \mu_N} \right\rangle$ as y_α where the tuple α is such that

$$\alpha = \left((\alpha^{(1)}_1, \dots, \alpha^{(1)}_{t^{(1)}}), \dots, (\alpha^{(N)}_1, \dots, \alpha^{(N)}_{t^{(N)}}) \right)$$

where each $\alpha^{(i)} \in \mathbb{N}_1^{t^{(i)}}$,*i.e.*, the N-tuple α is so that each tuple $\alpha^{(j)}$ contains at most one 1. Equation (2.53) can be rewritten as

$$y_\alpha = \int_K \mathbf{x}^\alpha d\mu(\mathbf{x}) \tag{2.56}$$

where α can only takes the values as described above, that is $\alpha \in \mathcal{A} \subset \mathbb{N}_N^n$. A state is then separable if its real coordinates can be written such that (2.56) is met. The problem of determining if $\hat{\rho} \in \mathcal{L}(\mathcal{H})$ is separable is then equivalent to an $\mathcal{A}K$-tms problem, that is, to determine if there exists a measure $d\mu(\mathbf{x})$ whose moments correspond to the real coordinates of $\hat{\rho}$, *i.e.*, if (2.52) holds for all $y_\alpha, \alpha \in \mathcal{A}$, for $\mathcal{A} \subset \mathbb{N}_N^n$. If there exists a representive measure, then the state can be written as (2.49), and is then separable.

Symmetric case

Consider now that every qudit have the same dimension d, and $t = d^2 - 1$. As presented in the last chapter, symmetric states are invariant under any permutation of the qudit states. For $\hat{\rho} \in \mathcal{P}(\mathcal{H})$, let

$$X_{\mu_1 \dots \mu_N} \equiv \mathrm{Tr}\{\hat{\rho} \hat{P}_S^\dagger \hat{\Lambda}_{\mu_1 \dots \mu_N} \hat{P}_S\}, \tag{2.57}$$

where $\hat{P}_S$ is the projector operator made of the Dicke states

$$\left| D_N^{(k)} \right\rangle = \mathcal{C} \sum_\pi \left| \underbrace{0 \dots 0}_{\text{k}} \underbrace{1 \dots 1}_{\text{N-k}} \right\rangle,$$

with $N - k$ excitations, and where $\mathcal{C}$ is a normalisation constant. $\hat{\rho}$ can then be expanded as [13]

$$\hat{\rho} = \mathcal{N} X_{\mu_1 \dots \mu_N} \hat{P}_S^\dagger \hat{\Lambda}_{\mu_1 \dots \mu_N} \hat{P}_S, \tag{2.58}$$

where $\mathcal{N}$ is the normalization constant such that $X_{0 \dots 0} = \mathrm{Tr}\, \hat{\rho} = 1$.

Since the state is symmetric, any permutation of the qubit states *i.e.*, of the indices $\mu_i \in \{0, \dots, t\}$, $i = 1, \dots, N$ in $X_{\mu_1 \dots \mu_N}$, leaves the tensor X in the symmetric subspace unchanged :

$$X_{\mu_1 \mu_2 \dots \mu_N} = X_{\mu_2 \mu_1 \dots \mu_N}, \tag{2.59}$$

which means that the local expectation values $X_{k;\mu_i}^{(i)}$ of each qubits are indistinguishable between them, *i.e.*, $X_{k;\mu_j}^{(i)} = X_{k;\mu_j}^{(i')}$ $\forall i, i', \mu_j$. It follow that the moment sequences $y^{(i)}$ of each qudit are indistinguishable between one another, that is $y_{k;\mu_j}^{(i)} = y_{k;\mu_j}^{(i')}$, $\forall i, i', \mu_j$, thus the different sets of variables associated to the subspaces $\mathcal{H}^{(i)}$ are indistinguishable between them. It follows that only one of the N different sets of variables needs to be considered, e.g., $x_j^{(1)}$ for $j = 1, \dots, t$. One has $x_j^{(i)} = x_j^{(i')}$, for all i, i', j In other words, the α in equation (2.56) is then $\in \mathbb{N}_N^n$. Indeed, since $x_1^{(1)} x_1^{(2)} = (x_1^{(1)})^2$ for the symmetric case. In summary, the total number of variables is divided by N compared to the general case of N-qudit of same dimension d, *i.e.*, $n = t$.

For a separable symmetric state, one has then, in terms of average values of the basis operators,

$$X_{\mu_1 \dots \mu_N} = \sum_k w_k \left(X_{k;\mu_1} \dots X_{k;\mu_N} \right) = \sum_k w_k \left(\prod_{i=1}^N X_{k;\mu_i} \right), \tag{2.60}$$

where $X_{k;0} = 1\ \forall k, j$, which can then be rewritten as

$$X_{\mu_1 \dots \mu_N} = \sum_k w_k \left(y_{k;\mu_1} \dots y_{k;\mu_N} \right) = \sum_k w_k \left(\prod_{i=1}^{N} y_{k;\mu_i} \right). \tag{2.61}$$

For a symmetric pure separable state, one has the simple form

$$X_{\mu_1 \dots \mu_N} = X_{\mu_1} \dots X_{\mu_N} = y_{\mu_1} \dots y_{\mu_N} \tag{2.62}$$

Equation (2.61) can be rewritten as [13] [27]

$$\begin{aligned} X_{\mu_1 \dots \mu_N} &= \int_K x_{\mu_1} \dots x_{\mu_N} d\mu(\mathbf{x}) \\ &= \int_K \mathbf{x}^{\alpha} d\mu(\mathbf{x}) \end{aligned} \tag{2.63}$$

with $x_0^{(i)} = 1$, $\mathbf{x} \in \mathbb{R}^t$, $\mathbf{x} = (x_1, \dots, x_t)$ the vector of t variables, and with $d\mu(\mathbf{x})$ a probability measure supported on $K = K^{(1)}$ defined by

$$d\mu(\mathbf{x}) = \sum_k w_k \delta(\mathbf{x} - \mathbf{y}_k) d\mathbf{x}. \tag{2.64}$$

Let us denote $X_{\mu_1 \dots \mu_N}$ as y_α where the tuple α is such that

$$\alpha = (\alpha_1, \dots, \alpha_t)$$

for $\alpha \in \mathbb{N}_N^3$. Equation (2.33) can be rewritten as

$$y_\alpha = \int_K \mathbf{x}^{\alpha} d\mu(\mathbf{x}) \tag{2.65}$$

where $\alpha \in \mathbb{N}_N^t$. Compared to the general case, $\alpha \in \mathcal{A}$ in (2.52) is replaced by $\alpha \in \mathbb{N}_N^t$. In other words, when $\hat{\rho}$ is symmetric, the separability problem is reduced to a K-tms problem, that is to determine if there exists a measure $d\mu(\mathbf{x})$ supported on K such that each moment of the moment sequence $y = (y_\alpha)_{\alpha \in \mathbb{N}_N^t}$ satisfies equation (2.65).

Exemple 2.4. Consider a two-qubit system as in example (2.3), for symmetric 2-qubit states. Since the two qubits are indistinguishable, $b_{k;1}^{(1)} = b_{k;1}^{(2)}$, $b_{k;2}^{(1)} = b_{k;2}^{(2)}$, $b_{k;3}^{(1)} = b_{k;3}^{(2)}$, thus $y_{k;1}^{(1)} = y_{k;1}^{(2)}$, $y_{k;2}^{(1)} = y_{k;2}^{(2)}$, $y_{k;3}^{(1)} = y_{k;3}^{(2)}$, and then $x_1^{(1)} = x_1^{(2)} = x_1$, $x_2^{(1)} = x_2^{(2)} = x_2$, $x_3^{(1)} = x_3^{(2)} = x_3$. If $\hat{\rho}$ is symmetric, then $\hat{\rho}$ is separable if there exists a representing

measure $d\mu(\mathbf{x})$ supported on $K = K^{(1)}$ such that each of the following equations holds :

$$X_{10} = X_{10} = y_{100} = \int_K x_1 d\mu(\mathbf{x}) = \int_K x_1^{(1)} d\mu(\mathbf{x}) = b_1^{(1)} = b_1^{(2)} = \int_K x_1^{(2)} d\mu(\mathbf{x})$$
$$X_{20} = X_{02} = y_{010} = \int_K x_2 d\mu(\mathbf{x}) = \int_K x_2^{(1)} d\mu(\mathbf{x}) = b_2^{(1)} = b_2^{(2)} = \int_K x_2^{(2)} d\mu(\mathbf{x})$$
$$X_{30} = X_{03} = y_{001} = \int_K x_3 d\mu(\mathbf{x}) = \int_K x_3^{(1)} d\mu(\mathbf{x}) = b_3^{(1)} = b_3^{(2)} = \int_K x_3^{(2)} d\mu(\mathbf{x})$$
$$X_{12} = X_{21} = y_{110} = \int_K x_1 x_2 d\mu(\mathbf{x}) = \int_K x_1^{(1)} x_2^{(2)} d\mu(\mathbf{x}) = b_1^{(1)} b_2^{(2)} = b_2^{(1)} b_1^{(2)} = \int_K x_2^{(1)} x_1^{(2)} d\mu(\mathbf{x})$$
$$X_{13} = X_{31} = y_{101} = \int_K x_1 x_3 d\mu(\mathbf{x}) = \int_K x_1^{(1)} x_3^{(2)} d\mu(\mathbf{x}) = b_1^{(1)} b_3^{(2)} = b_3^{(1)} b_1^{(2)} = \int_K x_3^{(1)} x_1^{(2)} d\mu(\mathbf{x})$$
$$X_{32} = X_{23} = y_{011} = \int_K x_1 x_3 d\mu(\mathbf{x}) = \int_K x_1^{(1)} x_3^{(2)} d\mu(\mathbf{x}) = b_1^{(1)} b_3^{(2)} = b_3^{(1)} b_1^{(2)} = \int_K x_3^{(1)} x_1^{(2)} d\mu(\mathbf{x})$$
$$X_{11} = y_{200} = \int_K x_1^2 d\mu(\mathbf{x}) = \int_K x_1^{(1)} x_1^{(2)} d\mu(\mathbf{x}) = b_1^{(1)} b_1^{(2)}$$
$$X_{22} = y_{020} = \int_K x_2^2 d\mu(\mathbf{x}) = \int_K x_2^{(1)} x_2^{(2)} d\mu(\mathbf{x}) = b_2^{(1)} b_2^{(2)}$$
$$X_{33} = y_{002} = \int_K x_3^2 d\mu(\mathbf{x}) = \int_K x_3^{(1)} x_3^{(2)} d\mu(\mathbf{x}) = b_3^{(1)} b_3^{(2)}.$$

If such a representing measure exists, then $\hat{\rho}$ is separable.

2.3 The $\mathcal{A}$K-truncated moment problem

2.3.1 Moment matrices and localizing matrices

Moment matrix

For a given integer $k \geq 0$, the *k-th order moment matrix* $M_k(y)$ is the real matrix $\in M_{t(k)}(\mathbb{R})$, whose rows and columns are indexed by the tuples $\alpha, \beta \in \mathbb{N}^n_k$, *i.e.*, $|\alpha|, |\beta| \leq k$, and with elements $M_k(y)_{\alpha,\beta}$ defined as

$$M_k(y)_{\alpha,\beta} = \int_{\mathbb{R}^n} \mathbf{x}^\alpha \mathbf{x}^\beta d\mu(\mathbf{x}) = \int_{\mathbb{R}^n} \mathbf{x}^{\alpha+\beta} d\mu(\mathbf{x}) = y_{\alpha+\beta}, \tag{2.66}$$

One has $M_k(y)_{\alpha,\beta} = M_k(y)_{\beta,\alpha}$, and $M_k(y)$ is thus a symmetric matrix.

From a given tms y of order d, moment matrix can be build up to a maximal order. Indeed, the k-th moment matrix $M_k(y)$ requires the moments y_α up to order

$$\max_{\alpha,\beta \in \mathbb{N}^n_k} \{\deg\left(\mathbf{x}^\alpha \mathbf{x}^\beta\right)\} = 2k,$$

thus leading to an upper bound for k:

$$k \leq d/2. \tag{2.67}$$

For a given k-order moment matrix $M_k(y)$, any k'-order moment matrices $M_{k'}(y)$, for all $k' \leq k$, is a submatrix of $M_k(y)$:

$$M_k(y) = \begin{pmatrix} M_{k'}(y) & B \\ B^T & C \end{pmatrix}. \tag{2.68}$$

Exemple 2.5. Let $n = 3$ and $d = 4$.

- $\mathbb{R}[\mathbf{x}] = \mathbb{R}[x_1, x_2, x_3]$,
- There are $\binom{3+4}{3} = \frac{(4+1)(4+2)(4+3)}{3!} = 35$ distinct triplet $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{N}^3_4$.
- $\mathbf{x}^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3} \in \mathbb{T}^3_4 = \{\mathbf{x}^\alpha | \alpha \in \mathbb{N}^3_4\}$. There are 35 distinct monomials in $\mathbb{T}^3_4$ (see Appendix A).
- The monomial basis is $\mathcal{B}^3_4$ is made of the 35 monomials of $\mathbb{T}^3_4$ sorted by degree, and within each degree in a lexicographic order (see Appendix A).

Let $d\mu(\mathbf{x})$ be a probability measure supported on $\mathbb{R}^3$. The order of the moment matrices can be of

$$k \leq 4/2 = 2,$$

that is $k = 0$, 1, or 2.

k=0

The moment matrix of order 0 is made of only one element:

$$y_{000} = \int x_1^0 x_2^0 x_3^0 d\mu(\mathbf{x}) = \int 1 d\mu(\mathbf{x}) = 1,$$

and then

$$M_0(y) = (1).$$

k=1

To construct the 1-st order moment matrix ($k = 1$), the moments up to order $d = 2k = 2$ are needed. The truncated moment sequence of degree 2 is the vector $y \equiv (y_\alpha)_{\alpha \in \mathbb{N}^3_2}$ made of the moments $y_\alpha, \forall \alpha \in \mathbb{N}^3_2$ of μ, that is $\forall \alpha$ such that $|\alpha| = \sum_i^3 \alpha_i \leq 2$. The number of elements in y is $\binom{n+d}{d} = \binom{3+2}{2} = 10$:

$$y = (y_{000}, y_{100}, y_{010}, y_{001}, y_{200}, y_{110}, , y_{101}, y_{020}, y_{011}, y_{002})$$

where

$$
\begin{aligned}
y_{100} &= \int_{\mathbb{R}^n} x_1^1 x_2^0 x_3^0 d\mu(\mathbf{x}) = \int_{\mathbb{R}^n} x_1 d\mu(\mathbf{x}),\\
y_{010} &= \int_{\mathbb{R}^n} x_1^0 x_2^1 x_3^0 d\mu(\mathbf{x}) = \int_{\mathbb{R}^n} x_2 d\mu(\mathbf{x}),\\
y_{001} &= \int_{\mathbb{R}^n} x_1^0 x_2^0 x_3^1 d\mu(\mathbf{x}) = \int_{\mathbb{R}^n} x_3 d\mu(\mathbf{x}),\\
y_{200} &= \int_{\mathbb{R}^n} x_1^2 x_2^0 x_3^0 d\mu(\mathbf{x}) = \int_{\mathbb{R}^n} x_1^2 d\mu(\mathbf{x}),\\
y_{110} &= \int_{\mathbb{R}^n} x_1^1 x_2^1 x_3^0 d\mu(\mathbf{x}) = \int_{\mathbb{R}^n} x_1^1 x_2^1 d\mu(\mathbf{x}),\\
y_{101} &= \int_{\mathbb{R}^n} x_1^1 x_2^0 x_3^1 d\mu(\mathbf{x}) = \int_{\mathbb{R}^n} x_1^1 x_3^1 d\mu(\mathbf{x}),\\
y_{020} &= \int_{\mathbb{R}^n} x_1^0 x_2^2 x_3^0 d\mu(\mathbf{x}) = \int_{\mathbb{R}^n} x_2^2 d\mu(\mathbf{x}),\\
y_{011} &= \int_{\mathbb{R}^n} x_1^0 x_2^1 x_3^1 d\mu(\mathbf{x}) = \int_{\mathbb{R}^n} x_2^1 x_3^1 d\mu(\mathbf{x}),\\
y_{002} &= \int_{\mathbb{R}^n} x_1^0 x_2^0 x_3^2 d\mu(\mathbf{x}) = \int_{\mathbb{R}^n} x_3^2 d\mu(\mathbf{x}).
\end{aligned}
$$

The 1-st order moment matrix is

$$
M_1(y) = \begin{pmatrix}
1 & y_{100} & y_{010} & y_{001}\\
y_{100} & y_{200} & y_{110} & y_{101}\\
y_{010} & y_{110} & y_{020} & y_{011}\\
y_{001} & y_{101} & y_{011} & y_{002}
\end{pmatrix}.
$$

One can observe that the 0-order moment matrix $M_0(y)$ of y is indeed a sub-matrix of the 1-st order moment matrix $M_1(y)$ of y :

$$
M_1(y) = \begin{pmatrix} M_0(y) & B\\ B^T & C \end{pmatrix} \tag{2.69}
$$

with

$$
B = \begin{pmatrix} y_{100} & y_{010} & y_{001}\\ y_{200} & y_{110} & y_{101} \end{pmatrix}, \tag{2.70}
$$

$$
B^T = \begin{pmatrix} y_{100} & y_{200}\\ y_{010} & y_{110}\\ y_{001} & y_{101} \end{pmatrix}, \tag{2.71}
$$

and

$$
C = \begin{pmatrix} y_{020} & y_{011}\\ y_{011} & y_{002} \end{pmatrix}. \tag{2.72}
$$

k=2

Similarly, to construct the 2-nd order moment matrix ($k = 2$), the moments up to order $d = 2k = 4$ are needed. The truncated moment sequence of degree 4 is the vector $y \equiv$

$(y_\alpha)_{\alpha\in\mathbb{N}^3_4}$ made of the moments $y_\alpha, \forall\alpha\in\mathbb{N}^3_4$ of $d\mu(\mathbf{x})$, that is $\forall\alpha$ such that $|\alpha|=\sum_i^3\alpha_i\leq 4$. The number of elements in y is $\binom{3+4}{4}=35$, and the 2-nd order moment matrix is (see Appendix A):

$$M_2(y)=\begin{pmatrix} 1 & y_{100} & y_{010} & y_{001} & y_{200} & y_{110} & y_{101} & y_{020} & y_{011} & y_{002}\\ y_{100} & y_{200} & y_{110} & y_{101} & y_{300} & y_{210} & y_{201} & y_{120} & y_{111} & y_{102}\\ y_{010} & y_{110} & y_{020} & y_{011} & y_{210} & y_{120} & y_{111} & y_{030} & y_{021} & y_{012}\\ y_{001} & y_{101} & y_{011} & y_{002} & y_{201} & y_{111} & y_{102} & y_{021} & y_{012} & y_{003}\\ y_{200} & y_{300} & y_{210} & y_{201} & y_{400} & y_{310} & y_{301} & y_{220} & y_{211} & y_{202}\\ y_{110} & y_{210} & y_{120} & y_{111} & y_{310} & y_{220} & y_{211} & y_{130} & y_{121} & y_{112}\\ y_{101} & y_{201} & y_{111} & y_{102} & y_{301} & y_{211} & y_{202} & y_{121} & y_{112} & y_{103}\\ y_{020} & y_{120} & y_{030} & y_{021} & y_{220} & y_{130} & y_{121} & y_{040} & y_{031} & y_{022}\\ y_{011} & y_{111} & y_{021} & y_{012} & y_{211} & y_{121} & y_{112} & y_{031} & y_{022} & y_{013}\\ y_{002} & y_{102} & y_{012} & y_{003} & y_{202} & y_{112} & y_{103} & y_{022} & y_{013} & y_{004}\end{pmatrix} \tag{2.73}$$

One can observe that $M_1(y)$ of y is a sub-matrix of the 2-nd order moment matrix $M_2(y)$ of y, as well as $M_0(y)$, since $M_0(y)$ is a sub-matrix of $M_1(y)$. Indeed, one can write

$$M_2(y)=\begin{pmatrix} M_1(y) & B\\ B^T & C\end{pmatrix} \tag{2.74}$$

with

$$B=\begin{pmatrix} y_{200} & y_{110} & y_{101} & y_{020} & y_{011} & y_{002}\\ y_{300} & y_{210} & y_{201} & y_{120} & y_{111} & y_{102}\\ y_{210} & y_{120} & y_{111} & y_{030} & y_{021} & y_{012}\\ y_{201} & y_{111} & y_{102} & y_{021} & y_{012} & y_{003}\\ y_{400} & y_{310} & y_{301} & y_{220} & y_{211} & y_{202}\end{pmatrix}, \tag{2.75}$$

$$B^T=\begin{pmatrix} y_{200} & y_{300} & y_{210} & y_{201} & y_{400}\\ y_{110} & y_{210} & y_{120} & y_{111} & y_{310}\\ y_{101} & y_{201} & y_{111} & y_{102} & y_{301}\\ y_{020} & y_{120} & y_{030} & y_{021} & y_{220}\\ y_{011} & y_{111} & y_{021} & y_{012} & y_{211}\\ y_{002} & y_{102} & y_{012} & y_{003} & y_{202}\end{pmatrix}, \tag{2.76}$$

and

$$C=\begin{pmatrix} y_{202} & y_{121} & y_{112} & y_{103}\\ y_{121} & y_{040} & y_{031} & y_{022}\\ y_{112} & y_{031} & y_{022} & y_{013}\\ y_{103} & y_{022} & y_{013} & y_{004}\end{pmatrix}. \tag{2.77}$$

2.3.2 Shifted moment sequence and localizing matrix

For any polynomial

$$g(\mathbf{x})=\sum_{\gamma\in\mathbb{N}^n_{t_g}} g_\gamma\mathbf{x}^\gamma$$

of degree $t_g = \deg(g(\mathbf{x})) \geq 1$, and a truncated moment sequence y of degree d, the truncated moment sequence of degree $d - \deg(g(\mathbf{x}))$ defined as

$$(g * y)_\alpha = \sum_{\gamma \in \mathbb{N}^n_{t_g}} g_\gamma y_{\alpha+\gamma}, \tag{2.78}$$

where $|\alpha| \leq t_g$, is called a *shifted truncated moment sequence*.

Let $g(\mathbf{x})$ be a polynomial of degree $\deg(g(\mathbf{x})) \geq 1$ and

$$d_g = \lceil \deg(g(\mathbf{x}))/2 \rceil, \tag{2.79}$$

where $\lceil x \rceil$ denotes the smallest integer equal or larger than x. For any integer $(k - d_g)$, the *kth-order localizing matrix* $M_{k-d_g}(g*y)$ of $g(\mathbf{x})$ is the $(k-d_g)$-th order moment matrix $M_{k-d_g}(g * y)$ whose elements are defined as

$$\begin{aligned} M_{k-d_g}(g * y)_{\alpha,\beta} &= (g * y)_{\alpha+\beta} \\ &= \sum_{\gamma \in \mathbb{N}^n_{t_g}} g_\gamma y_{\alpha+\beta+\gamma}, \end{aligned} \tag{2.80}$$

for $(k - d_g) \geq 0$, $\alpha, \beta \in \mathbb{N}^n_{k-d_g}$, *i.e.*, $|\alpha|, |\beta| \leq (k - d_g)$.

From a given tms y and a polynomial $g(\mathbf{x})$, localizing matrices do not exist for any integer k. Indeed, the $(k - d_g)$-th moment matrix $M_{k-d_g}(y)$ requires the moments y_α up to order:

$$d \geq \deg\left(g(\mathbf{x})\right) + 2(k - d_g).$$

This leads to an upper bound for $k - d_g$:

$$\begin{aligned} & \deg(g(\mathbf{x})) + 2k - 2d_g \leq d \\ \Leftrightarrow \quad & 2k - 2d_g \leq d - \deg(g(\mathbf{x})) \\ \Leftrightarrow \quad & k - d_g \leq \frac{(d - \deg(g(\mathbf{x})))}{2}. \end{aligned}$$

Since $(k - d_g) \geq 0$,

$$d_g \leq k \leq \frac{(d - \deg(g(\mathbf{x})))}{2} + d_g, \tag{2.81}$$

and with

$$\lfloor (d - \deg(g(\mathbf{x})))/2 \rfloor = \lfloor d/2 \rfloor - d_g,$$

where $\lfloor x \rfloor$ is the largest integer smaller than x, localizing matrices for a given tms y of order d and a polynomial $g(\mathbf{x})$ exist for any integer k such that

$$d_g \leq k \leq d/2. \tag{2.82}$$

The definition of d_g has been chosen in such a way that the upper bound $k \leq d/2$ is the same as that for the kth-order moment matrix. Any k'-th order localizing matrix $M_{k'-d_g}(g * y)$ for $k' \leq k$ is as sub-matrix of the k-th order localizing matrix $M_{k'-d_g}(g * y)$.

Exemple 2.6. Let $n = 3$ and $d = 4$, $d\mu(\mathbf{x})$ a measure supported on $\mathbb{R}^3, \alpha \in \mathbb{N}^3_4, \mathbf{x}^\alpha \in \mathbb{T}^3_4$. Consider the polynomial

$$g(\mathbf{x}) = 1 - x_1^2 - x_2^2 - x_3^2 \tag{2.83}$$

of degree $\deg(g(\mathbf{x})) = 2$, and $d_g = \lceil 2/2 \rceil = 1$. The values of k for which the moment matrix $M_{k-d_g}(g * y)$ exists are

$$1 \leq k \leq 4/2,$$

that is, $k = 1$ or 2.

k=1

The 1st-order ($k = 1$) localizing matrix $M_{1-1}(g * y) = M_0(g * y)$ is

$$M_0(g * y) = 1 - y_{200} - y_{020} - y_{002}$$

k=2

The 2nd-order ($k = 2$) localizing matrix $M_{2-1}(g * y) = M_1(g * y)$ is

$$M_1(g * y) = \begin{pmatrix} 1 - y_{200} - y_{020} - y_{002} & y_{100} - y_{300} - y_{020} - y_{002} & y_{010} - y_{200} - y_{030} - y_{002} & y_{001} - y_{200} - y_{020} - y_{003} \\ y_{100} - y_{300} - y_{020} - y_{002} & y_{200} - y_{400} - y_{020} - y_{002} & y_{110} - y_{300} - y_{030} - y_{002} & y_{101} - y_{300} - y_{020} - y_{003} \\ y_{010} - y_{200} - y_{030} - y_{002} & y_{110} - y_{300} - y_{030} - y_{002} & y_{020} - y_{200} - y_{040} - y_{002} & y_{011} - y_{200} - y_{030} - y_{003} \\ y_{001} - y_{200} - y_{020} - y_{003} & y_{101} - y_{300} - y_{020} - y_{003} & y_{011} - y_{200} - y_{030} - y_{003} & y_{002} - y_{200} - y_{020} - y_{004} \end{pmatrix}. \tag{2.84}$$

2.3.3 The truncated moment problem

For a given probability measure $d\mu(\mathbf{x})$ on $\mathbb{R}^n$, one can find its moment sequence $y = (y)_{\alpha \in \mathbb{N}^n}$ made of the moments given by equation (2.25), and obtain informations on the probability measure $d\mu(\mathbf{x})$. The moment problem is the inverse problem: given a (truncated) sequence of moments, one can try to find a measure from the knowledge of its moments. Formally, the (truncated) moment problem is to find *conditions* under which there exists a measure $d\mu(\mathbf{x})$ such that each y_α of a (truncated) moment sequence y can be *represented* as an integral of the form

$$y_\alpha = \int_{\mathbb{R}^n} \mathbf{x}^\alpha d\mu(\mathbf{x}). \tag{2.85}$$

If such a measure exists, $d\mu(\mathbf{x})$ is called a *representing measure*.

The K-truncated moment problem

If the support of the unknown measure $d\mu(\mathbf{x})$ is a subset K of $\mathbb{R}^n$ defined by multivariate polynomials $g_i(\mathbf{x})$ in the variables $(x_1, \ldots, x_n)$:

$$K \equiv \{\mathbf{x} \in \mathbb{R}^n | g_1(\mathbf{x}) \geq 0, \ldots, g_m(\mathbf{x}) \geq 0\}, \tag{2.86}$$

that is, a semialgebraic set of $\mathbb{R}^n$, the truncated moment problem is called a *K-truncated moment problem*, or a *K-tms problem*. In other words, in a K-tms problem, $d\mu(\mathbf{x})$ has to satisfy

$$y_\alpha = \int_K \mathbf{x}^\alpha d\mu(\mathbf{x}) \tag{2.87}$$

for all $\alpha \in \mathbb{N}^n_d$ and for $\mathbf{x}^\alpha \in \mathbb{T}^n_d$, with $d\mu(\mathbf{x})$ supported on $K \subset \mathbb{R}^n$

A generalisation of the K-tms problem is the *$\mathcal{A}$K-tms problem* where the moments y_α are known only for a finite subset $\mathcal{A} \subset \mathbb{N}^n_d$ of degree $|\alpha| \leq d$. In other words, (2.87) has to be fulfilled only for $\alpha \in \mathcal{A}$ rather than for all $|\alpha| \leq d$.

The rest of this section presents necessary, and necessary *and* sufficient conditions for a truncated moment sequence to admit a representing measure.

Necessary conditions

Theorem 2.1. *An order-d tms $y \equiv (y_\alpha)_{\alpha \in \mathbb{N}^n_d}$ admits a representing measure such that (2.85) holds for all y_α if the k-th order moment matrix $M_k(y)$ for $k \leq d/2$ is positive semidefinite.*

Proof. If

$$y_\alpha = \int_K \mathbf{x}^\alpha d\mu(\mathbf{x})$$

holds for all y_α, then for any polynomial $p(\mathbf{x})$ in $\mathbb{R}[\mathbf{x}]$

$$p(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}^n_k} p_\alpha \mathbf{x}^\alpha,$$

of degree k or less, one has

$$\begin{aligned}
\mathbf{p}^T M_k(y)\mathbf{p} &= \sum_{\alpha,\beta \in \mathbb{N}^n_k} p_\alpha y_{\alpha+\beta} p_\beta \\
&= \sum_{\alpha,\beta \in \mathbb{N}^n_k} p_\alpha p_\beta \int_K \mathbf{x}^{\alpha+\beta} d\mu(\mathbf{x}) \\
&= \sum_{\alpha,\beta \in \mathbb{N}^n_k} \int_K (p_\alpha \mathbf{x}^\alpha)(p_\beta \mathbf{x}^\beta) d\mu(\mathbf{x}) \\
&= \int_K p(\mathbf{x})^2 d\mu(\mathbf{x}) \\
&\geq 0,
\end{aligned}$$

where $\mathbf{p}$ is the column vector made of the element of the sequence p. $M_k(y)$ is thus a positive semidefinite matrix. □

Since moment matrices $M_{k'}(y)$ of order k' are sub matrices of $M_k(y)$ for $k' \geq k$, one can consider the largest possible value for k, which is the upper bound $d/2$ in the equation (2.67), to get the strongest necessary conditions:

Theorem 2.2. *An order-d tms $y \equiv (y_\alpha)_{\alpha \in \mathbb{N}^n_d}$ admits a representing measure such that (2.85) holds for all y_α if the moment matrix $M_{\lfloor d/2 \rfloor}(y)$ is positive-semidefinite.*

A similar necessary condition exists for localizing matrices.

Theorem 2.3. *If an order-d tms $y = (y_\alpha)_{\alpha \in \mathbb{N}^n_d}$ admits a representing measure such that (2.85) holds for all y_α, then any kth order localizing matrix is necessarily positive semidefinite.*

Proof. If

$$y_\alpha = \int_K \mathbf{x}^\alpha d\mu(\mathbf{x})$$

holds, then for then for any polynomial $g(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$

$$g(\mathbf{x}) = \sum_{\gamma \in \mathbb{N}^n_{\deg(g(\mathbf{x}))}} g_\gamma \mathbf{x}^\gamma,$$

of degree k or less where

$$K = \{\mathbf{x} \in \mathbb{R}^n | g_1(\mathbf{x}) \geq 0, \ldots, g_m(\mathbf{x}) \geq 0\}, \tag{2.88}$$

and for any polynomial

$$p(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}^n_{k-d_g}} p_\alpha \mathbf{x}^\alpha,$$

of degree $k - d_g$ or less, one has

$$\begin{aligned}
\mathbf{p}^T M_{k-d_g}(g * y)\mathbf{p} &= \sum_{\alpha,\beta \in \mathbb{N}^n_{k-d_g}} p_\alpha \sum_{\gamma \in \mathbb{N}^n_{deg(g(\mathbf{x}))}} g_\gamma y_{\alpha+\beta+\gamma} p_\beta \\
&= \sum_{\alpha,\beta \in \mathbb{N}^n_{k-d_g}} p_\alpha p_\beta \sum_{\gamma \in \mathbb{N}^n_{deg(g(\mathbf{x}))}} g_\gamma \int_K x^{\alpha+\beta+\gamma} d\mu(\mathbf{x}) \\
&= \sum_{\alpha,\beta \in \mathbb{N}^n_{k-d_g}} \sum_{\gamma \in \mathbb{N}^n_{deg(g(\mathbf{x}))}} \int_K (g_\gamma \mathbf{x}^\gamma)(p_\alpha \mathbf{x}^\alpha)(p_\beta \mathbf{x}^\beta) d\mu(\mathbf{x}) \\
&= \int_K g(\mathbf{x}) p(\mathbf{x})^2 d\mu(\mathbf{x}) \\
&\geq 0.
\end{aligned}$$

since $g(\mathbf{x})$ is positive on K by definition, where $\mathbf{p}$ is the column vector made of the elements of the sequence p. M_{k-d_g} is thus positive-semidefinite □

Since moment matrices $M_{k'}(g * y)$ of order k' are sub-matrices of $M_k(g * y)$ for $k' \geq k$, one can consider the largest possible value for k, which is the upper bound $d/2$ of equation (2.82), to get the strongest necessary conditions:

Theorem 2.4. *If an order-d tms $y \equiv (y_\alpha)_{\alpha \in \mathbb{N}^n_d}$ admits a representing measure such that (2.85) holds for all y_α, then all localizing matrix $M_{\lfloor d/2 \rfloor - d_{g_i}}(g_i * y)$ for each polynomial $g_i(\mathbf{x}) \in R[\mathbf{x}]$ as in (2.88),and with $d_{g_i} = \lceil deg(g_i(\mathbf{x})) \rceil$, is necessarily positive-semidefinite.*

Sufficient condition

The following sufficient condition was obtained in [16] for even order truncated moment sequences. The formulation below comes from [17] and presented again in [13, 14]. Proofs of the theorem can be found in [16, 40].

Theorem 2.5. *If an order $2k$ tms $z = (z_\beta)_{\beta \in \mathbb{N}^n_{2k}}$ is such that its kth order moment matrix and all kth order localizing matrices are positive, and if additionally*

$$\operatorname{rank}(M_k(z)) = \operatorname{rank}(M_{k-d_0}(z)), \tag{2.89}$$

that is, the moment matrix $M_k(z)$ is a flat extension of the moment matrix $M_{k-d_0}(z)$, with

$$d_0 = \max_{1\leq i\leq m}\{1, \lceil deg(g_i(\mathbf{x}))/2\rceil\}, \tag{2.90}$$

for $g_i(\mathbf{x})$, then the tms $(z_\beta)_{\beta\in\mathbb{N}^n_{2k}}$ admits a representing measure composed of

$$r = \text{rank}\,(M_k(z))$$

delta functions.

Necessary and sufficient condition

Since the latter condition is only sufficient, a representing measure does not necessarily satisfy the rank condition. One can however search for an *extension* of y that satisfies it. An extension of an order d tms $y \equiv (y_\alpha)_{\alpha\in\mathbb{N}^n_d}$ is defined as any order $2k$ tms $(z_\beta)_{\beta\in\mathbb{N}^n_{2k}}$ of degree $2k$ with $2k > d$ such that $z_\alpha = y_\alpha$ for all $|\alpha| \leq d$. An extension $(z_\beta)_{\beta\in\mathbb{N}^n_{2k}}$ is called *flat* if it satisfies $\text{rank}\,(M_k(z)) = \text{rank}\,(M_{k-d_0}(z))$ with d_0 as in (2.90). If $(z_\beta)_{\beta\in\mathbb{N}^n_{2k}}$ satisfies theorem (2.5) above, then it has a representing measure, and so does $(y_\alpha)_{\alpha\in\mathbb{N}^n_d}$ as a restriction of $(z_\beta)_{\beta\in\mathbb{N}^n_{2k}}$. One can then formulate a necessary and sufficient condition for the existence of a representing measure. The following necessary and sufficient condition was obtained in [16]. The formulation below comes from [17] and presented again in [14] [13]. A proof of the theorem can be found in [16].

Theorem 2.6. *A tms $y = (y_\alpha)_{\alpha\in\mathbb{N}^n_d}$ admits a representing measure supported on K if and only if there exists a flat extension $(z)_{\beta\in\mathbb{N}^n_{2k}}$ with $2k > d$ such that $M_k(z_\beta) \succeq 0$ and $M_{k-d_{g_i}}(g_i * z) \succeq 0$ with $d_{g_i} = \lceil \deg(g_i(\mathbf{x}))/2\rceil$ for $i = 1, \ldots, m$.*

The necessary and sufficient condition in Theorem (2.6) has been generalized for $\mathcal{A}K$-tms in [18]. A proof can be found in [14, 18].

Theorem 2.7. *An $\mathcal{A}-$tms $(y_\alpha)_{\alpha\in\mathcal{A}}$, $\mathcal{A} \subset \mathbb{N}^n_d$ admits a representing measure supported on K if and only if there exists a flat extension $z = (z_\beta)_{\beta\in\mathbb{N}^n_{2k}}$ with $2k > d$ such that $M_k(z) \succeq 0$ and $M_{k-d_{g_i}}(g_i * z) \succeq 0$, with $d_{g_i} = \lceil \deg(g_i(\mathbf{x}))/2\rceil$ for $i = 1, \ldots, m$.*

Exemple 2.7. Let $n = 3$ and $d = 4$, $d\mu(\mathbf{x})$ a measure supported on $\mathbb{R}^3, \alpha \in \mathbb{N}^3_4$, $\mathbf{x}^\alpha \in \mathbb{T}^3_4$. Consider the following K-truncated moment problem: Given the tms $y = (y_\alpha)_{\alpha\in\mathbb{N}^3_4}$, what are the conditions under which y admits a representing measure $d\mu(\mathbf{x})$ over $\mathbb{R}^3$ supported on the subset $K \subset \mathbb{R}^3$

$$K = \{\mathbf{x} \in \mathbb{R}^n | 1 - x_1^2 - x_2^2 - x_3^2 \geq 0\}, \tag{2.91}$$

where $g(\mathbf{x}) = \sum_{\gamma\in\mathbb{N}^n_2} g_\gamma \mathbf{x}^\gamma = 1 - x_1^2 - x_2^2 - x_3^2$ is a polynomial of degree 2, such that

$$y_\alpha = \int_K \mathbf{x}^\alpha d\mu(\mathbf{x})$$

holds for all y_α of y. Theorem (2.6) states that y admits a representing measure if there exists any extension $z = (z_\beta)_{\beta\in\mathbb{N}^n_{2k}}$ of degree $2k$ with

$$2k > d = 4 \Leftrightarrow k > 2,$$

such that $z_\alpha = y_\alpha$ for all $|\alpha| \leq d = 4$, and that satisfies the rank condition $\operatorname{rank} M_k(z) = \operatorname{rank} M_{k-d_0}(z)$ where

$$d_0 = \max_{1 \leq i \leq m} \{1, \lceil deg(g_i(\mathbf{x}))/2 \rceil\} = \max\{1, \lceil 1 \rceil\} = 1,$$

and such that

$$M_k(z) \succeq 0$$

and

$$M_{k-d_g}(g * z) \succeq 0,$$

with $d_g = \lceil \deg(g(\mathbf{x}))/2 \rceil \ = 1$.

k=3

Let us construct the conditions for the smallest k possible, that is $k = 3$. y admits a representing measure if and only if the following conditions hold:

- All the moments z_β up to order 4 of the extension $z = (z_\beta)_{\beta \in \mathbb{N}^3_{2k=6}}$ are the same as the moments y_α in the tms $y = (y_\alpha)_{\alpha \in \mathbb{N}^3_4}$.
- The ranks of the moment matrices $M_3(z)$ and $M_2(z)$ are the same, *i.e.*, $M_3(z)$ is a flat extension of $M_2(z)$.
- $M_3(z)$ is positive semi-definite.
- $M_2(g * z)$ is positive semi-definite.

k=4

Let us construct the conditions for the smallest k possible, that is $k = 4$. y admits a representing measure if and only if the following conditions hold:

- All the moments z_β up to order 4 of the extension $z = (z_\beta)_{\beta \in \mathbb{N}^3_8}$ are the same as the moments y_α in the tms $y = (y_\alpha)_{\alpha \in \mathbb{N}^3_8}$.
- The ranks of the moment matrices $M_4(z)$ and $M_3(z)$ are the same, *i.e.*, $M_4(z)$ is a flat extension of $M_3(z)$.
- $M_4(z)$ is positive semi-definite.
- $M_3(g * z)$ is positive semi-definite.

The conditions for $k > 4$ are analogous. One can keep increasing the value k until the 4 conditions above are satisfied for a given k. If the conditions are satisfied for some $k^* > 2$, then the tms y admits a representing measure composed of $\operatorname{rank} M_{k^*}(z)$ delta functions.

From theorem 2.7, one can derive a necessary and sufficient condition for the separability of $\hat{\rho}$.

Theorem 2.8. *A state $\hat{\rho}$ is separable if and only if its coordinates $X_{\mu_1 \ldots \mu_N}$ correspond to a tms $(y_\alpha)_{\alpha \in \mathcal{A}}$ such that there exists a flat extension $z = (z_\beta)_{\beta \in \mathbb{N}^n_{2k}}$ with $2k > d$ such that $M_k(z) \succeq 0$ and $M_{k-d_{g_i}}(g_i * z) \succeq 0$, $g_i(\mathbf{x}) \in K$ the polynomial inequality constraints derived from the positivity of $\hat{\rho}$, with $d_{g_i} = \lceil \deg(g_i(\mathbf{x}))/2 \rceil$ for $i = 1, \ldots, m$.*

As an example consider the 2-qubit case presented in example 2.3. For the K-tms problem, that is, when the state is symmetric, the conditions to be satisfied are presented in the example 2.7 above. For the $\mathcal{A}K$-tms for the general case, for $k = 3$, y admits a representing measure if and only if the following conditions hold:

- All the moments z_β up to order 4 of the extension $z = (z_\beta)_{\beta \mathbb{N}^3_{2k=6}}$ are the same as the moments y_α in the tms $y = (y_\alpha)_{\alpha \in \mathcal{A} \subset \mathbb{N}^3_4}$.
- The ranks of the moment matrices $M_3(z)$ and $M_2(z)$ are the same, *i.e.*, $M_3(z)$ is a flat extension of $M_2(z)$.
- $M_3(z)$ is positive semi-definite.
- $M_2(g_0^{(1)} * z)$ is positive semi-definite where $g_0^{(1)}(\mathbf{x}) = 1 - (x_1^{(1)})^2 - (x_2^{(1)})^2 - (x_3^{(1)})^2 = 1 - x_1^2 - x_2^2 - x_3^2$.
- $M_2(g_0^{(2)} * z)$ is positive semi-definite where $g_0^{(2)}(\mathbf{x}) = 1 - (x_1^{(2)})^2 - (x_2^{(2)})^2 - (x_3^{(2)})^2 = 1 - x_4^2 - x_5^2 - x_6^2$.

The conditions for higher values of k are analogous.

The mapping allows some flexibility on the problem and its constraints. For instance, for a 3 qubit system, one can determine if the two first qubits are separable with respect to the third qubit while ignoring any entanglement between the first two ones. This can be done by taking the first two qubits as a 4-level system. There are then two sets of variable, 15 for the two qubits, 3 for the second one, then $n = 18$ and $N = 2$. One can also impose a symmetry between the two first qubits only, by equating the first two sets of variables.

Chapter 3

Semidefinite optimization for the separability problem

The aim of this chapter is to describe how one can solve a moment problem, thus the separability problem, using semidefinite optimization. The first section presents the basic notions related to semidefinite optimization. First a description of optimization problems is presented following [41]. The concepts of global solutions, convexity, semidefinite optimization, and the dual theory of linear programming are then presented, followed by a description of how one can relax a polynomial optimization. An algorithm to solve a truncated moment problem is then presented. The second section of this chapter describes a semidefinite optimization algorithm to solve the separability problem as a truncated moment problem. Results of our implementation of the algorithm are then presented.

3.1 Semidefinite optimization for the truncated moment problem

3.1.1 Optimization problems

Throughout, $\mathbb{R}^n$ is considered as the set of n-tuples presented as column vectors.

Generalities

An *optimization problem* is the minimization or maximization of a function subjects to constraints in its variables. Mathematically, it is formulated as:

$$\begin{array}{ll} \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimize}} & f(\mathbf{x}) \\ \text{subject to} & h_i(\mathbf{x}) = 0, \quad i \in \mathcal{E}, \\ & g_j(\mathbf{x}) \geq 0, \quad j \in \mathcal{I}. \end{array} \tag{3.1}$$

where

- $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ is the real vector of *variables* (also called *unknowns* or *parameters*),

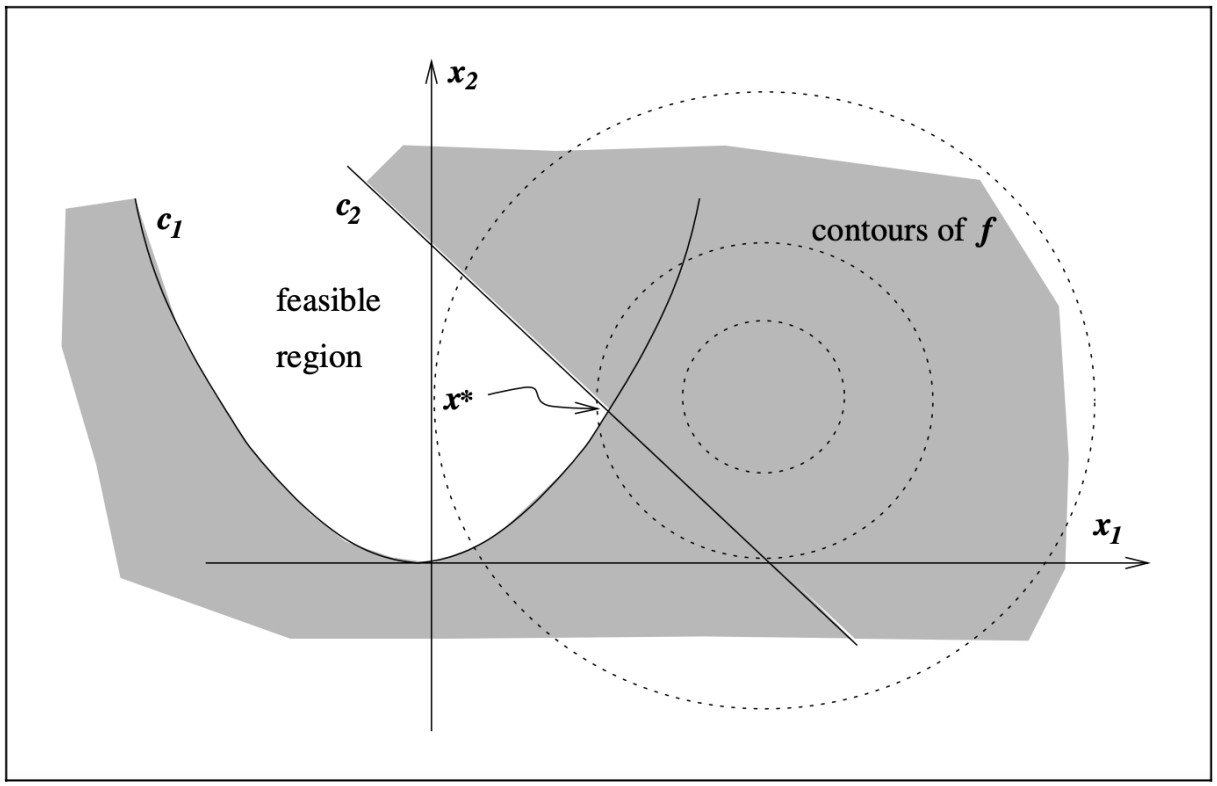

Figure 3.1: Geometrical representation of the problem (3.2) where the dotted lines represent the contours of the objective function f, c_1 and c_2 the constraints, and $\mathbf{x}^*$ is the solution of the problem. The shaded part represents the infeasible region. (Figure taken from [41])

- $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ is the scalar function we want to maximize or minimize, usually called the *objective function*,
- $\forall i, j,\ h_i(\mathbf{x}), g_j(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ are the scalar *constraints* functions of $\mathbf{x}$ that define equalities and inequalities respectively the unknown vector $\mathbf{x}$ must satisfy,
- $\mathcal{E}$ and $\mathcal{I}$ are finite sets of indices for equality and inequality constraints respectively.

Optimization problems can be classified depending on the nature of the objective function, their constraints, and the number of variables. An important distinction is made between problems that have constraints and those that do not. If $\mathcal{E} = \mathcal{I} = \emptyset$, the problem is called an *unconstrained optimization* problem. If at least $\mathcal{E}$ or $\mathcal{I}$ is $\neq \emptyset$, the problem is called a *constrained optimization* problem. These constraints define a subset of points $K \subset \mathbb{R}^n$ called the *feasible region*. If the problem is an unconstrained problem, the feasible region is $\mathbb{R}^n$. When the objective function is a polynomial $p(\mathbf{x})$, the problem is called a *polynomial optimization*. The solution of the optimization problem is denoted as $\mathbf{x}^* \equiv (x_1^*, x_2^*, \ldots, x_n^*)^T$, and the minimum value of the polynomial is denoted as $p(\mathbf{x}^*) = p^{\min}$. The problem has no solutions if the feasible region is empty (called the *infeasible* case) or if the objective function is unbounded below on the feasible region (called the *unbounded* case).

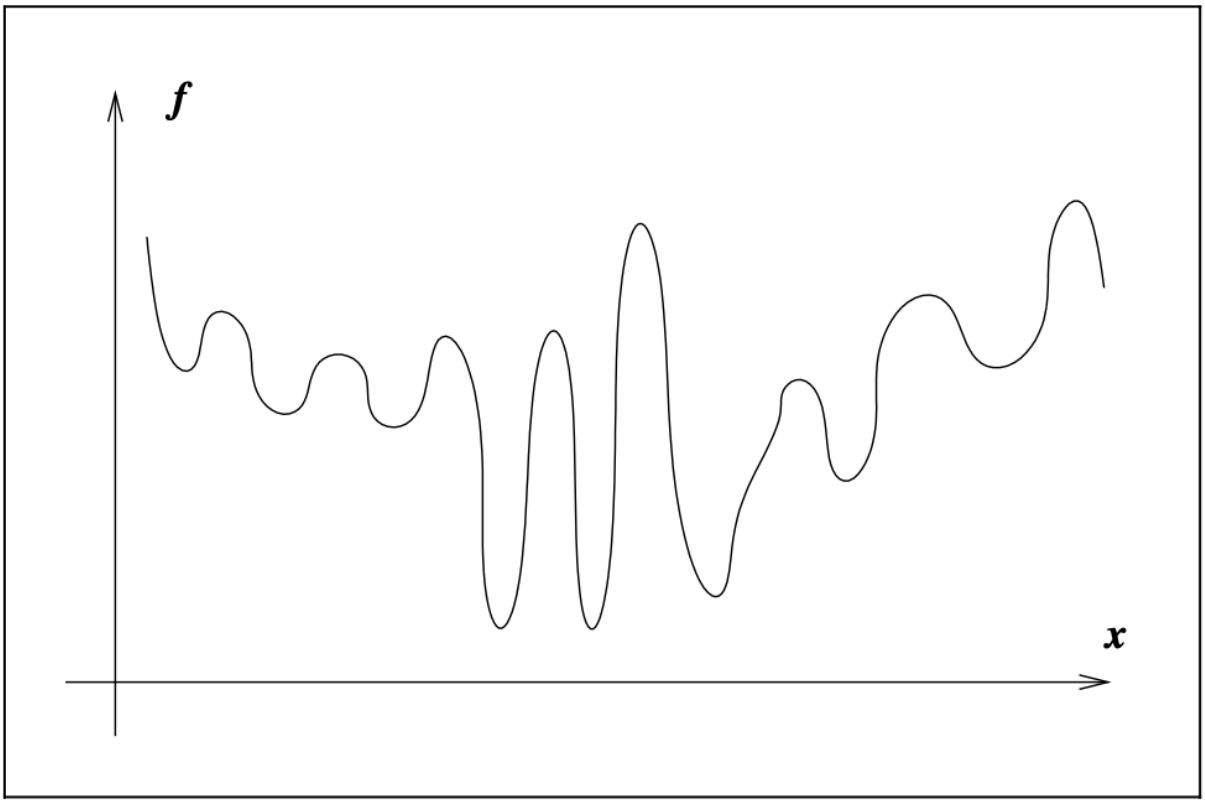

Figure 3.2: Graph of a function with many local minima.(Figure taken from [41])

As an example, consider the following problem:

$$\begin{array}{ll} \min\limits_{\mathbf{x} \in \mathbb{R}^2} & (x_1 - 2)^2 + (x_2 - 1)^2 \\ \text{s.t.} & x_1^2 - x_2 \leq 0, \\ & x_1 + x_2 \leq 2 \end{array} \tag{3.2}$$

We can rewrite it in the form of (3.1) by identifying

- $\mathbf{x} = (x_1, x_2)^T$,
- $f(\mathbf{x}) = (x_1 - 2)^2 + (x_2 - 1)^2$,
- $g_1(\mathbf{x}) = -x_1^2 + x_2$,
- $g_2(\mathbf{x}) = -x_1 - x_2 + 2$,
- $\mathcal{I} = \{1, 2\}$ and $\mathcal{E} = \emptyset$

A geometrical representation of the problem is shown on figure 3.1. The contours of f is the set of points for which $f(\mathbf{x})$ as a constant value.

Global solutions

Solutions $\mathbf{x}^* = (x_1^*, \ldots, x_n^*)$ that minimizes the value of the objective function can be either local solution, or global solutions. If $f(\mathbf{x}^*)$ is smaller than all other feasible nearby

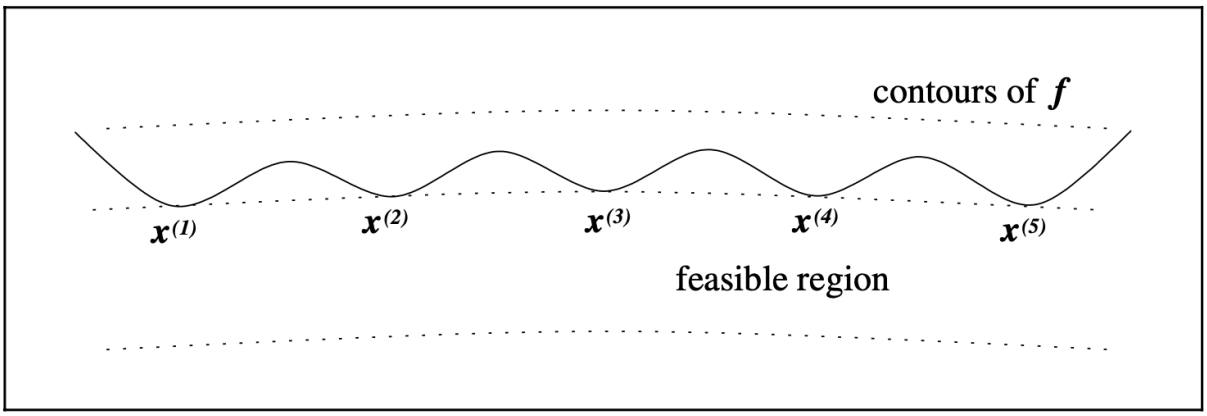

Figure 3.3: Representation of the minimization problem (3.3) where $\mathbf{x} = (\mathbf{x}^{(i)})$ is the vector of the local solutions (Figure taken from [41]).

points, the solution is called a *local* solution. If $f(\mathbf{x}^*)$ is the point with lowest function value among all feasible points, the solution is called a *global* solution. Depending on the problem, global solutions can be difficult to recognize and locate. Figure 3.2 shows a function with many local minimums, where programs that solve optimization problems tend to be "trapped".

Constraints might improve the situation. The feasible region they define may exclude local minima. It may then be easier to locate the global minima from the remaining minima. They, however, can make the problem more difficult. Indeed, consider the optimization

$$\begin{array}{ll} \underset{\mathbf{x} \in \mathbb{R}^2}{\text{minimize}} & (x_2 + 100)^2 + 0.01x_1^2 \\ \text{subject to} & x_2 - \cos(x_1) \geq 0, \quad . \end{array} \tag{3.3}$$

When one considers the problem without the constraint $g(\mathbf{x}) = x_2 - \cos(x_1) \geq 0$, it has the unique solution $(0, -100)^T$ while with the constraint, there are local solutions near the points $(k\pi, -1)$, for $k \in \{\pm 1, \pm 3, \pm 5, \dots\}$ as shown in Figure 3.3.

Convex programming

A set $\mathcal{S} \in \mathbb{R}^n$ is a *convex set* if for any two points $x, y \in \mathcal{S}$ we have

$$\alpha x + (1 - \alpha)y \in \mathcal{S}, \forall \alpha \in [0, 1]. \tag{3.4}$$

A function f is a *convex function* if its domain $\mathcal{S}$ is convex, and if for any two points $x, y \in \mathcal{S}$,

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y) \tag{3.5}$$

is satisfied $\forall \alpha \in [0, 1]$. If $-f$ is convex, then the function is said to be *concave*.

If the objective function f in the optimization problem (3.1) and its feasible region are both convex, any local solution of the problem is a global solution. As an example, the linear function $f(\mathbf{x}) = \mathbf{c}^T\mathbf{x} + \alpha$ for any constant vector $\mathbf{c} \in \mathbb{R}^n$ and α a scalar, is convex. The quadratic function $f(\mathbf{x}) = \mathbf{x}^T H\mathbf{x}$ where H is a symmetric positive semidefinite matrix is also convex. The problem (3.1) in which the objective function is convex, the equality constraint functions are linear, and the inequality constraint functions are concave, is called a *convex programming*. It is the latter that is of interest in this work.

Linear programming

A *linear* program is a an optimization problem where the objective function and all the constraints are linear functions of the variables. Every optimization problem of this kind can be written as [42]

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & A\mathbf{x} = \mathbf{b}, \\ & \mathbf{x} \geq 0. \end{aligned} \tag{3.6}$$

where $\mathbf{x}$ and $\mathbf{c}$ are vectors $\in \mathbb{R}^n$, $\mathbf{b}$ is a vector $\in \mathbb{R}^m$, and A is an $m \times n$ matrix. The inequality $\mathbf{x} \geq 0$ means that every component x_i of the vector $\mathbf{x}$ are ≥ 0. As an example

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^3} \quad & 3x_1 + 5x_2 + x_3 \\ \text{s.t.} \quad & x_1 + 3x_2 + 5x_3 = 2, \\ & x_1 + 9x_2 + 4x_3 = 1, \\ & x_1 \geq 0, \\ & x_2 \geq 0, \\ & x_3 \geq 0. \end{aligned} \tag{3.7}$$

where

- $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$ the vector which contains the parameters,
- $\mathbf{c} = \begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix} \in \mathbb{R}^3$ the vector which contains the coefficients,
- $A = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 9 & 4 \end{bmatrix}$ the 2×3 matrix which contains the coefficients for the inequality constraints,
- $\mathbf{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \in \mathbb{R}^2$ is the vector which contains the left side of the inequality constraints.

This formulation is called the *standard form*. Since $\mathbf{c}^T\mathbf{x}$ is a convex function, linear programming is a convex programming. If the objective function or some of the constraints are nonlinear functions of the variables, the problem is called a *nonlinear programming* problem.

3.1.2 Semidefinite programming

Semidefinite programming (SDP) is a generalization of linear programming. Consider X a symmetric $n \times n$ matrix. A linear function of X is a real valued function of the form

$$\sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} \tag{3.8}$$

with $c_{ij} \in \mathbb{R}$ $(i, j = 1, \ldots, n)$. Let C be the $n \times n$ matrix with element C_{ij} defined as $C_{ij} = c_{ij}$, $\forall i, j$. Equation (3.8) can be equivalently written as $\mathrm{Tr}(CX)$.

A *Semidefinite program* is an optimization problem of the form

$$\begin{aligned} \min_{X \in \mathcal{S}^n} \quad & \mathrm{Tr}(CX) \\ \text{s.t.} \quad & \mathrm{Tr}(A_i^T X) = b_i (i = 1, \ldots, m), \\ & X \succeq 0. \end{aligned} \tag{3.9}$$

where C, A_i $(i = 1, \ldots, m)$, and $\mathbf{b} = (b_1, \ldots, b_m)^T$ are given, and C is a symmetric matrix $\in M_n(\mathbb{R})$. The m linear equality constraints $\mathrm{Tr}(A_i^T X)$ are called *linear matrix inequalities* (LMI).

Semidefinite programs are convex programs, that is, the solution is the global minimum of the problem. They are a generalisation of linear programming. Indeed, on can always formulate (3.6) in the form (3.9). As an example, problem (3.7) can be rewritten in the form (3.9) where

- $X = \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix}$ is the 2×2 symmetric matrix which contains the parameters, with the identification $x_{11} = x_1, x_{12} = x_2$, and $x_{22} = x_3$,
- $C = \begin{bmatrix} 3 & 2,5 \\ 2,5 & 1 \end{bmatrix}$ is the 2×2 symmetric matrix which contains the coefficients of the inequality constraints,
- $A_1 = \begin{bmatrix} 1 & 3 \\ 0 & 5 \end{bmatrix}$, $A_2 = \begin{bmatrix} 1 & 9 \\ 0 & 4 \end{bmatrix}$, the 2×2 matrices which contain the coefficients for the inequality constraints,
- $\mathbf{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \in \mathbb{R}^2$ is the vector which contains the left side of the inequality constraints.

Dual theory of linear programming

The above formulation of the linear programs (3.6) and (3.9) is called the *primal* formulation. One can rewrite these problems in a so-called *dual* form. The dual semidefinite program of (3.9) is

$$\begin{aligned} \max_{\mathbf{y} \in \mathbb{R}^m} \quad & \mathbf{b}^T \mathbf{y} \\ \text{s.t.} \quad & \sum_{i=1}^{m} A_i y_i + S = C, \\ & S \succeq 0. \end{aligned} \tag{3.10}$$

Let p_P^* denote the optimal objective function value of the primal problem with solution X^*, and p_D^* the one for the dual problem with solution $(\mathbf{y}^*, S^*)$. The primal and dual problem are said *strictly feasible* if there exists such X^* for the primal problem, and $(\mathbf{y}^*, S^*)$ for the dual problem respectively. One has $p_P^* = -\infty$ if the primal problem is infeasible, and $p_D^* = \infty$ if the dual problem is infeasible. For any SDP, one has

$$p_P^* \leq p_D^*, \tag{3.11}$$

and

$$\mathrm{Tr}(CX^*) - \sum_{i=1}^{m} b_i y_i^* = \mathrm{Tr}(S^*X^*) \geq 0, \tag{3.12}$$

called the *duality gap*. If (3.12) holds, it is called a *weak duality*. It allows one to bound the optimal values of the dual problem by choosing a valid variable of the primal problem, and conversely. An important problem in dual theory is to identify sufficient conditions that ensure a zero duality gap, that is where (3.12) is a strict equality and called a *strong duality*, thus (3.11) is a strict equality. One can show that any SDP satisfying the following conditions, called the *Slater's conditions*, has a strong duality between its primal and dual problem [24] :

- If the primal (respectively the dual) problem is feasible and the dual (respectively the primal) problem is strictly feasible, then $p_P^* = p_D^*$, *i.e.*, the strong duality holds. There exist then a valid choice $X^* \succeq 0$ for the dual problem with $p_P^* = \langle C, X^* \rangle$ (respectively there exists a valid choice $(\mathbf{y}^*, S^* \succeq 0)$ for the primal problem with $p_D^* = \mathbf{b}^T\mathbf{y}^*$).
- If both primal and dual problems are strictly feasible, the strong duality holds and there exists valid choices $X^* \succeq 0$ and $(\mathbf{y}^*, S^* \succeq 0)$ such that $p_P^* = p_D^* = \langle C, X^* \rangle = \mathbf{b}^T\mathbf{y}^*$.

In other words, strong duality allows to identify the optimal value of an SDP by choosing valid variables of both primal and dual problem.

3.1.3 Moment relaxation for polynomial optimization

The global optimization of a polynomial $p(\mathbf{x}) = \mathbf{c}^T\mathbf{x}$

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \quad & p(\mathbf{x}) \\ \text{s.t.} \quad & h_i(\mathbf{x}) = 0, \\ & g_j(\mathbf{x}) \geq 0. \end{aligned} \tag{3.13}$$

where $h_i(\mathbf{x})$ and $g_j(\mathbf{x})$ are polynomials is NP-hard to solve in general [14], and as described in the previous section, when the problem is not convex (non linear), it may be hard for solvers to find and locate global minimum. However, one can approximate and reformulate the problem (3.13) in such a way that it becomes convex. This process is called a *convex relaxation*. Works in [35] describe how one can construct a hierarchy of convex relaxations using representations of nonnegative polynomials as sum of squares, and the dual theory of moments. Indeed, one can show that a sequence of moments of nonnegative measure corresponds to positive linear functionals on $\mathbb{R}[\mathbf{x}]$. The following approximation of the problem (3.13) was first proposed by Lasserre [35]. The presentation here follows the one described in [14]. Consider a polynomial $p(\mathbf{x})$ of degree d and the subset K defined by polynomial equalities and inequalities. For all points $\mathbf{x}^* \in K$ minimizing $p(\mathbf{x})$, and $d\mu(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}^*)d\mathbf{x}$ the Dirac measure, one can write

$$\begin{aligned} \int p(\mathbf{x}) d\mu(\mathbf{x}) &= \int p(\mathbf{x}) \delta(\mathbf{x} - \mathbf{x}^*) d\mathbf{x} \\ &= p(\mathbf{x}^*) \\ &= p^{min}, \end{aligned} \tag{3.14}$$

and thus

$$p^{min} \geq \min_{d\mu(\mathbf{x})} \int p(\mathbf{x}) d\mu(\mathbf{x}). \tag{3.15}$$

where the minimum is taken over all probability measure $d\mu(\mathbf{x})$ on $\mathbb{R}^n$. Since $p(\mathbf{x}) \geq p^{min}$ for all $\mathbf{x} \in K$, for $d\mu(\mathbf{x})$ a probability measure supported on K, *i.e.*, $\int_k d\mu(\mathbf{x}) = 1$,

$$\begin{aligned} \int_K p(\mathbf{x}) d\mu(\mathbf{x}) &\geq \int_K p^{min} d\mu(\mathbf{x}) \\ &= p^{min}, \end{aligned} \tag{3.16}$$

which leads to the following results

$$p^{min} = \min_{\mathbf{x} \in K} p(\mathbf{x}) = \min_{d\mu(\mathbf{x})} \int_K p(\mathbf{x}) d\mu(\mathbf{x}), \tag{3.17}$$

which, with $p(\mathbf{x}) = \sum_\alpha p_\alpha \mathbf{x}^\alpha$, can be written as

$$\begin{aligned} p^{min} = \min_{d\mu(\mathbf{x})} \int p(\mathbf{x}) d\mu(\mathbf{x}) &= \min_{d\mu(\mathbf{x})} \sum_\alpha \int p_\alpha \mathbf{x}^\alpha d\mu(\mathbf{x}) \\ &= \min_{(y_\alpha)_{\alpha \in \mathbb{N}^n_d}} \sum_\alpha p_\alpha y_\alpha \\ &= \min_{\mathbf{y} \in \mathbb{R}^{\mathbb{N}^n_d}} \mathbf{p}^T \mathbf{y}, \end{aligned} \tag{3.18}$$

where $\mathbf{p}$ is the column vector made of all the coefficient elements p_α of its sequence $(p_\alpha)_{\alpha \in \mathbb{N}^n_d}$, and $\mathbf{y}$ is the column vector made of the moments

$$y_\alpha = \int_K \mathbf{x}^\alpha d\mu(\mathbf{x})$$

in the moment sequence $y = (y_\alpha)_{\alpha \in \mathbb{N}^n_d}$ of the measure $d\mu(\mathbf{x})$, which is a functional of $\mathbf{x}^\alpha$. In other words, the problem of minimizing the polynomial $p(\mathbf{x})$ with solution $\mathbf{x}^*$ is equivalent to the problem of minimizing the linear functional $\mathbf{p}^T\mathbf{y}$, that is, solving the optimization problem

$$\begin{aligned} \min_{\mathbf{y} \in \mathbb{R}^{\mathbb{N}^n_d}} \quad & \mathbf{p}^T \mathbf{y} \\ \text{s.t.} \quad & y_0 = 1, \\ & y \text{ has a representing measure on } K, \end{aligned} \tag{3.19}$$

where the constraint $y_0 = y_{\alpha=(0,\ldots,0)} = 1$ ensures that $d\mu(\mathbf{x})$ is a probability measure on K,

$$\int_K y_0 d\mu(\mathbf{x}) = \int_K 1 d\mu(\mathbf{x}) = 1,$$

and with solution $\delta(\mathbf{x} - \mathbf{y}^*)d\mathbf{x}$. This in turns means that solving the K-truncated moment problem, that is to find a representing measure $d\mu(\mathbf{x})$ supported on a given subset K such that the moments y_α of the the given truncated moment sequence $y = (y_\alpha)_{\alpha \in \mathbb{N}^n_d}$ can be represented as $y_\alpha = \int_K \mathbf{x}^\alpha d\mu(\mathbf{x})$, amounts to solve the optimization problem (3.19). Indeed, if a minimizer $\mathbf{y}^*$ is found, y has a representing measure, as a constraint to the problem. If it does not exists such a representing measure, the optimization problem is infeasible, thus has no solution.

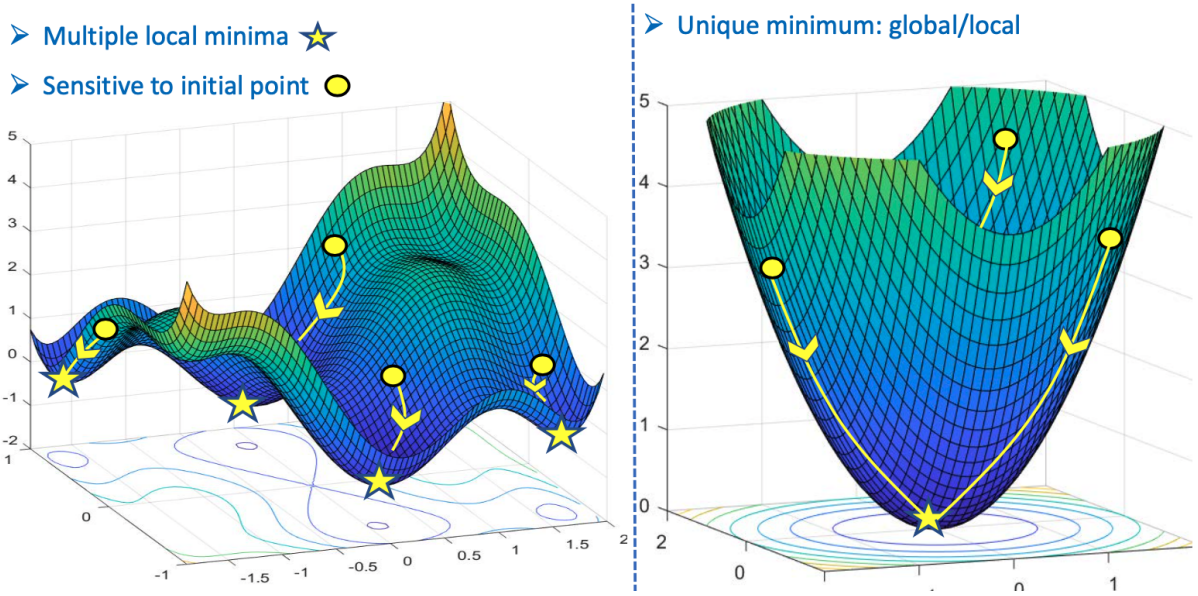

Figure 3.4: Graphs of a function with multiples local minima (on the right) and with a unique minima (on the left) (Figure taken from [43]).

3.1.4 Semidefinite algorithm for the K-tms problem

Recall Theorem 2.6: a truncated moment sequence $y = (y_\alpha)_{\alpha \in \mathbb{N}^n_d}$ admits a representing measure supported on the subset K defined by

$$K = \{\mathbf{x} \in \mathbb{R}^n | g_1(\mathbf{x}) \geq 0, \ldots, g_m(\mathbf{x}) \geq 0\},$$

where $g_i(\mathbf{x})$ are polynomials if and only if there exists an extension $(z)_{\beta \in \mathbb{N}^n_{2k}}$ with $2k > d$ such that $z_\alpha = y_\alpha$ for all $|\alpha| \leq d$, and that satisfies the rank condition $\text{rank}\,(M_k(z)) = \text{rank}\,(M_{k-d_0}(z))$ where

$$d_0 = \max_{1 \leq i \leq m} \{1, \lceil deg(g_i(\mathbf{x}))/2 \rceil\},$$

and such that localizing matrices $M_{k-d_{g_i}}(z)$ for $i \in \{1, \ldots, m\}$ are positive semi-definite. In other words, to solve the K-tms problem amounts to construct a positive semidefinite moment matrix $M_k(z)$ with some entry given by $z_\alpha = y_\alpha$ for $\alpha \in \mathbb{N}^n_d$, and with constraints on the moment matrices and localizing matrices linear in the z_α. This problem is a semidefinite optimization where the variables are the z_β for $\beta \in \mathbb{N}^n_{2k}$ with the smallest extension order $k_0 = \lfloor d/2 \rfloor + 1$. The flatness condition $\text{rank}\,(M_k(z)) = \text{rank}\,(M_{k-d_0}(z))$ with $d_0 = \max_{1 \leq i \leq m}\{1, \lceil deg(g(\mathbf{x})_i)/2 \rceil\}$ cannot be directly implemented in the SDP as a constraint. To implement the flatness condition, one can consider the semidefinite optimization problem [18]:

$$\begin{aligned} \min_{\mathbf{z} \in \mathbb{R}^{\mathbb{N}^n_k}} \quad & \mathbf{R}^T \mathbf{z} \\ \text{s.t.} \quad & y_0 = 1, \\ & M_k(z) \succeq 0, \\ & M_{k-d_i}(g_i * z) \succeq 0, (i \in \{1, \ldots, m\}), \\ & z_\alpha = y_\alpha \qquad \alpha \in \mathbb{N}^n_d, \end{aligned} \tag{3.20}$$

where $\mathbf{R}$ is the column vector made of the randomly chosen coefficients R_α of the polynomial

$$R(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}^n_{k_0}} R_\alpha x^\alpha, \tag{3.21}$$

where $R(\mathbf{x})$ is taken as a sum of square polynomial of degree $2k_0$ to ensure that $\mathbf{R}^T\mathbf{z}$ has a global minimum.

One can then implement an algorithm using the semidefinite optimization (3.20). The algorithm works as follow: first, the SDP runs for the extension order $k = k_0$, that is, the lowest order possible. There are then 3 cases depending on the outcome of the SDP:

- If the SDP is infeasible, meaning all constraints cannot be satisfied, the tms y admits no representing measure.
- If the SDP is feasible, and for that value k the rank condition $\operatorname{rank} M_k(z) = \operatorname{rank} M_{k-d_0}(z)$ is met, *i.e.*, there exist a flat extension z of y, the tms y admits a representing measure.
- If the SDP is feasible, but for that value k the rank condition $\operatorname{rank} M_k(z) = \operatorname{rank} M_{k-d_0}(z)$ is not met, there exists no flat extension z at the order k and the SDP remains inconclusive. One can then run another SDP with a different R_α, or increase the order k by one, and this until the SDP is either feasible or not.

According to Theorem 2.5, when a feasible flat extension z is found, z admits a representing measure composed of $r = \operatorname{rank} M_k(z)$ delta functions :

$$d\mu(\mathbf{x}) = \sum_{k=1}^{r} w_k \delta\left(\mathbf{x} - \mathbf{y}^*_k\right) d\mathbf{x} \tag{3.22}$$

where :

- r is finite,
- $w_j > 0$,
- $\mathbf{y^*}_k \in K$.

Suppose the SDP finds an extension $z = z^*$ at an order k that satisfies the rank condition, it is possible to obtain an explicit decomposition of $M_k(z^*)$ with rank r of the form

$$M_k(z^*)_{\alpha\beta} = \sum_{j=i}^{r} w_j (\mathbf{x}^*_j)^\alpha (\mathbf{x}^*_j)^\beta, \tag{3.23}$$

with $\alpha, \beta \in \mathbb{N}^n_k$, $w_j \geq 0$, and $\mathbf{x}^*_j \in K$. The method to do so is described in Appendix B.

3.2 Semidefinite optimization for the separability problem

3.2.1 Separability problem algorithm

Determining whether a state $\hat{\rho} \in \mathcal{P}(\mathcal{H})$ of a N-qudit system is separable, *i.e.*, if $\hat{\rho}$ can be written as a convex linear combination of pure states :

$$\hat{\rho} = \sum_k w_k \left(\rho_k^{(1)} \otimes \cdots \otimes \rho_k^{(N)} \right),$$

amounts to run the semidefinite optimization algorithm described above for the optimization problem (3.20), where the moments y_α are the real coordinates of $\hat{\rho}$ in the tensorial representation, and where the feasible region is the subset K defined by the polynomial constraints derived from the positivity of $\hat{\rho}$. In terms of separability of $\hat{\rho}$, the three different outcomes of an SDP for an order k are :

- If the SDP is infeasible, the state is entangled.
- If the SDP is feasible, and for that value k the rank condition $\text{rank}\,(M_k(z)) = \text{rank}\,(M_{k-d_0}(z))$ is met, the state is separable.
- If the SDP is feasible, but for that value k the rank condition $\text{rank}\,(M_k(z)) = \text{rank}\,(M_{k-d_0}(z))$ is not met, the problem remains inconclusive. One can then run another SDP with a different R_α, or increase the order k by one, and this until the SDP is either feasible or not.

Figure 3.5 illustrate the semidefinite algorithm for the separability problem.

If the state is separable, one can extract a the optimal values $\mathbf{y}_k^*$ whose entries are the elements of the convex combination

$$x_{\mu_1\mu_2\ldots\mu_N} = \sum_k w_k \left(y_{k;\mu_1}^{(1)} \ldots x_{y;\mu_N}^{(N)} \right),$$

that is, the elements of the Bloch vector of every pure state $\hat{\rho}_k^{(i)}, i \in \{1, \ldots, N\}$. As an example, for a 2 qubit system as presented in example (2.8), the optimal values $\mathbf{y}_k^*$ is a vector made of the 6 elements $\left(b_{k;1}^{(1)}, b_{k;2}^{(1)}, b_{k;3}^{(1)}, b_{k;1}^{(2)}, b_{k;2}^{(2)}, b_{k;3}^{(2)}\right)$ where the first 3 are the Bloch vector elements of the 1st qubit pure state k, and the last 3 are the Bloch vector of the 2^{nd}qubit pure states k. If the 2 qubit state is symmetric, then only one Bloch vector is given *i.e.*, $\mathbf{x}_k^*$ is of three variables, since they are indistinguishable from one qubit to another. The method to extract the optimal solutions is described in Appendix (B).

In summary, if the semidefinite optimization whose variables is the flat extension of the truncated moment sequence given by the local expectation values of the basis operators of a given state is feasible, then there exists a representing atomic probability measure. It then means that the global expectation values of the basis operators can be written as a convex combination of product of the local expectation values, which means that the state is a convex combination of product state, and is then separable. If the optimization is not feasible, the state is entangled.

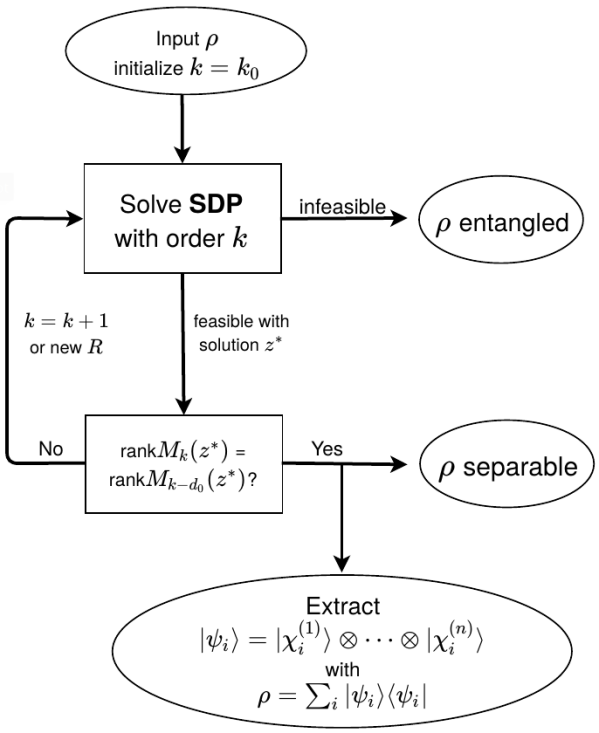

Figure 3.5: Illustration of the semidefinite algorithm for the separability problem.

Exemple 3.1. Consider a 2 qubit systems as described in example (2.8). There are 6 variables $(x_1, x_2, x_3, x_4, x_5, x_6)$ and 16 moment constraint defined by the real coordinates of the state in the tensorial representation. As an example, consider the randomly generated separated pure state $\hat{\rho}$

$$\begin{pmatrix} 0.41+0i & -0.16+0.12i & -0.24+0.32i & 0.001-0.198i \\ -0.16-0.12i & 0.1+0i & 0.19-0.056i & -0.06+0.08i \\ -0.24-0.32i & 0.19+0.056i & 0.39+0i & -0.15+0.115i \\ 0.001+0.198i & -0.06-0.08i & -0.15-0.115i & 0.1+0i \end{pmatrix} \tag{3.24}$$

It's real coordinates are

$$X = \begin{pmatrix} 1.0000 & -0.6368 & -0.4745 & 0.6077 \\ -0.6013 & 0.3829 & 0.2853 & -0.3654 \\ -0.7985 & 0.5085 & 0.3789 & -0.4852 \\ 0.0281 & -0.0179 & -0.0133 & 0.0171 \end{pmatrix} \tag{3.25}$$

The outcome of the SDP gives the optimal solutions

$$\mathbf{y}^* = \begin{pmatrix} -0.6 & -0.8 & 0.03 & -0.64 & -0.47 & 0.61 \end{pmatrix} \tag{3.26}$$

where the three first elements is the Bloch vector of the first qubit, and the last three the

one for the second qubit. It follows that

$$\begin{aligned}\rho^{(1)} =& \frac{1}{2}\hat{\mathbb{1}} + \frac{-0.6}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} + \frac{-0.8}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} + \frac{0.03}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\\ =& \begin{pmatrix}0.51+0i & -0.3+0.4i\\ -0.3-0.4i & 0.485+0i\end{pmatrix}\end{aligned}$$

and

$$\begin{aligned}\rho^{(2)} =& \frac{1}{2}\hat{\mathbb{1}} + \frac{-0.64}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} + \frac{-0.47}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} + \frac{0.61}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\\ =& \begin{pmatrix}0.805+0i & -0.32+0.24i\\ -0.32-0.24i & 0.195+0i\end{pmatrix},\end{aligned}$$

and one can check that $\hat{\rho} = \hat{\rho}^{(1)} \otimes \hat{\rho}^{(2)}$. For mixed states, one can construct each pure states similarly, where the weights can then be computed.

3.2.2 Results

The rest of this section present different results of the SDP algorithm for different quantum system. We implemented the algorithm in Matlab [44] using the package Gloptipoly 3.10 [45]. To solve the SDP, we use two different solvers, namely SeDuMi 1.3.5 [46] and Mosek 9.3.7 [47].

To test our algorithm, we created different kinds of states in the sense that we knew beforehand whether they are entangled or separable. The states were constructed using the functions in the package QUBIT4MATLAB V3.0 written by Géza toth [48] :

- **Separated pure states** : Created by making the tensor product of N random state vectors of dimension d. These random states are created from a uniform distribution on the complex sphere of radius 1 implemented in QUBIT4MATLAB. For symmetric states, the state is N times the tensor product of a single random state vector of dimension d.

- **Separated mixed states** : Created by making a random combination of p random separated pure states constructed as described above, with random weights. For symmetric states, the p random pure states are symmetric.

- **Entangled pure states** : We first created two random pure state vectors $|\phi_1\rangle$ and $|\phi_2\rangle$ of dimension d. We then created two tensor product states $|\psi_1\rangle$ and $|\psi_2\rangle$ constructed by making N times the tensor product of $|\phi_1\rangle$, and N times tensor product of $|\phi_2\rangle$ respectively. The two tensor product states are then summed and normalized. More specifically, the state created is

  $$\begin{aligned}|\psi\rangle &= \mathcal{N}(|\psi_1\rangle + |\psi_2\rangle)\\ &= \mathcal{N}(|\phi_1 \dots \phi_1\rangle + |\phi_2 \dots \phi_2\rangle)\end{aligned}$$

  with $\mathcal{N}$ a normalization constant. One can observe that $|\psi\rangle$ is symmetric.

- **Entangled mixed states** : Created by making a random combination of p random entangled states constructed as described above, with random weights.

Timings

We run the sdp algorithm for 100 of each of the 4 kinds of states above for qubit systems, up to 10 qubits. Every constructed separable states were detected as separable (the SDP was feasible), and every constructed entangled states were detected as entangled (the SDP was infeasible). We computed the average time it took for the algorithm to run for systems made of 2 to 10 symmetric qubits, and for the 4 kinds of states. The mixed states for both entangled and separable states are a mixture of 3 pure states, with random weights. When the SDP is inconclusive, up to 5 different R are tested before increasing the order k. For a set of 4 different kinds of states, the SDP algorithm was performed using the two different solvers SeDuMi and Mosek. They were performed on a laptop computer on a Linux distribution, equipped with a 2.6GHz processor and 16GB RAM. The results are displayed in the table below.

Table 3.1: Average time in seconds of the algorithm for $N \in \{2, \ldots, 10\}$ symmetric qubits.

		2	3	4	5	6	7	8	9	10
Sep. pure	Sedumi	0.13	0.15	0.25	0.42	1.1	2.38	5.89	15.45	33.5
	Mosek	0.18	0.18	0.21	0.29	0.45	0.9	2.09	4.9	12.11
Ent. pure	Sedumi	0.11	0.13	0.21	0.46	1.12	3.08	7.28	17.2	42.5
	Mosek	0.15	0.16	0.19	0.24	0.39	0.8	1.96	5.56	11.75
Sep. mixed	Sedumi	0.14	0.16	0.24	0.5	1.23	3.11	8.01	19.42	44.59
	Mosek	0.18	0.19	0.21	0.29	0.46	0.97	2.28	5.56	13.45
Ent. mixed	Sedumi	0.11	0.14	0.22	0.46	1.1	2.95	7.28	17.71	42.39
	Mosek	0.15	0.16	0.18	0.24	0.4	0.83	1.95	4.86	11.73

All entangled and separable states were detected by the algorithm in a single run. We observe that the performance of SeDuMi and Mosek are very similar for up to systems made of 4 qubits. For systems of > 4 qubits, Mosek starts to be more performant. It is more than 3 times faster for systems made of 10 Qubits.

For a set of 4 symmetric, the SDP was performed using both the symmetric method, and the general method to compare them for the same state. The results are displayed in the tables below for 2 Qubits, 3 Qubit, and 2 Qutrits.

We observe that the general method, that is when the variables of the subsystems are not equated, it takes significantly more time to detect entanglement/separability for systems of 3 qubits and 2 qutrits. We again observe that Mosek is significantly faster than SeDuMi when it comes to the general method, up to 20 times faster.

Table 3.2: Timing in seconds for a separable symmetric pure state solved with the general method and the symmetric method

Sep. pure		2 Qubits	3 qubits	2 qutrits
Symmetric Method	Sedumi	0.18	0.2	0.72
	Mosek	0.16	0.24	0.37
General Method	Sedumi	0.25	317.9	222.5
	Mosek	0.19	17.53	12.64

Table 3.3: Timing in seconds for a separable symmetric mixed state solved with the general method and the symmetric method

Sep. mixed		2 Qubits	3 qubits	2 qutrits
Symmetric Method	Sedumi	0.15	0.19	0.98
	Mosek	0.18	0.21	0.97
General Method	Sedumi	0.25	380.75	440.75
	Mosek	0.22	25.4	49.78

Table 3.4: Timing in seconds for a entangled symmetric pure state solved with the general method and the symmetric method

Ent. pure		2 Qubits	3 qubits	2 qutrits
Symmetric Method	Sedumi	0.14	0.15	0.48
	Mosek	0.16	0.16	0.24
General Method	Sedumi	0.2	320.95	272.26
	Mosek	0.16	17.44	12.34

Table 3.5: Timing in seconds for a entangled symmetric mixed state solved with the general method and the symmetric method

Ent. mixed		2 Qubits	3 qubits	2 qutrits
Symmetric Method	Sedumi	0.13	0.14	0.62
	Mosek	0.15	0.16	0.25
General Method	Sedumi	0.18	323.38	289
	Mosek	0.17	19.38	12.31

Probability of separability

We generated 2 random states chosen from two different measures, namely the *Hilbert-Schmidt measure* and the *Bures measure*. For 1000 random 2 qubit states tested on the H-S measure, 243 were separable, while 72 were separable for the Bures measure. Our results are consistent with results in [49] stating that 24,24 % of 2 qubit states for the H-S measure are separated, and 7,3% for the Bures measure. For 3 qubits, non were separable states for both measure. Even though our sample of states was only of 1000 random states, we can conclude that the number of separable states is close to 0. Results in [49] shows that a random state made of one qubit and one qutrit has a 3,7% probability to be separable for the H-S measure, and 0.1% for the Bures measure. It shows that the probability for random states to be separable significantly decreases with the size of the system. Our results of 0 separable states for 3 qubits system for both measures are consistent with this idea.

Conclusion

The separability has been widely studied since its discovery. Many criteria and algorithms for the separability of states have been developed. The aim of this work was to present a necessary and sufficient condition for separability of arbitrary states with an arbitrary number of constituents, and arbitrary symmetries between the subparts, by mapping it onto a truncated moment problem, and solved using semidefinite optimization.

In the first chapter, we presented the basic notions of quantum mechanics used in this manuscript. The first section summarized the concepts of Hilbert spaces, quantum states, operator spaces and how one can write any operators in the basis made of the GGM operators in a convenient tensor notation. A description of symmetric states was then exposed and how one can represent them in the symmetric subspace using the Dicke states. The notion of density operators and the difference between pure and mixed states was then exposed. The Bloch representation of qubit, qudit, and multiple qudit states was then introduced. The separability problem was then exposed and, with the criterion of the equivalence between two operators, a necessary and sufficient condition for separability of states in terms of product of local expectation values of the basis operators was obtained.

In the second chapter, algebraic preliminaries on monomials, polynomials, and matrices were presented, followed by the notions of moment and (truncated) moment sequences. The mapping between a state and a truncated moment sequence whose moments are the local expectation values of basis operators was then exposed. It lead to the equivalence that a separable state can be written as an integral form whose probability measure is an atomic measure, and where its first order moments are given by the local expectation values of the basis operators for a given state. The latter meant that a given state is separable if and only if there exists an atomic representing measure whose first order moments are given by the local expectation values of the state. This problem is called a truncated moment problem. To derive a necessary and sufficient condition on the existence of such a measure, we presented the notions of moment matrices, localizing matrices and flat extensions of truncated moment sequences. A necessary and sufficient condition on the separability of a state was then given.

The third chapter introduced the notions of optimizations problems, convex programming, linear programming and semidefinite programming. The dual theory of linear programming was then exposed, followed by the presentation on how one can solve a moment problem using semidefinite optimization. Solving a moment problem amounted to detect if an SDP whose variables are a flat extension of the truncated moment sequence given by the local expectation values of the basis operators for a given states is feasible, then there exists a representing measure, which meant that the state can be written as convex combination of product states, thus the state is separable. An algorithm was then presented to solve the separability problem. To conclude the chapter and this manuscript,

results on our implementation of the algorithm in Matlab were then exposed. We exposed the average timing it took for an SDP to be solved for 4 different kinds of states. We then presented our results on the probability that a random state is separable for two qubits and three qubits, for two different measures. When the dimension and the number of subparts of a system increase, we observed that the probability for a random state to be separable significantly decreases. For a 1000 random 3-qubit states tested, none were separable.

Mapping the separability problem onto a truncated moment problem provides a necessary and sufficient condition for a state to be separable. It provides a certificate of separability and easily accommodates with missing data, which makes it a powerful tool for applications and experiments. The idea of mapping the separability problem onto a truncated moment problem has also been applied to the more general problem of quantum channel separability [27].

The codes we developed within the framework of this master thesis could be further used in a variety of applications when detection of separability of mixed states is required. In particular when decoherence plays a significant role in the dynamics of multipartite systems.

Appendix A

Sets, basis, and truncated moment sequence

For $n = 3$ and $d = 4$,

$$\begin{aligned}\mathbb{N}_4^3 = \{&(0,0,0),(1,0,0),(0,1,0),(0,0,1),(2,0,0),(1,1,0),(1,0,1),(0,2,0),\\ &(0,1,1),(0,0,2),(3,0,0),(2,1,0),(2,0,1),(1,2,0),(1,1,1),(1,0,2),\\ &(0,3,0),(0,2,1),(0,1,2),(0,0,3),(4,0,0),(3,1,0),(3,0,1),(2,2,0),\\ &(2,1,1),(2,0,2),(1,0,3),(1,2,1),(1,1,2),(1,0,3),(0,4,0),(0,3,1),\\ &(0,2,2),(0,1,3),(0,0,4)\}\end{aligned} \tag{A.1}$$

$$\begin{aligned}\mathbb{T}_4^3 = \{&1, x_1, x_2, x_3, x_1^2, x_1x_2, x_1x_3, x_2^2, x_2x_3, x_3^2, x_1^3, x_1^2x_2, x_1^2x_3, x_1x_2^2,\\ &x_1x_2x_3, x_1x_3^2, x_2^3, x_2^2x_3, x_2x_3^2, x_3^3x_1^4, x_1^3x_2, x_1^3x_3, x_1^2x_2^2, x_1^2x_2x_3,\\ &x_1^2x_3^2, x_1x_2^3, x_1x_2^2x_3, x_1x_2x_3^2, x_1x_3^3, x_2^4, x_2^3x_3, x_2^2x_3^2, x_2x_3^3, x_3^4\}\end{aligned} \tag{A.2}$$

$$\begin{aligned}\mathcal{B}_4^3 = ((1),&(x_1,x_2,x_3),(x_1^2,x_1x_2,x_1x_3,x_2^2,x_2x_3,x_3^2),\\ &(x_1^3,x_1^2x_2,x_1^2x_3,x_1x_2^2,x_1x_2x_3,x_1x_3^2,x_2^3,x_2^2x_3,x_2x_3^2,x_3^3)\\ &(x_1^4,x_1^3x_2,x_1^3x_3,x_1^2x_2^2,x_1^2x_2x_3,x_1^2x_3^2,x_1x_2^3,x_1x_2^2x_3,x_1x_2x_3^2,x_1x_3^3,\\ &x_2^4,x_2^3x_3,x_2^2x_3^2,x_2x_3^3,x_3^4))\end{aligned} \tag{A.3}$$

$$\begin{aligned}y = (y_\alpha)_{\alpha\in\mathbb{N}_4^3} = \{&1, y_{100}, y_{010}, y_{001}, y_{200}, y_{110}, y_{101}, y_{020}, y_{011}, y_{002}, y_{300}, y_{210},\\ &y_{201}, y_{120}, y_{111}, y_{102}, y_{030}, y_{021}, y_{012}, y_{003}, y_{400}, y_{310}, y_{301},\\ &y_{220}, y_{211}, y_{202}, y_{103}, y_{121}, y_{112}, y_{103}, y_{040}, y_{031}, y_{022}, y_{013}, y_{004}\}\end{aligned} \tag{A.4}$$

Appendix B

Extracting globally optimal solutions

The content of this appendix comes from [50]. Let the tms z^* be a flat extension satisfying the rank condition $\operatorname{rank} M_k(z^*) = \operatorname{rank} M_{k-d_0}(z^*)$. Since the condition holds, z^* is the vector of a $\operatorname{rank} M_k(z^*)$-atomic measure supported on K. The moment matrix $M_k(z^*)$ can then be constructed as

$$M_k(z^*) = \sum_{j=1}^{r} B((\mathbf{x}_j^*)^\alpha) \left(B((\mathbf{x}_j^*)^\alpha)\right)^T = V^*(V^*)^T$$

for $\alpha \in \mathbb{N}_d^n$, where $r = \operatorname{rank} M_k(z^*)$ and

$$V^* = \begin{pmatrix} B_k((\mathbf{x}^*(1))^\alpha) & B_k((\mathbf{x}^*(2))^\alpha) & \dots & B((\mathbf{x}^*(r)))^\alpha \end{pmatrix} \tag{B.1}$$

for $B_k((\mathbf{x}_j^*)^\alpha)$ the column vector whose elements are made of the elements of the monomial basis in the same order, up to order k, and $x_j^*, j = \dots, r$ are r global minimizers of the objective function. One can extract a *Cholesky* factor V of $M_k(\mathbf{z}^*)$, that is, a matrix V with r columns such that

$$M_k(\mathbf{z}^*) = VV^T. \tag{B.2}$$

Since both V and V^* span the same linear subspace, the solution extraction algorithm amounts to transform V into V^* using suitable column operations. First, one can reduce V to a column echelon form as

$$U = \begin{bmatrix} 1 & & & & \\ x & & & & \\ 0 & 1 & & & \\ 0 & 0 & 1 & & \\ x & x & x & & \\ & \vdots & & \ddots & \\ 0 & 0 & 0 & \cdots & 1 \\ x & x & x & \cdots & x \\ & \vdots & & & \vdots \\ x & x & x & \cdots & x \end{bmatrix}. \tag{B.3}$$

Each row in U corresponds to a monomial $\mathbf{x}^\alpha$ in the monomial basis by construction of the moment matrix. The first non-zero elments in each columns, called the *pivot element*,

correspond to the monomials $\mathbf{x}^{\beta_j}, j = 1, 2, \ldots, r$ of the basis that generates the solutions. Let

$$w = \left[\mathbf{x}^{\beta_1}, \mathbf{x}^{\beta_2} \cdots \mathbf{x}^{\beta_r}\right]^T \tag{B.4}$$

denote the generating basis of the r solutions. For all solutions $\mathbf{x} = \mathbf{x}_j^*, j = 1, 2 \ldots, r$ one has

$$B_k((\mathbf{x}^*)^\alpha) = Uw. \tag{B.5}$$

Extracting the optimal solutions the amounts to solve (B.5), i.e., a polynomial system of equations. To solve these polynomial equations, one can extract from U the $r \times r$ multiplication matrix N_i made of the coefficients of the monomials $x_i\mathbf{x}^{\beta^j}, j = 1, \ldots, r$ in the generating basis for each first degree monomials $x_i, i = 1, 2, \ldots, n$, i.e.,

$$N_i w = x_i w. \tag{B.6}$$

One can show that the entries of the solutions $\mathbf{x}^*(j), j = 1, \ldots, r$ are common eigenvalues of multiplication matrices $N_i, i = 1, \ldots, n$. One can build a random combination of the N_i as

$$N = \sum_{i=1}^{n} \lambda_i N_i \tag{B.7}$$

for $\lambda_i, i = 1, \ldots, n$ are non-negative numbers such that $\sum_i \lambda_i = 1$. The i-th entry $\mathbf{x}_{j;i}^*$ of $\mathbf{x}_j^* \in \mathbb{R}^n$ is given by

$$\mathbf{x}_{j;i}^* = q_j^T N_i q_j, \tag{B.8}$$

where the q_i are the elements of an orthogonal matrix $Q = \begin{bmatrix} q_1 & q_2 & \ldots & q_r \end{bmatrix}$, i.e., $q_i^T q_i = 1$ and $q_i^T q_j = 0$ for $i \neq j$ such that

$$N = QTQ^T, \tag{B.9}$$

where T is an upper triangular matrix with eigenvalues of N sorted increasingly along the diagonal, also called the *Schur* decomposition. This procedure has been implemented in the Matlab package Gloptipoly 3 [45]

Bibliography

[1] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?," *Phys. Rev.*, vol. 47, pp. 777–780, May 1935.

[2] "Discussion of probability relations between separated systems," vol. 31, no. 4.

[3] J. S. Bell, "On the einstein podolsky rosen paradox," *Physics Physique Fizika*, vol. 1, pp. 195–200, Nov 1964.

[4] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels," *Phys. Rev. Lett.*, vol. 70, pp. 1895–1899, Mar 1993.

[5] A. K. Ekert, "Quantum cryptography based on bell's theorem," *Phys. Rev. Lett.*, vol. 67, pp. 661–663, Aug 1991.

[6] M. A. and N. I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.

[7] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement," *Reviews of Modern Physics*, vol. 81, pp. 865–942, jun 2009.

[8] D. Bruss, "Characterizing entanglement," *Journal of Mathematical Physics*, vol. 43, pp. 4237–4251, sep 2002.

[9] O. Gühne and G. Toth, "Entanglement detection," *Physics Reports*, vol. 474, pp. 1–75, apr 2009.

[10] M. Horodecki, P. Horodecki, and R. Horodecki, "Separability of mixed states: necessary and sufficient conditions," *Physics Letters A*, vol. 223, pp. 1–8, nov 1996.

[11] A. Neven and T. Bastin, "The quantum separability problem is a simultaneous hollowisation matrix analysis problem," *Journal of Physics A: Mathematical and Theoretical*, vol. 51, p. 315305, jun 2018.

[12] E. Wolfe and S. F. Yelin, "Certifying separability in symmetric mixed states of n qubits, and superradiance," *Phys. Rev. Lett.*, vol. 112, p. 140402, Apr 2014.

[13] F. Bohnet-Waldraff, D. Braun, and O. Giraud, "Entanglement and the truncated moment problem," *Physical Review A*, vol. 96, Sep 2017.

[14] M. Laurent, *Sums of squares, moment matrices and optimization over polynomials*, vol. 149, pp. 155–270. New York: Springer, 2009.

[15] K. Schmüdgen, "The k-moment problem for compact semi-algebraic sets," *Mathematische Annalen*, vol. 289, no. 1, pp. 203–206, 1991.

[16] R. E. Curto and L. A. Fialkow, "Truncated k-moment problems in several variables," *Journal of Operator Theory*, vol. 54, no. 1, pp. 189–226, 2005.

[17] J. W. Helton and J. Nie, "A semidefinite approach for truncated k-moment problems," *Foundations of Computational Mathematics*, vol. 12, pp. 851–881, Sept. 2012.

[18] J. Nie, "The a-truncated k-moment problem," *Foundations of Computational Mathematics*, vol. 14, pp. 1243–1276, Oct. 2014.

[19] G. Bernd and J. Matousek, *Approximation algorithms and semidefinite programming*. Heidelberg New York: Springer-Verlag Berlin Heidelberg, 2012.

[20] L. László, "On the shannon capcity of a graph," *Information Theory, IEEE Transactions on*, vol. 25, pp. 1 – 7, 02 1979.

[21] M. X. Goemans and D. P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," *J. ACM*, vol. 42, p. 1115–1145, nov 1995.

[22] R. L. Kosut, A. Shabani, and D. A. Lidar, "Robust quantum error correction via convex optimization," *Phys. Rev. Lett.*, vol. 100, p. 020502, Jan 2008.

[23] Y. Eldar, "A semidefinite programming approach to optimal unambiguous discrimination of quantum states," *IEEE Transactions on Information Theory*, vol. 49, pp. 446–456, feb 2003.

[24] X. Wang, "Semidefinite optimization for quantum information," July 2018.

[25] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, "Complete family of separability criteria," *Phys. Rev. A*, vol. 69, p. 022308, Feb 2004.

[26] F. Hulpke and D. Bruß, "A two-way algorithm for the entanglement problem," *Journal of Physics A: Mathematical and General*, vol. 38, pp. 5573–5579, jun 2005.

[27] N. Milazzo, D. Braun, and O. Giraud, "Truncated moment sequences and a solution to the channel separability problem," *Physical Review A*, vol. 102, nov 2020.

[28] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Mécanique quantique: Tome 1*. CNRS/EDP Sciences, 2018.

[29] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Mécanique quantique: Tome 2*. CNRS/EDP Sciences, 2018.

[30] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Mécanique quantique: Tome 3, Fermions, bosons, photons, corrélations et intrication*. CNRS/EDP Sciences, 2019.

[31] J. D. Cresser, "Quantum Physics Notes." `http://physics.mq.edu.au/~jcresser/Phys304/Handouts/QuantumPhysicsNotes.pdf`, August 2011.

[32] J. von Neumann, *Mathematical Foundations of Quantum Mechanics: New Edition*. Princeton University Press, 2018.

[33] R. A. Bertlmann and P. Krammer, "Bloch vectors for qudits," *Journal of Physics A: Mathematical and Theoretical*, vol. 41, p. 235303, may 2008.

[34] S. Hartmann, "Generalized dicke states," *Quantum information and computation*, vol. 16, 01 2012.

[35] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 796–817, 2001.

[36] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge New York: Cambridge University Press, 2012.

[37] G. Landi and A. Zampini, *Linear algebra and analytic geometry for physical sciences*. Cham, Switzerland: Springer, 2018.

[38] A. Kholevo, A. Holevo, M. Shirokov, and R. Werner, "On the notion of entanglement in hilbert spaces," *Russian Mathematical Surveys - RUSS MATH SURVEY-ENGL TR*, vol. 60, 04 2005.

[39] I. Bengtsson and K. Zyczkowski, "Geometry of quantum states: An introduction to quantum entanglement," *Geometry of Quantum States: An Introduction to Quantum Entanglement*, 01 2006.

[40] M. Laurent, "Revisiting two theorems of curto and fialkow on moment matrices," *Proceedings of the American Mathematical Society*, vol. 133, no. 10, pp. 2965–2976, 2005.

[41] J. Nocedal and S. J. Wright, *Numerical optimization*. New York: Springer, 2006.

[42] F. Bohnet-Waldraff, "Entanglement and quantumness-new numerical approaches," 2017.

[43] A. Jasour, ""risk aware and robust nonlinear planning", course notes for mit 16.s498, rarnop.mit.edu," 2019.

[44] MATLAB, *version 9.9.0 (R2020b)*. Natick, Massachusetts: The MathWorks Inc., 2020.

[45] D. Henrion, J.-B. Lasserre, and J. Lofberg, "GloptiPoly 3: moments, optimization and semidefinite programming," *Optimization Methods and Software*, vol. 24, pp. pp. 761–779, Aug. 2009.

[46] J. F. Sturm, "Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones," *Optimization methods and software*, vol. 11, no. 1-4, pp. 625–653, 1999.

[47] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 9.3.7*, 2019.

[48] G. Toth, "Qubit4matlab v3.0: A program package for quantum information science and quantum optics for matlab," *Computer Physics Communications*, vol. 179, pp. 430–437, sep 2008.

[49] A. Khvedelidze and I. Rogojin, "On the geometric probability of entangled mixed states," *Journal of Mathematical Sciences*, vol. 209, 09 2015.

[50] D. Henrion, *Positive polynomials in control*. Berlin New York, N.Y: Springer, 2005.