Bounding separable rank via polynomial optimization

Sander Gribling
IRIF, Université de Paris

Monique Laurent
CWI and Tilburg University

Andries Steenkamp
CWI

February 2, 2021

Contents

1 Introduction

For a given integer $d \in \mathbb{N}$, let $\mathcal{H}^d$ be the space of complex Hermitian $d \times d$ matrices. We consider the following subcone of $\mathcal{H}^d \otimes \mathcal{H}^d$:

$$\mathcal{SEP}_d := \text{cone}\{xx^* \otimes yy^* \colon x \in \mathbb{C}^d,\ y \in \mathbb{C}^d,\ \|x\| = \|y\| = 1\}, \tag{1}$$

also denoted simply as $\mathcal{SEP}$ when the dimension d is not important.This cone is of special interest in the area of quantum information theory: its elements are known as the *separable states* on $\mathcal{H}^d \otimes \mathcal{H}^d$ and a positive semidefinite matrix $\rho \in \mathcal{H}^d \otimes \mathcal{H}^d$ that does not belong to $\mathcal{SEP}$ is said to be *entangled*. Entangled states can be used to observe quantum, non-classical behaviours. It is therefore important to be able to decide membership in the cone $\mathcal{SEP}$.

Gurvits [?] has shown that linear optimization over the cone $\mathcal{SEP}$ is NP-hard. **to check**. Several necessary criteria are known; the positive partial transpose criterion [?, ?], TODO. Doherty, Parrilo, and Spedalieri derived a complete hierarchy of outer approximations to the set $\mathcal{SEP}$ [?]. One can interpret their hierarchy in the language of moments of distributions: ρ is separable if there exists an atomic measure on the bi-sphere whose fourth-degree moments agree with ρ (see Section 4.3 for details). A certificate of $\rho \notin \mathcal{SEP}$ is called an *entanglement witness*. Several of the above-mentioned criteria can provide such witnesses. Another way to obtain an entanglement witness is to find a linear functional $W \in \mathcal{H}^d \otimes \mathcal{H}^d$ such that $\text{Tr}(W\rho) > \max\{\text{Tr}(W\sigma) \colon \sigma \in \mathcal{SEP}\}$. This shows the importance of linear optimization over the set $\mathcal{SEP}$.

In this work we consider a related problem: given a state $\rho \in \mathcal{SEP}$, what is the smallest integer $r \in \mathbb{N}$ such that there exist vectors $a_1, \ldots, a_r, b_1, \ldots, b_r \in \mathbb{C}^d$ for which

$$\rho = \sum_{\ell=1}^{r} a_\ell a_\ell^* \otimes b_\ell b_\ell^*. \tag{2}$$

This smallest integer r is called the *separable rank* of ρ and denoted as $\text{rank}_{\text{sep}}(\rho)$. The separable rank has been previously studied in [?, ?, ?] (where it is sometimes also called the *length* of ρ). We approach the problem of determining the separable rank from the moment perspective. We use the observation that, if $\text{rank}_{\text{sep}}(\rho) = r$ and ρ admits the decomposition (2), then the sum of the r atomic measures at the vectors $(a_\ell, b_\ell) \in \mathbb{C}^d \times \mathbb{C}^d$ is a measure μ whose expectation $\int 1 d\mu$ is equal to r and whose fourth-degree moments correspond to the entries of ρ. Moreover, this measure may be assumed to be supported on the set

$$\mathcal{V}_\rho = \{(x, y) \in \mathbb{C}^d \times \mathbb{C}^d : \|x\|_\infty^2, \|y\|_\infty^2 \leq \sqrt{\rho_{\max}},\ xx^* \otimes yy^* \preceq \rho\}, \tag{3}$$

where $\rho_{\max}$ denotes the largest diagonal entry of ρ. We then obtain a lower bound on the separable rank of ρ by minimizing the expectation $\int 1\, d\mu$ over all measures μ that are supported on $\mathcal{V}_\rho$ and have fourth-degree moments corresponding to ρ. A similar approach has previously been used to study other "factorization ranks" such as the rank of tensors [], the nonnegative rank, the completely positive rank, the positive semidefinite rank, and the completely positive semidefinite rank [FP16, GdLL19]. It is important to note that viewing separability of ρ as a moment problem on the bi-sphere does not (straightforwardly) lead to a bound on the separable rank. Indeed, for a measure μ on the bi-sphere whose fourth-degree moments correspond to ρ we necessarily have $\int 1\, d\mu = \text{Tr}(\rho)$. It is thus important to use another scaling for the points (a_ℓ, b_ℓ) entering a separable decomposition of ρ.

When using moment methods one typically works with measures supported on semi-algebraic sets, i.e., sets described by polynomial inequalities on the variables. In our approach this is also the case. Indeed the set $\mathcal{V}_\rho$ is semi-algebraic since one can encode the condition $xx^* \otimes yy^* \preceq \rho$ by requiring all principal minors of $\rho - xx^* \otimes yy^*$ to be nonnegative. This would lead to a description of the set $\mathcal{V}_\rho$ with a number of polynomial constraints that is exponential in d. Instead, we will give a more direct way to deal with constraints of the form $G(x) \succeq 0$ for matrices $G(x)$ whose entries are polynomials in x. We show that all the usual properties of the Lasserre hierarchy are preserved (e.g., asymptotic convergence, finite convergence under "flatness"). In Section 4.2 we show how this type of constraints can also be used to strengthen bounds on the completely positive rank given in [GdLL19].

We then discuss several extensions of our approach and some connections to the Doherty-Parrilo-Spedalieri approach. Note that the separable rank of $\rho \in \mathcal{H}^d \otimes \mathcal{H}^d$ is upper bounded by $d^4 = \dim_{\mathbb{C}}(\mathcal{H}^d \otimes \mathcal{H}^d)$ (using Carathéodory's theorem). We can leverage this fact and the asymptotic convergence of our hierarchy of lower bounds to detect non-membership to $\mathcal{SEP}$: given a matrix $\rho \notin \mathcal{SEP}$, there is a level of our hierarchy which is infeasible or one that provides a lower bound on $\mathrm{rank}_{\mathrm{sep}}(\rho)$ which is strictly larger than d^4. Our hierarchy can thus also be used to provide a type of entanglement witnesses. We discuss this extension in **??**.

Our approach can also be adapted to the notion of *mixed separable rank*, where one tries to find factorizations of the form $\rho = \sum_{\ell=1}^r A_\ell \otimes B_\ell$ with A_ℓ, B_ℓ Hermitian positive semidefinite matrices and r as small as possible, see **??**. In [**?**] it was shown that, if ρ is diagonal, then its mixed separable rank is equal to the nonnegative rank of the associated $d \times d$ matrix (consisting of the diagonal entries of ρ). It is known that computing the nonnegative rank of a matrix is hard; NP-hardness was shown by Vavasis [Vav09], and more recently $\exists\mathbb{R}$-hardness was shown by Shitov [Shi16]. Hence the mixed separable rank is an NP-hard parameter.

Finally, recent work of Dressler, Nie, and Yang [] showed how to improve the Doherty-Parrilo-Spedalieri approach by observing a global symmetry: for any real number ϕ we have $xx^* = (e^{\mathbf{i}\phi}x)(e^{\mathbf{i}\phi}x)^*$ and of course $\|e^{\mathbf{i}\phi}x\| = \|x\|$. In the Doherty-Parrilo-Spedalieri approach we may thus restrict to the part of the bi-sphere where $x_1, y_1 \in \mathbb{R}$ and even $x_1, y_1 \geq 0$. The same degree of freedom is present in our moment approach and we may therefore also restrict the support of our measures. By doing so, we can work with a smaller number of real variables ($4d-2$ instead of $4d$), see Section 4.1.

SG: TODO: mention Harrow et al. paper

2 Preliminaries on polynomial optimization

2.1 Polynomials, linear functionals and moment matrices

We first fix some notation that we use throughout. $\mathbb{N}$ denotes the set of nonnegative integers. We set $[n] = \{1, 2, ..., n\}$ for an integer $n \geq 1$ and $|\alpha| = \sum_{i=1}^n \alpha_i$ for $\alpha \in \mathbb{N}^n$.

For a complex matrix X we denote its transpose by X^T and its conjugate transpose by X^*. For a scalar $a \in \mathbb{C}$ its conjugate is $a^* = \overline{a}$ and its modulus is $|a| = \sqrt{a^* a}$. The vector space $\mathbb{C}^n$ is equipped with the scalar product $\langle x, y \rangle = x^* y = \sum_{i=1}^n x_i^* y_j$ for $x, y \in \mathbb{C}^n$ and the Euclidean norm of $x \in \mathbb{C}^n$ is $\|x\| = \sqrt{x^* x}$. Analogously, $\mathbb{C}^{n\times n}$ is equipped with the trace inner product $\langle X, Y \rangle = \mathrm{Tr}(X^* Y) = \sum_{i,j=1}^n \overline{X}_{ij} Y_{ij}$ and $\|X\| = \sqrt{\langle X, X \rangle}$ for $X \in \mathbb{C}^{n\times n}$. A matrix $X \in \mathbb{C}^{n\times n}$ is called Hermitian if $X^* = X$ and we let $\mathcal{H}^n$ denote the space of complex Hermitian $n \times n$ matrices, A matrix $X \in \mathcal{H}^n$ is positive semidefinite (denoted $X \succeq 0$) if $v^*Av \geq 0$ for all $v \in \mathbb{C}^n$. We let $\mathcal{H}^n_+$ denote the cone of Hermitian positive semidefinite matrices.

For a set S in a vector space we let $\mathrm{cone}(S)$ and $\mathrm{conv}(S)$ denote, respectively, its conic hull and its convex hull.

Polynomials. We consider polynomials in n complex variables $x_1, \ldots, x_n$ and their conjugates $\overline{x_1}, \ldots, \overline{x_n}$. For $\alpha, \beta \in \mathbb{N}^n$ we use the short-hand $\mathbf{x}^\alpha \overline{\mathbf{x}}^\beta$ to denote the monomial

$$\mathbf{x}^\alpha \overline{\mathbf{x}}^\beta = \prod_{i=1}^n x_i^{\alpha_i} \prod_{j=1}^n \overline{x_j}^{\beta_j}.$$

The degree of this monomial, denoted by $\deg(\mathbf{x}^\alpha \overline{\mathbf{x}}^\beta)$, is equal to $|\alpha| + |\beta| = \sum_{i=1}^n \alpha_i + \beta_i$. We collect the set of all monomials of degree at most $t \in \mathbb{N} \cup \{\infty\}$ in the vector $[\mathbf{x}, \overline{\mathbf{x}}]_t$ (using some given ordering of the monomials) and also set $[\mathbf{x}, \overline{\mathbf{x}}] = [\mathbf{x}, \overline{\mathbf{x}}]_\infty$. We interpret $[\mathbf{x}, \overline{\mathbf{x}}]_t$ as a set when we write $\mathbf{x}^\alpha \overline{\mathbf{x}}^\beta \in [\mathbf{x}, \overline{\mathbf{x}}]_t$. Taking the complex linear span of all monomials in $[\mathbf{x}, \overline{\mathbf{x}}]_t$ gives the space of polynomials with complex coefficients and degree at most t:

$$\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_t := \mathrm{Span}\,\{m \mid m \in [\mathbf{x}, \overline{\mathbf{x}}]_t\} = \Big\{ \sum_{m \in [\mathbf{x}, \overline{\mathbf{x}}]_t} a_m m : a_m \in \mathbb{C} \Big\}.$$

For $t = \infty$ we obtain the full polynomial ring in $\mathbf{x}, \overline{\mathbf{x}}$ over $\mathbb{C}$, also denoted as $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$. So any polynomial $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ is of the form $p = \sum_{\alpha,\beta} p_{\alpha,\beta} \mathbf{x}^\alpha \overline{\mathbf{x}}^\beta$, where only finitely many coefficients $p_{\alpha,\beta}$ are nonzero; its *degree* is the maximum degree of the monomials occurring in p with a nonzero coefficient, i.e., $\deg(p) = \max_{p_{\alpha,\beta} \neq 0} \deg(\mathbf{x}^\alpha \overline{\mathbf{x}}^\beta)$. For convenience let $\mathbb{C}_0^{\mathbb{N}^n \times \mathbb{N}^n}$ denote the set of vectors $\mathbf{a} = (a_{\alpha,\beta})_{(\alpha,\beta) \in \mathbb{N}^n \times \mathbb{N}^n}$ that have only finitely many nonzero entries. Then any polynomial p can be written as $p = \mathbf{a}^*[\mathbf{x}, \overline{\mathbf{x}}]$, where we set $\mathbf{a} = (\overline{p}_{\alpha,\beta}) \in \mathbb{C}_0^{\mathbb{N}^n \times \mathbb{N}^n}$ (the conjugate of the vector of coefficients of p).

Conjugation on complex variables extends linearly to polynomials: for $p = \sum_{\alpha,\beta} p_{\alpha,\beta} \mathbf{x}^\alpha \overline{\mathbf{x}}^\beta$ we define its conjugate polynomial $\overline{p} = \sum_{\alpha,\beta} \overline{p}_{\alpha,\beta} \overline{\mathbf{x}}^\alpha \mathbf{x}^\beta$. Then, p is called *Hermitian* if $p = \overline{p}$. Hermitian polynomials only take real values: $p(x) \in \mathbb{R}$ for all $x \in \mathbb{C}^n$. We denote the space of Hermitian polynomials by $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h$. For instance, the polynomial $p = x + \overline{x}$ is Hermitian as well as $p = \mathbf{i}x - \mathbf{i}\overline{x}$, but $q = x - \overline{x}$ is not Hermitian (note $q(\mathbf{i}) = 2\mathbf{i} \notin \mathbb{R}$), where $\mathbf{i} = \sqrt{-1} \in \mathbb{C}$.

To capture positivity on the ring of polynomials, we work with the cone of Hermitian sums of squares, denoted by Σ. Any polynomial of the form $q\overline{q}$ (for some $q \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$) is called a *Hermitian square* and Σ is then the conic hull of Hermitian squares. For any integer $t \in \mathbb{N}$ we let $\Sigma_{2t} = \mathrm{cone}\{p\overline{p} \mid p \in [\mathbf{x}, \overline{\mathbf{x}}]_t\} = \Sigma \cap \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_{2t}$ denote the cone of Hermitian sums of squares with degree at most $2t$.

The dual space of polynomials. The algebraic dual of the ring of polynomials $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ is the vector space of all linear functionals on $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$. To clarify, a linear functional L on $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ is a linear map from $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ to $\mathbb{C}$. For every $t \in \mathbb{N} \cup \{\infty\}$ we denote the dual space of $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_t$ by $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_t^*$, defined as

$$\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_t^* = \{L : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_t \to \mathbb{C} \mid L \text{ is linear}\}.$$

We again abbreviate $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_\infty^*$ by $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^*$. A linear functional $L \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_t^*$ is called *Hermitian* if $L(\overline{p}) = \overline{L(p)}$ for all $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_t$. A (Hermitian) linear functional $L \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_{2t}^*$ is called *positive* if it maps Hermitian squares to nonnegative real numbers, i.e., if $L(p\overline{p}) \geq 0$ for all $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_t$.

Example of linear functionals. For any $a \in \mathbb{C}^n$ we can define the *evaluation functional at* a, denoted $L_a \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^*$, by

$$L_a(p) = p(a) \text{ for every } p \in C[\mathbf{x}, \overline{\mathbf{x}}].$$

It is easy to see that L_a is Hermitian and positive.

Linear functionals applied to polynomial-valued matrices. It will also be useful to apply linear functionals to polynomial-valued matrices entrywise. That is, for a polynomial-valued matrix $G \in (\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}])^{m\times m}$ and a linear functional $L \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^*$ we define

$$L(G) := \Big(L(G_{ij})\Big)_{i,j\in[m]} \in \mathbb{C}^{m\times m}.$$

Moment matrices. As an example of the above, applying a linear functional to the (infinite) matrix $[\mathbf{x}, \overline{\mathbf{x}}][\mathbf{x}, \overline{\mathbf{x}}]^*$ leads to the notion of moment matrix. Given $L \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^*_{2t}$, where $t \in \mathbb{N} \cup \{\infty\}$, we define the *moment matrix* of L by

$$M_t(L) := L([\mathbf{x}, \overline{\mathbf{x}}]_t[\mathbf{x}, \overline{\mathbf{x}}]_t^*) = \big(L(m\overline{m'})\big)_{m,m'\in[\mathbf{x},\overline{\mathbf{x}}]_t}.$$

If t is finite then the moment matrix is said to be *truncated* at *order* t. Note that L is Hermitian if and only if its moment matrix $M_t(L)$ is Hermitian. Similarly, L is positive if and only if its moment matrix $M_t(L)$ is positive semidefinite. Indeed, for any $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_t$, written as $p = \mathbf{a}^*[\mathbf{x}, \overline{\mathbf{x}}]_t \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ with $\mathbf{a} \in \mathbb{C}_0^{\mathbb{N}^n\times\mathbb{N}^n}$, we have $\overline{p} = [\mathbf{x}, \overline{\mathbf{x}}]_t^*\mathbf{a}$ and thus

$$L(p\overline{p}) = L(\mathbf{a}^*[\mathbf{x}, \overline{\mathbf{x}}]_t[\mathbf{x}, \overline{\mathbf{x}}]_t^*\mathbf{a}) = \mathbf{a}^*L([\mathbf{x}, \overline{\mathbf{x}}]_t[\mathbf{x}, \overline{\mathbf{x}}]_t^*)\mathbf{a} = \mathbf{a}^*M_t(L)\mathbf{a}. \tag{4}$$

If $t = \infty$ we write $M(L)$ instead of $M_\infty(L)$. Observe that the moment matrix of an evaluation functional L_a equals $[a, \overline{a}]_t[a, \overline{a}]_t^*$ and thus it has rank 1. Hence, if L is a linear combination of evaluation functionals, then its moment matrix has finite rank.

Polynomial localizing maps gL. Given a polynomial $g \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ and a linear functional $L \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^*$ we can define a new linear functional $gL \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^*$ by

$$\begin{aligned} gL : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] &\to \mathbb{C} \\ p &\mapsto L(gp). \end{aligned}$$

In this way, we can say that g acts on $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^*$ by mapping L to gL. Often constraints are phrased in terms of the positivity of gL. As stated before, positivity of gL can be characterized by positive semidefiniteness of its moment matrix:

$$gL \text{ is positive} \iff L(g \cdot [\mathbf{x}, \overline{\mathbf{x}}][\mathbf{x}, \overline{\mathbf{x}}]^*) = M(gL) \succeq 0. \tag{5}$$

If both g and L are Hermitian then gL is Hermitian and hence $M(gL)$ is Hermitian. If L is an evaluation map at a point $a \in \mathbb{C}^n$ for which $g(a) \geq 0$, then gL is a positive map since we have $(gL)(p\overline{p}) = g(a)|p(a)|^2 \geq 0$. In the literature $M(gL)$ is often called a *localizing moment matrix.*

Polynomial-matrix localizing maps $G \otimes L$. We introduce a natural generalization of the above notion of localizing map for polynomial-valued matrices. Consider a Hermitian matrix $G \in (\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}])^{m\times m}$ whose entries are polynomials in the complex variables $\mathbf{x}, \overline{\mathbf{x}}$. Since we already know how a single polynomial acts on $L \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^*$, the matrix-valued map

$$\begin{aligned} G \otimes L : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] &\to \mathbb{C}^{m\times m} \\ p &\mapsto \big((G_{ij}L)(p)\big)_{i,j\in[m]} \end{aligned}$$

is well-defined. Then $G \otimes L$ maps polynomials to $m \times m$ matrices and we have:

$$(G \otimes L)(p) = \big((G_{ij}L)(p)\big)_{i,j=1}^m = \big(L(G_{ij}p)\big)_{i,j=1}^m = L(Gp).$$

As before, we say that $G \otimes L$ is *positive* if it maps positive elements (i.e., Hermitian squares $p\overline{p}$) to positive elements (i.e., Hermitian positive semidefinite $m \times m$ matrices), i.e., if the following holds:

$$\big(G \otimes L\big)(p\overline{p}) = L(Gp\overline{p}) = \big(L(G_{ij}p\overline{p})\big)_{i,j=1}^m \succeq 0 \text{ for all } p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]. \tag{6}$$

Given the definition of the moment matrix $M(gL) = L(g[\mathbf{x}, \overline{\mathbf{x}}][\mathbf{x}, \overline{\mathbf{x}}]^*)$ it makes sense to define the *moment matrix* of $G \otimes L$ by

$$M(G \otimes L) := L(G \otimes [\mathbf{x}, \overline{\mathbf{x}}][\mathbf{x}, \overline{\mathbf{x}}]^*). \tag{7}$$

Remark 1. *When $L = L_a$ is the evaluation map at some vector $a \in \mathbb{C}^n$ the moment matrix $M(G \otimes L_a)$ has indeed a tensor product structure since we have*

$$M(G \otimes L_a) = L_a(G \otimes [\mathbf{x}, \overline{\mathbf{x}}][\mathbf{x}, \overline{\mathbf{x}}]^*) = G(a) \otimes [a, \overline{a}][a, \overline{a}]^* = L_a(G) \otimes L_a([\mathbf{x}, \overline{\mathbf{x}}][\mathbf{x}, \overline{\mathbf{x}}]^*).$$

In particular, if $G(a) \succeq 0$ then we have $M(G \otimes L_a) \succeq 0$. Therefore, $M(G \otimes L) \succeq 0$ when L is a conic combination of evaluation maps at points for which G is positive semidefinite.

The analogue of Eq. (5) does not extend to the matrix case: it does not seem to be the case that $G \otimes L$ is positive if and only if its moment matrix $M(G \otimes L)$ is positive semidefinite, only the 'if part' holds in general. To see this we present alternative characterizations for positivity of $G \otimes L$ and $M(G \otimes L)$.

Lemma 2. *$G \otimes L$ is positive, i.e., Eq. (6) holds, if and only if any of the following equivalent conditions holds:*

$$v^*L(Gp\overline{p})v = L(v^*Gv \cdot p\overline{p}) \geq 0 \textit{ for all } v \in \mathbb{C}^m \textit{ and } p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}], \tag{8}$$

$$M((v^*Gv)L) \succeq 0 \quad \textit{for all } v \in \mathbb{C}^m, \tag{9}$$

$$(v \otimes \mathbf{a})^* \, M(G \otimes L) \, (v \otimes \mathbf{a}) \geq 0 \textit{ for all } v \in \mathbb{C}^m \textit{ and } \mathbf{a} \in \mathbb{C}_0^{\mathbb{N}^n \times \mathbb{N}^n}. \tag{10}$$

Proof. The equivalence of Eq. (6) and Eq. (8) is clear. The equivalence of Eq. (8) and Eq. (9) follows using Eq. (4) applied to each localizing map $(v^*Gv)L$. To see the equivalence of Eq. (8) and Eq. (10), write any polynomial $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ as $p = \mathbf{a}^*[\mathbf{x}, \overline{\mathbf{x}}]$ with $\mathbf{a} = (a_{\alpha,\beta}) \in \mathbb{C}_0^{\mathbb{N}^n \times \mathbb{N}^n}$, and observe that, for any $v \in \mathbb{C}^m$, we have:

$$v^*L(Gp\overline{p})v = v^*L(G \cdot \mathbf{a}^*[\mathbf{x}, \overline{\mathbf{x}}][\mathbf{x}, \overline{\mathbf{x}}]^*\mathbf{a}) = (v \otimes \mathbf{a})^*L(G \otimes [\mathbf{x}, \overline{\mathbf{x}}][\mathbf{x}, \overline{\mathbf{x}}]^*)v \otimes \mathbf{a} = (v \otimes \mathbf{a})^*M(G \otimes L)v \otimes \mathbf{a},$$

using the definition of $M(G \otimes L)$ from Eq. (7). $\square$

Lemma 3. *$M(G \otimes L) \succeq 0$ if and only if any of the following equivalent conditions holds:*

$$w^* M(G \otimes L) w \geq 0 \text{ for all } w \in \mathbb{C}^m \otimes \mathbb{C}_0^{\mathbb{N}^n \times \mathbb{N}^n}, \tag{11}$$

$$L(\vec{\mathbf{p}}^* G \vec{\mathbf{p}}) \geq 0 \text{ for all } \vec{\mathbf{p}} = (p_1, \ldots, p_m) \in (\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}])^m. \tag{12}$$

Proof. Given a vector $w = (w_{i,(\alpha,\beta)})_{i,(\alpha,\beta)} \in \mathbb{C}^m \otimes \mathbb{C}_0^{\mathbb{N}^n \times \mathbb{N}^n}$ define, for each $i \in [m]$, the vector $\mathbf{a}_i = (w_{i,(\alpha,\beta)})_{(\alpha,\beta)} \in \mathbb{C}_0^{\mathbb{N}^n \times \mathbb{N}^n}$ and the corresponding polynomial $p_i = \mathbf{a}_i^* [\mathbf{x}, \overline{\mathbf{x}}]$, and define the vector $\vec{\mathbf{p}} = (p_1, \ldots, p_m) \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^m$. Then one can verify that

$$w^* M(G \otimes L) w = w^* L(G \otimes [\mathbf{x}, \overline{\mathbf{x}}][\mathbf{x}, \overline{\mathbf{x}}]^*) w = \sum_{i,j=1}^{m} G_{ij} p_i \overline{p}_j = \vec{\mathbf{p}}^* G \vec{\mathbf{p}},$$

from which follows the equivalence of Eq. (11) and Eq. (12). $\square$

Note that Eq. (10) is the restriction of Eq. (11), where we restrict to vectors w in tensor product form $w = v \otimes \mathbf{a}$, so Eq. (11) implies Eq. (10). In addition, we recover Eq. (8) if, in Eq. (12), we restrict the polynomials $p_1, \ldots, p_m$ to be of the form $p_i = v_i p$ (for $i \in [m]$) for some polynomial $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ and some vector $v = (v_1, \ldots, v_m) \in \mathbb{C}^m$. This shows again that Eq. (8) is more restrictive than Eq. (12). Summarizing we have the following implication.

Lemma 4. *If $M(G \otimes L) \succeq 0$ then $G \otimes L$ is positive.*

The above notion of localizing moment matrix can also be defined for truncated linear functionals, one only needs to suitably define its order. Let $L \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_{2t}$ with $t \in \mathbb{N}$ and $g \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_\delta$ be a polynomial with degree δ. Then the truncated localizing moment matrix of order $t - \lceil \delta/2 \rceil$:

$$M_{t - \lceil d/2 \rceil}(gL) = L(g[\mathbf{x}, \overline{\mathbf{x}}]_{t - \lceil \delta/2 \rceil} [\mathbf{x}, \overline{\mathbf{x}}]^*_{t - \lceil \delta/2 \rceil})$$

is well-defined since we apply L to polynomials of degree at most $\delta + 2(t - \lceil \delta/2 \rceil) \leq 2t$. In the same way, if G is a polynomial-valued matrix with degree δ, i.e., $\delta = \max_{i,j} \deg(G_{ij})$, then one can define the truncated moment matrix of order $t - \lceil \delta/2 \rceil$:

$$M_{t - \lceil \delta/2 \rceil}(G \otimes L) = L(G \otimes [\mathbf{x}, \overline{\mathbf{x}}]_{t - \lceil \delta/2 \rceil} [\mathbf{x}, \overline{\mathbf{x}}]^*_{t - \lceil \delta/2 \rceil}).$$

Remark 5. *Note that requiring positivity of the moment matrix $M(G \otimes L)$ not only provides a stronger condition than requiring positivity of $G \otimes L$, but it is also a condition that is computationally easier to check. To make this concrete we consider the truncated case when $G \in (\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_\delta)^{m \times m}$ and $L \in (\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_{2t+\delta})^*$, so that $M_t(G \otimes L)$ is well-defined. Then, the condition $M_t(G \otimes L) \succeq 0$ implies any of the following two equivalent conditions (the truncated analogs of* (8) *and* (9)*):*

$$L(v^* G v \cdot p\overline{p}) \geq 0 \;\; \text{for all } v \in \mathbb{C}^m \text{ and } p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_t, \tag{13}$$

$$M_t((v^* G v) L) \succeq 0 \;\; \text{for all } v \in \mathbb{C}^m. \tag{14}$$

Hence, while it is computationnally easy to check whether $M_t(G \otimes L) \succeq 0$, it is not clear whether one can check efficiently Eq. (14)*, which requires verifying positive semidefiniteness of infinitely many moment matrices. For this reason we will select the stronger moment matrix positivity condition when defining our new hierarchy of bounds for the separable rank, but we note that the weaker* *positivity condition of the localizing map will be sufficient to establish convergence properties of the bounds.*

Finally let us point out that positivity of the moment matrix $M(G \otimes L)$ is the analog on the moment side of the dual notion of sum-of-squares matrix (aka SoS-matrix) polynomials. Recall that a polynomial-valued matrix $S \in (\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_{2t})^{m\times m}$ is called a SoS-matrix polynomial if it is of the form $S = RR^*$ for some polynomial-valued matrix $R \in (\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_t)^{m\times k}$ and $k \in \mathbb{N}$ or, equivalently, if $S = \sum_{h=1}^{k} \vec{\mathbf{p_h}}\vec{\mathbf{p_h}}^*$ for some $\vec{\mathbf{p_h}} \in (\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_t)^m$ and $k \in \mathbb{N}$. From this follows that condition (12) is equivalent to

$$L(\langle G, S\rangle) \geq 0 \quad \textit{for any SoS-matrix polynomial } S \in (\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_{2t})^{m\times m},$$

i.e., L is nonnegative on the cone generated by $\langle G, S\rangle$ where $S \in (\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_{2t})^{m\times m}$ is a SoS-matrix polynomial. Expressing polynomials in this cone is often used in the literature (see, e.g., []), however the dual moment formulation given here had not been observed before to the best of our knoweldege.

Link to complete positivity of $G \otimes L$. SG: TODO ML: Maybe we defer this to a later 'concluding' section in order not to break the flow here?

2.2 The moment method

We now state several widely used definitions and results from polynomial optimization that we will need to design our hierarchy of bounds on the separable rank.

Given a set of Hermitian polynomials $S \subseteq \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h$ we define the *positivity domain* of S as

$$\mathscr{D}(S) := \{u \in \mathbb{C}^n \mid g(u) \geq 0 \text{ for every } g \in S\}. \tag{15}$$

Given a Hermitian polynomial-valued matrix $G \in (\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}])^{m\times m}$ we define the polynomial set

$$S_G = \{v^* G v : v \in \mathbb{C}^d, \|v\| = 1\} \subseteq \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h, \tag{16}$$

so that the set

$$\mathscr{D}(S_G) = \{u \in \mathbb{C}^n \mid G(u) \succeq 0\} \tag{17}$$

corresponds to the positivity domain of G. For $t \in \mathbb{N} \cup \{\infty\}$ and $S \subseteq \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h$ the set

$$\mathcal{M}_{2t}(S) := \text{cone}\{gp\overline{p} \mid p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}],\ g \in S \cup \{1\},\ \deg(gp\overline{p}) \leq 2t\}$$

denotes the *quadratic module* generated by S, *truncated* at order $2t$ when $t \in \mathbb{N}$. If $t = \infty$ we simply write $\mathcal{M}(S)$. The quadratic module $\mathcal{M}(S)$ is said to be *Archimedean* if, for some scalar $R > 0$,

$$R - \sum_{i=1}^{n} x_i \overline{x_i} \in \mathcal{M}(S). \tag{18}$$

Hence a quadratic module is Archimedean if it contains an algebraic certificate of boundedness of the associated positivity domain. The next lemma shows that, in the case when the algebraic certificate in (18) belongs to $\mathcal{M}_2(S)$, the linear functionals that are nonnegative on $\mathcal{M}(S)$ are bounded. Its proof is standard (and easy) and thus omitted.

Lemma 6. *Let $S \subseteq \mathbb{C}[\mathbf{x},\overline{\mathbf{x}}]^h$ be such that $R - \sum_{i=1}^n x_i\overline{x_i} \in \mathcal{M}_2(S)$ for some $R > 0$. For any $t \in \mathbb{N}$ assume $L_t \in \mathbb{C}[\mathbf{x},\overline{\mathbf{x}}]_{2t}^*$ is nonnegative on $\mathcal{M}_{2t}(S)$. Then we have*

$$|L_t(w)| \leq R^{|w|/2} L_t(1) \text{ for all } w \in [\mathbf{x},\overline{\mathbf{x}}]_{2t}.$$

Moreover, if

$$\sup_{t\in\mathbb{N}} L_t(1) < \infty, \tag{19}$$

then $\{L_t\}_{t\in\mathbb{N}}$ has a point-wise converging subsequence in $\mathbb{C}[\mathbf{x},\overline{\mathbf{x}}]^*$.

Linear functionals and measures. The following result is central to our approach. It is a complex analogue of a combination of results of Putinar [Put93] and Tchakaloff [Tch57]. In the real case this result can be found, e.g., in [GdLL19, Theorem 7], where a full proof is given. For completeness, in Appendix A we will show how to derive from it the following complex analogue[1].

Theorem 7. *Let $S \subseteq \mathbb{C}[\mathbf{x},\overline{\mathbf{x}}]^h$ such that $\mathcal{M}(S)$ is Archimedean and let $G \in (\mathbb{C}[\mathbf{x},\overline{\mathbf{x}}])^{m\times m}$ be a Hermitian polynomial-valued matrix. Assume $L \in \mathbb{C}[\mathbf{x},\overline{\mathbf{x}}]^*$ is nonnegative on $\mathcal{M}(S)$ and $G \otimes L$ is positive or, equivalently, L is nonnegative on $\mathcal{M}(S \cup S_G)$, with S_G as in Eq. (16). Then, for any integer $k \in \mathbb{N}$, there exists a linear functional $\widehat{L} \in \mathbb{C}[\mathbf{x},\overline{\mathbf{x}}]^*$ such that*

$$\widehat{L}(p) = L(p) \text{ for every } p \in \mathbb{C}[\mathbf{x},\overline{\mathbf{x}}]_k, \tag{20}$$

$$\widehat{L} = \sum_{\ell=1}^{K} \mu_\ell L_{v_\ell}, \tag{21}$$

for some integer $K \geq 1$, scalars $\mu_1, \mu_2, ..., \mu_K > 0$ and vectors $v_1, v_2, ..., v_K \in \mathscr{D}(S) \cap \mathscr{D}(S_G)$.

We also recall the following well-known result, which, under the 'flatness' condition (22), permits to claim existence of an atomic decomposition as in Eq. (21) for a truncated linear functional. We refer to Appendix A for details on how derive this complex version from the usual real version, as discussed, e.g., in [GdLL19].

Theorem 8. *Given integers $t \geq \delta \geq 1$, consider $L \in \mathbb{C}[\mathbf{x},\overline{\mathbf{x}}]_{2t}^*$, $S \subseteq \mathbb{C}[\mathbf{x},\overline{\mathbf{x}}]_{2\delta}^h$, a Hermitian polynomial matrix $G \in (\mathbb{C}[\mathbf{x},\overline{\mathbf{x}}]_{2\delta})^{m\times m}$ and the corresponding polynomial set S_G in Eq. (16). Assume that L is nonnegative on $\mathcal{M}_t(S \cup S_G)$ and that L is δ-flat, which means that it satisfies*

$$\mathrm{rank}(M_t(L)) = \mathrm{rank}(M_{t-\delta}(L)). \tag{22}$$

Then L can be extended to a linear functional $\widehat{L} \in \mathbb{C}[\mathbf{x},\overline{\mathbf{x}}]^*$, which is a conic combination of evaluation functionals at points in $\mathscr{D}(S) \cap \mathscr{D}(S_G)$. In other words, Eq. (20) holds for $k = 2t$ and Eq. (21) holds with all $\mu_\ell > 0$ and $v_\ell \in \mathscr{D}(S) \cap \mathscr{D}(S_G)$.

3 A hierarchy of lower bounds on the separable rank

In this section we show how to use polynomial optimization techniques to obtain a hierarchy of lower bounds on the separable rank.

[1]Of course this result remains valid when allowing several matrix-valued positivity constraints $G_\ell \otimes L$ instead of a single one, and the various polynomial-valued matrices G_ℓ may have different sizes. In addition it would suffice to assume that the quadratic module $\mathcal{M}(S \cup S_G)$ is Archimedean.

3.1 The parameter τ_{sep}

Consider a separable state $\rho \in \mathcal{SEP}_d$. As defined earlier, its separable rank is the smallest integer $r \in \mathbb{N}$ for which there exist (nonzero) vectors $a_1, \ldots, a_r, b_1, \ldots, b_r \in \mathbb{C}^d$ such that

$$\rho = \sum_{\ell=1}^{r} a_\ell a_\ell^* \otimes b_\ell b_\ell^*. \tag{23}$$

We mention several properties that are satisfied by the vectors a_ℓ, b_ℓ entering such a decomposition. First of all, the vectors a_ℓ, b_ℓ clearly satisfy the positivity condition

$$\rho - a_\ell a_\ell^* \otimes b_\ell b_\ell^* \succeq 0 \quad \text{for all } \ell \in [r]. \tag{24}$$

Let

$$\rho_{\max} = \max_{i,j \in [d]} \rho_{ij,ij}$$

denote the maximum diagonal entry of ρ. Then, in view of (24), the vectors a_ℓ, b_ℓ also satisfy $|(a_\ell)_i|^2 |(b_\ell)_j|^2 \le \rho_{ij,ij}$ for all $i, j \in [d]^2$, which implies the following boundedness conditions

$$\|a_\ell\|_\infty^2 \cdot \|b_\ell\|_\infty^2 \le \rho_{\max} \quad \text{and} \quad \|a_\ell\|_2^2 \cdot \|b_\ell\|_2^2 \le \mathrm{Tr}(\rho) \quad \text{for all } \ell \in [r].$$

Note that we may rescale the vectors a_ℓ, b_ℓ so that additional properties can be assumed. For instance we may rescale them so that $\|a_\ell\|_\infty = \|b_\ell\|_\infty$, in which case we may assume without loss of generality that

$$\|a_\ell\|_\infty^2, \|b_\ell\|_\infty^2 \le \sqrt{\rho_{\max}} \quad \text{for all } l \in [r]. \tag{25}$$

Another possibility is rescaling so that $\|a_\ell\|_2 = \|b_\ell\|_2$, in which case we could instead assume that

$$\|a_\ell\|_2^2, \|b_\ell\|_2^2 \le \sqrt{\mathrm{Tr}(\rho)} \quad \text{for all } l \in [r]. \tag{26}$$

Yet another possibility would be to rescale so that $\|b_\ell\|_2 = 1$ for all ℓ, in which case we would have[3]

$$\|b_\ell\|_2 = 1, \; \|a_\ell\|_2^2 \le \mathrm{Tr}(\rho) \; \text{ for all } l \in [r]. \tag{27}$$

To fix ideas we will assume we apply the first recaling (25), so that each (a_ℓ, b_ℓ) belongs to the set

$$\mathcal{V}_\rho := \Big\{ (x, y) \in \mathbb{C}^d \times \mathbb{C}^d \mid xx^* \otimes yy^* \preceq \rho, \; \|x\|_\infty, \|y\|_\infty \le \rho_{\max}^{1/4} \Big\}, \tag{28}$$

we will consider the impact of doing other rescalings later on in the paper. Thus we have

$$\frac{1}{r}\rho = \frac{1}{r} \sum_{\ell=1}^{r} a_\ell a_\ell^* \otimes b_\ell b_\ell^* \in \mathrm{conv}\{xx^* \otimes yy^* : (x, y) \in \mathcal{V}_\rho\},$$

which motivates defining the following parameter

$$\tau_{\mathrm{sep}}(\rho) := \inf \Big\{ \lambda : \lambda > 0, \frac{1}{\lambda}\rho \in \mathrm{conv}\{xx^* \otimes yy^* : (x, y) \in \mathcal{V}_\rho\} \Big\}. \tag{29}$$

From the above discussion this parameter gives a lower bound on the separable rank:

[2] ML: We could consider/try numerically adding these sharper constraints?

[3] ML: Not yet sure whether this is the best presentation, to be decided later when we see how the text develops..

Lemma 9. *For any $\rho \in \mathcal{SEP}_d$, we have $\tau_{sep}(\rho) \leq \mathrm{rank}_{\mathrm{sep}}(\rho)$. Moreover, if $\rho \notin \mathcal{SEP}_d$ then $\tau_{sep}(\rho) = \mathrm{rank}_{\mathrm{sep}}(\rho) = \infty$.*

The parameter $\tau_{sep}(\rho)$ does not seem any easier to compute than the separable rank. It however enjoys an additional convexity property that the combinatorial parameter $\mathrm{rank}_{\mathrm{sep}}(\rho)$ does not have. In the next section we will present a hierarchy of lower bounds on $\mathrm{rank}_{\mathrm{sep}}(\rho)$, constructed using tools from polynomial optimization. Since these bounds arise from convex programs they will in fact also lower bound the (weaker) parameter $\tau_{sep}(\rho)$.

3.2 Polynomial optimization approach for τ_{sep} and $\mathrm{rank}_{\mathrm{sep}}$

As above, let $\rho \in \mathcal{SEP}_d$ be given, together with a decomposition (23) with $r = \mathrm{rank}_{\mathrm{sep}}(\rho)$, where we assume that the points (a_ℓ, b_ℓ) belong to the set $\mathcal{V}_\rho$ in (28). We explain how to define bounds for $\mathrm{rank}_{\mathrm{sep}}(\rho)$ by using the moment method from Section 2.2.

For this let us consider the linear functional

$$L = \sum_{\ell=1}^{r} L_{(a_\ell, b_\ell)}, \tag{30}$$

the sum of the evaluation functionals at the points entering the decomposition (23). Then L acts on the polynomial space $\mathbb{C}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]$, where it is now convenient to denote the $2d$ variables as $\mathbf{x} = (x_1, \ldots, x_d)$ and $\mathbf{y} = (y_1, \ldots, y_d)$, corresponding to the 'bipartite' structure in Eq. (23). By construction, L corresponds to a finite atomic measure supported on the set $\mathcal{V}_\rho$. Moreover we have

$$L(1) = \sum_{\ell=1}^{r} L_{(a_\ell, b_\ell)}(1) = \sum_{\ell=1}^{r} 1 = r = \mathrm{rank}_{\mathrm{sep}}(\rho)$$

and the fourth-degree moments are given by the entries of ρ:

$$L(\mathbf{x}\mathbf{x}^* \otimes \mathbf{y}\mathbf{y}^*) = \rho.$$

In addition, since each (a_ℓ, b_ℓ) belongs to the set $\mathcal{V}_\rho$, it follows that

$$M(G_\rho \otimes L) = L(G_\rho \otimes [\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}][\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]^*) \succeq 0 \quad \text{and} \quad L \geq 0 \text{ on } \mathcal{M}(S_\rho),$$

after defining the polynomial-valued Hermitian matrix

$$G_\rho(\mathbf{x}, \mathbf{y}) := \rho - \mathbf{x}\mathbf{x}^* \otimes \mathbf{y}\mathbf{y}^* \in \mathbb{C}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_4^{d^2 \times d^2} \tag{31}$$

and the localizing set

$$S_\rho = \left\{ \sqrt{\rho_{\max}} - x_i \overline{x}_i, \sqrt{\rho_{\max}} - y_i \overline{y}_i : i \in [d] \right\} \subseteq \mathbb{C}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_2^h. \tag{32}$$

To see that $M(G_\rho \otimes L) \succeq 0$ use Remark 1. Recall also the definition of the set of localizing polynomials corresponding to (31):

$$S_{G_\rho} = \{ v^* G_\rho v : v \in \mathbb{C}^d \otimes \mathbb{C}^d \} \subseteq \mathbb{C}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_4^h.$$

Then, by construction, the combined positivity domains of the sets S and S_{G_ρ} recover $\mathcal{V}_\rho$:

$$\mathscr{D}(S_{G_\rho}) \cap \mathscr{D}(S_\rho) = \mathcal{V}_\rho = \Big\{(x, y) \in \mathbb{C}^d \times \mathbb{C}^d \mid xx^* \otimes yy^* \preceq \rho,\ \|x\|_\infty, \|y\|_\infty \leq \rho_{\max}^{1/4}\Big\}.$$

Moreover, let us recall for further reference that, as explained in Lemmas 2 and 4,

$$M(G_\rho \otimes L) \succeq 0 \implies L \geq 0 \text{ on } \mathcal{M}(S_{G_\rho}). \tag{33}$$

The above observations motivate introducing the following parameters: for any integer $t \in \mathbb{N} \cup \{\infty\}$, $t \geq 2$, define

$$\begin{aligned}\xi_t^{\rm sep}(\rho) := \inf \Big\{ L(1) \mid\ & L \in \mathbb{C}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_{2t}^*, \\ & L(xx^* \otimes yy^*) = \rho, \\ & L \geq 0 \text{ on } \mathcal{M}_{2t}(S), \\ & L(G_\rho \otimes [\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_{t-2} [\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_{t-2}^*) \succeq 0 \Big\}.\end{aligned} \tag{34}$$

For $t = \infty$ the parameter $\xi_\infty^{\rm sep}(\rho)$ involves linear functionals acting on the full polynomial space $\mathbb{C}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]$. In addition, we let $\xi_*^{\rm sep}(\rho)$ denote the parameter obtained by adding the constraint $\text{rank}(M(L)) < \infty$ to the definition of $\xi_\infty^{\rm sep}(\rho)$.

As is well-known, for finite $t \in \mathbb{N}$, the bound $\xi_t^{\rm sep}(\rho)$ can be expressed as a semidefinite program since nonnegativity of L on the truncated quadratic module $\mathcal{M}_{2t}(S_\rho)$ can be encoded through positive semidefiniteness of the moment matrix $M_t(L) = L([\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_t [\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_t^*)$ and the localizing moment matrices $M_{t-1}(gL) = L(g[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_{t-1} [\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_{t-1}^*)$ for all $g \in S_\rho$.

By the above discussion, for any $\rho \in \mathcal{SEP}_d$ we have the following chain of inequalities:

$$\xi_2^{\rm sep}(\rho) \leq \xi_3^{\rm sep}(\rho) \leq \cdots \leq \xi_\infty^{\rm sep}(\rho) \leq \xi_*^{\rm sep}(\rho) \leq \text{rank}_{\rm sep}(\rho) < \infty. \tag{35}$$

We will now show that the bounds $\xi_t^{\rm sep}(\rho)$ in fact converge to the parameter $\tau_{\rm sep}(\rho)$. In a first step we observe that the parameters $\xi_t^{\rm sep}(\rho)$ converge to $\xi_\infty^{\rm sep}(\rho)$ and after that we show that $\xi_\infty^{\rm sep}(\rho) = \xi_*^{\rm sep}(\rho) = \tau_{sep}(\rho)$.[4]

Lemma 10. *Let $\rho \in \mathcal{SEP}_d$. The infimum is attained in problem (34) for any integer $t \geq 2$ or $t = \infty$, and we have $\lim_{t\to\infty} \xi_t^{\rm sep}(\rho) = \xi_\infty^{\rm sep}(\rho)$.*

Proof. First we show that problem (34) attains its optimum. For this note that, in view of Eq. (35), we may restrict the optimization to linear functionals L satisfying $L(1) \leq \text{rank}_{\rm sep}(\rho)$. By the definition of S_ρ in (32), the quadratic module $\mathcal{M}(S_\rho)$ is Archimedean, with $R - \sum_{i=1}^d (x_i \overline{x}_i + y_i \overline{y}_i) \in \mathcal{M}_2(S_\rho)$ and $R = 2d$. As L is nonnegative on $\mathcal{M}_{2t}(S_\rho)$, we can apply Lemma 6 and conclude that

$$|L(w)| \leq R^{|w|/2} L(1) \ \text{ for any } \ w \in [\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_{2t}.$$

Hence we are optimizing a linear objective function over a compact set and thus the optimum is attained. So, for each integer $t \geq 2$, let L_t be an optimum solution of problem (34). As $\sup_t L_t(1) \leq \text{rank}_{\rm sep}(\rho) < \infty$, we can conclude from Lemma 6 that there exists a linear functional $L \in \mathbb{C}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]^*$ which is the limit of a subsequence of the sequence $(L_t)_t$. Then L is feasible for $\xi_\infty^{\rm sep}(\rho)$, which implies $\xi_\infty^{\rm sep}(\rho) \leq L(1) = \lim_{t\to\infty} L_t(1) = \lim_t \xi_t^{\rm sep}(\rho)$. Note that this L is optimal for $\xi_\infty^{\rm sep}(\rho)$. ∎

[4] ML: Be precise below what assumption we make on ρ; maybe some results hold also for ρ not separable?

Lemma 11. *For any $\rho \in \mathcal{H}^d \otimes \mathcal{H}^d$ we have $\xi_\infty^{\rm sep}(\rho) = \xi_*^{\rm sep}(\rho) = \tau_{\rm sep}(\rho)$.*

Proof. As $\xi_\infty^{\rm sep}(\rho) \le \xi_*^{\rm sep}(\rho)$ it suffices to show that $\xi_*^{\rm sep}(\rho) \le \tau_{sep}(\rho)$ and $\tau_{sep}(\rho) \le \xi_\infty^{\rm sep}(\rho)$.

First we show $\xi_*^{\rm sep}(\rho) \le \tau_{sep}(\rho)$. If $\tau_{sep}(\rho) = \infty$ there is nothing to prove. So assume we have a feasible solution: $\rho = \lambda \sum_{\ell=1}^K \mu_l a_\ell a_\ell^* \otimes b_\ell b_\ell^*$, where $(a_\ell, b_\ell) \in \mathcal{V}_\rho$, $\mu_\ell > 0$ and $\sum_\ell \mu_\ell = 1$. Define the linear functional $L = \lambda \sum_{\ell=1}^K \mu_\ell L_{(a_\ell,b_\ell)}$. Then L is feasible for $\xi_*^{\rm sep}(\rho)$ with $L(1) = \lambda$. Hence, $\xi_*^{\rm sep}(\rho) \le L(1) = \lambda$, which shows $\xi_*^{\rm sep}(\rho) \le \tau_{sep}(\rho)$.

Now we show $\tau_{\rm sep}(\rho) \le \xi_\infty^{\rm sep}(\rho)$. If $\xi_\infty^{\rm sep}(\rho) = \infty$ there is nothing to prove. So assume L is a feasible solution to $\xi_\infty^{\rm sep}(\rho)$. Then, in view of Eq. (33), $L \ge 0$ on $\mathcal{M}(S_\rho \cup S_{G_\rho})$. As $\mathcal{M}(S_\rho)$ is Archimedean we can apply Theorem 7 (with $k = 4$) and conclude that the restriction of L to $\mathbb{C}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_4$ is a conic combination of evaluations at points in $\mathscr{D}(S_\rho) \cap \mathscr{D}(S_{G_\rho}) = \mathcal{V}_\rho$. In other words, there exist $(a_\ell, b_\ell) \in \mathcal{V}_\rho$ and scalars $\mu_\ell > 0$ such that $L(p) = \sum_{\ell=1}^K \mu_\ell p(a_\ell, b_\ell)$ for any $p \in \mathbb{C}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_4$. In particular, we have $L(1) = \sum_{\ell=1}^K \mu_\ell$ and $\rho = L(\mathbf{x}\mathbf{x}^* \otimes \mathbf{y}\mathbf{y}^*) = \sum_{\ell=1}^K \mu_\ell \, a_\ell a_\ell^* \otimes b_\ell b_\ell^*$. This implies that $\frac{1}{L(1)}\rho$ belongs to $\mathrm{conv}\{xx^* \otimes yy^* : (x, y) \in \mathcal{V}_\rho\}$ and thus $\tau_{sep}(\rho) \le L(1)$, showing $\tau_{sep}(\rho) \le \xi_\infty^{\rm sep}(\rho)$. $\square$

As observed earlier already, since $\mathcal{SEP}_d$ is a d^4-dimensional cone, by the Carathéodory theorem we have $\mathrm{rank}_{\rm sep}(\rho) \le d^4$ for any $\rho \in \mathcal{SEP}_d$. Based on this one can also use the bounds $\xi_t^{\rm sep}(\rho)$ to test non-membership in $\mathcal{SEP}_d$: $\rho \notin \mathcal{SEP}_d$ if and only if $\xi_t^{\rm sep}(\rho) > d^4$ for some integer t.

Lemma 12. *Let $\rho \in \mathcal{H}^d \otimes \mathcal{H}^d$. Then, $\rho \in \mathcal{SEP}_d$ if and only if $\xi_t^{\rm sep}(\rho) \le d^4$ for all integers $t \ge 2$.*

Proof. The 'only if' part follows from the fact that $\xi_t^{\rm sep}(\rho) \le \mathrm{rank}_{\rm sep}(\rho) \le d^4$ when $\rho \in \mathcal{SEP}_d$. Conversely, assume $\xi_t^{\rm sep}(\rho) \le d^4$ for all integers $t \ge 2$. Then, we can use the same argument as in the proof of Lemma 10 and conclude the existence of $L \in \mathbb{C}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]^*$ feasible for $\xi_\infty^{\rm sep}(\rho)$, so that $\xi_\infty^{\rm sep}(\rho) \le L(1) < \infty$. Then, in view of Lemma 11, we have $\tau_{sep}(\rho) < \infty$, which shows that ρ is separable. $\square$

Finally, as a direct application of Theorem 8, for an integer $t \ge 2$ we can conclude finite convergence $\xi_t^{\rm sep}(\rho) = \tau_{\rm sep}(\rho)$ if we have a 2-flat optimal solution. Recall that a linear functional $L \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_{2t}^*$ is δ-flat if

$$\mathrm{rank}(M_t(L)) = \mathrm{rank}(M_{t-\delta}(L)).$$

In particular, 2-flatness is an efficiently verifiable property of an optimal solution to $\xi_t^{\rm sep}(\rho)$.

Lemma 13. *Let $\rho \in \mathcal{SEP}_d$ and $t \ge 2$. If $L \in \mathbb{C}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_{2t}^*$ is an optimal solution for $\xi_t^{\rm sep}(\rho)$ and 2-flat, then $\xi_t^{\rm sep}(\rho) = \tau_{\rm sep}(\rho)$.*

Proof. Recall that $\xi_t^{\rm sep}(\rho) \le \xi_*^{\rm sep}(\rho)$ for all $t \ge 2$, and in Lemma 11 we have seen $\xi_*^{\rm sep}(\rho) = \tau_{\rm sep}(A)$. Theorem 8 shows L can be extended to a linear functional $\widehat{L} \in \mathbb{C}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]^*$ which is a conic combination of evaluation functionals at points in $\mathcal{V}_\rho$. This means that $\widehat{L}$ is feasible for $\xi_*^{\rm sep}(\rho)$ and therefore $\xi_*^{\rm sep}(\rho) \le \xi_t^{\rm sep}(\rho)$. It follows that $\xi_t^{\rm sep}(\rho) = \tau_{\rm sep}(A)$. $\square$

4 Extensions and connections to other problems

Here we explain some simple extensions of the approach given in the previous section to related problems and we connect it to prior work on both factorization ranks and the set of separable states.

Multipartite separable states. The approach generalizes in a straightforward way to the separable rank of *multipartite* separable quantum states where for n-partite quantum states we require ρ to belong to the set

$$\operatorname{cone}\{x_1x_1^* \otimes x_2x_2^* \otimes \ldots \otimes x_nx_n^* : x_1,\ldots,x_n \in \mathbb{C}^d, \|x_i\| = 1 \ (i\in[n])\}.$$

In addition, one can use a different local dimension d_i for each part (i.e., $x_i \in \mathbb{C}^{d_i}$).

Mixed separable rank. In the definition of the set of separable quantum states we required a decomposition into *pure states*, of the form $xx^* \otimes yy^*$. Alternatively, one could define separability in terms of a decomposition using *mixed states* $X \otimes Y$ where $X, Y \in \mathcal{H}_+^d$. It is easy to see (using the spectral decomposition of a Hermitian positive semidefinite matrix) that $\mathcal{SEP}_d$ can alternatively be defined as

$$\mathcal{SEP}_d = \operatorname{cone}\{X \otimes Y : X, Y \in \mathcal{H}_+^d\}.$$

The associated factorization rank, called the *mixed separable rank* and denoted $\operatorname{rank}_{\rm sep}^{\rm mix}(\rho)$, is defined as the smallest integer $r \in \mathbb{N}$ for which there exist matrices $A_1,\ldots,A_r,B_1,\ldots,B_r \in \mathcal{H}_+^d$ such that

$$\rho = \sum_{\ell=1}^r A_\ell \otimes B_\ell.$$

The mixed separable rank has been studied previously in []. Using the same approach as in Section 3 one can define lower bounds on $\operatorname{rank}_{\rm sep}^{\rm mix}(\rho)$. Compared to the hierarchy of lower bounds on $\operatorname{rank}_{\rm sep}(\rho)$ we use more variables. A natural starting point is to introduce one variable for each entry of the matrices X and Y in $\mathcal{H}_+^d$ leading to a set of variables $\{x_{i,j} : i,j \in [d]\} \cup \{y_{i,j} : i,j\in[d]\}$. However, since the matrices are Hermitian we may work modulo the ideal corresponding to the equations $x_{i,j} = \overline{x_{j,i}}$ and $y_{i,j} = \overline{y_{j,i}}$ for all $i,j\in[d]$. This is equivalent to using the $2\binom{d+1}{2}$ variables

$$\{x_{i,j} : i,j \in [d], i \le j\} \cup \{y_{i,j} : i,j \in [d], i \le j\}. \tag{36}$$

In what follows we use X and Y to denote the $d \times d$ Hermitian matrices corresponding to the variables $x_{i,j}$ and $y_{i,j}$ ($i \le j$) respectively, that is the entries of X and Y are given by

$$X_{i,j} := \begin{cases} x_{i,j} & \text{if } i \le j \\ \overline{x_{j,i}} & \text{if } i > j \end{cases}, \qquad Y_{i,j} := \begin{cases} y_{i,j} & \text{if } i \le j \\ \overline{y_{j,i}} & \text{if } i > j \end{cases} \tag{37}$$

To make a converging hierarchy one needs to make sure that the variables $x_{i,j}$ and $y_{i,j}$ are bounded. One can do so analogously to Eq. (25). We first observe that $x_{i,i}y_{j,j} \le \rho_{\max}$ for all $i,j\in[d]$ and then use the fact that for positive semidefinite matrices the off-diagonal entries can be bounded by the diagonal entries $|x_{i,j}|^2 \le x_{i,i}x_{j,j}$. Together this allows one to add constraints analogous to Eq. (25):

$$|x_{i,j}|, |y_{i,j}| \le \sqrt{\rho_{\max}} \qquad \text{for all } i,j \in [d] \text{ with } i \le j. \tag{38}$$

Combining the above localizing conditions with the polynomial-matrix localizing conditions corresponding to X, Y, and $\rho - X \otimes Y$, this leads to bounds of the form

$$\begin{aligned}
\xi_t^{\mathrm{sep,mix}} := \inf \Big\{ L(1) \mid\ & L \in \mathbb{C}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_{2t}^*, \\
& L(X \otimes Y) = \rho, \\
& L \geq 0 \text{ on } \mathcal{M}_{2t}(\{\sqrt{\rho_{\max}} - x_{i,j}\overline{x_{i,j}}, \sqrt{\rho_{\max}} - y_{i,j}\overline{y_{i,j}} : i, j \in [d], i \leq j\}), \\
& L(X \otimes [\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_{t-1}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_{t-1}^*) \succeq 0, \\
& L(Y \otimes [\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_{t-1}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_{t-1}^*) \succeq 0, \\
& L((\rho - X \otimes Y) \otimes [\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_{t-1}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_{t-1}^*) \succeq 0 \Big\}.
\end{aligned} \tag{39}$$

Here we use the variables defined in Eq. (36) and the matrices X and Y defined through Eq. (37). Analogously to Section 3, these bounds converge to

$$\tau_{\mathrm{sep}}^{\mathrm{mix}}(\rho) := \inf \Big\{ \lambda : \lambda > 0, \frac{1}{\lambda}\rho \in \mathrm{conv}\{X \otimes Y : (X, Y) \in \mathcal{V}_{\rho,\mathrm{mix}}\} \Big\} \tag{40}$$

where

$$\mathcal{V}_{\rho,\mathrm{mix}} = \{(X, Y) : X, Y \in \mathcal{H}_+^d,\ |X_{i,j}|, |Y_{i,j}| \leq \sqrt{\rho_{\max}}\ (i, j \in [d]), \rho - X \otimes Y \succeq 0\}.$$

4.1 Reducing the number of variables

SG: TODO! copy-paste from the intro Finally, recent work of Dressler, Nie, and Yang [] showed how to improve the Doherty-Parrilo-Spedalieri approach by observing a global symmetry: for any real number ϕ we have $xx^* = (e^{\mathbf{i}\,\phi}x)(e^{\mathbf{i}\,\phi}x)^*$ and of course $\|e^{\mathbf{i}\,\phi}x\| = \|x\|$. In the Doherty-Parrilo-Spedalieri approach we may thus restrict to the part of the bi-sphere where $x_1, y_1 \in \mathbb{R}$ and even $x_1, y_1 \geq 0$. The same degree of freedom is present in our moment approach and we may therefore also restrict the support of our measures. By doing so, we can work with a smaller number of real variables ($4d - 2$ instead of $4d$), see Section 4.1.

4.2 Completely positive matrices / a connection to ξ_t^{cp}

For a given integer $d \in \mathbb{N}$, the cone of completely positive $d \times d$ matrices is defined as

$$\mathcal{CP}_d := \mathrm{cone}\{xx^T : x \in \mathbb{R}_{\geq 0}^d\}.$$

The cone of completely positive matrices and its dual, the cone of copositive matrices, are well known for their expressive power. For example, many NP-hard problems can be formulated as linear optimization problems over these cones [**?**, **?**]. We refer to [BSM03] for many structural properties about the cone $\mathcal{CP}_d$. As in the case of separable states, given a completely positive matrix A one can ask what is the smallest integer $r \in \mathbb{N}$ such that A admits a decomposition of the form

$$A = \sum_{\ell=1}^{r} a_\ell a_\ell^T$$

for entrywise nonnegative vectors $a_\ell \in \mathbb{R}^d_{\geq 0}$ ($\ell \in [r]$). The smallest such r is called the *completely positive rank* of A and denoted as $\mathrm{rank}_{\mathrm{cp}}(A)$. In [FP16] the authors defined the parameter $\tau_{\mathrm{cp}}(A)$ as

$$\tau_{\mathrm{cp}}(A) := \inf \left\{ \lambda : \lambda > 0, \frac{1}{\lambda}\rho \in \mathrm{conv}\{xx^T : x \in \mathbb{R}^d_{\geq 0},\ xx^T \leq A,\ xx^T \preceq A\} \right\} \tag{41}$$

to lower bound the completely positive rank (they also gave an SDP-relaxation $\tau^{\mathrm{sos}}_{\mathrm{cp}}(A)$). In [GdLL19] the first two authors, together with de Laat, studied (among others) the completely positive rank from the polynomial optimization perspective; we derived a hierarchy of semidefinite programming relaxations $\xi^{\mathrm{cp}}_t(A)$. There we used that if $xx^T \preceq A$, then also $(xx^T)^{\otimes \ell} \preceq A^{\otimes \ell}$ for all $\ell \in \mathbb{N}$ and therefore the following constraints are valid

$$L((xx^T)^{\otimes \ell}) \preceq A^{\otimes \ell} \qquad \text{for all } \ell \in \mathbb{N}. \tag{42}$$

In [GdLL19] it was shown that the following hierarchy converges to $\tau_{\mathrm{cp}}(A)$ as $t \to \infty$:

$$\begin{aligned} \xi^{\mathrm{cp}}_t(A) := \inf \Big\{ L(1) \mid\ & L \in [\mathbf{x}]^*_{2t}, \\ & L(xx^T) = A, \\ & L \geq 0 \text{ on } \mathcal{M}_{2t}(\{\sqrt{A_{ii}}x_i - x_i^2 : i \in [d]\}), \\ & L \geq 0 \text{ on } \mathcal{M}_{2t}(\{A_{ij} - x_i x_j : i,j \in [d], i \neq j\}), \\ & L((xx^T)^{\otimes \ell}) \preceq A^{\otimes \ell} \text{ for all } \ell \in [t] \Big\}. \end{aligned} \tag{43}$$

The same holds if we replace the last constraint with

$$L \geq 0 \qquad \text{on } \mathcal{M}_{2t}(\{v^T(A - xx^T)v : v \in \mathbb{R}^d\}). \tag{44}$$

Using the same reasoning as in Section 3.2, we see that we can strengthen $\xi^{\mathrm{cp}}_t(A)$ by adding the constraint

$$M((A - xx^*) \otimes L) := L((A - xx^T) \otimes [\mathbf{x}]_{t-1}[\mathbf{x}]_{t-1}^T) \succeq 0. \tag{45}$$

In particular, Lemmas 2 and 4 show that Eq. (45) implies Eq. (44). We now show that Eq. (45) in fact implies Eq. (42) which means that adding Eq. (45) strengthens both approaches provided in [GdLL19] (we present numerical examples in **??**). To do so, we introduce the following notation. Let $\langle x \rangle$ denote the vector of noncommutative monomials in the variables $x_1, \ldots, x_d$. Then we can define the *noncommutative localizing matrix*

$$M^{\mathrm{nc}}((A - xx^*) \otimes L) := L((A - xx^*) \otimes \langle x \rangle \langle x \rangle^*).$$

Note that $M((A - xx^*) \otimes L) \succeq 0$ if and only if $M^{\mathrm{nc}}((A - xx^*) \otimes L) \succeq 0$ (since the latter is obtained by duplicating rows/columns of the former).

Lemma 14. *For $A \in \mathbb{R}^{d \times d}$ and $L \in \mathbb{R}[\mathbf{x}]^*$ we have the following. If $L(xx^*) = A$ and $M((A - xx^*) \otimes L) \succeq 0$, then Eq. (42) holds, i.e.,*

$$L((xx^*)^{\otimes \ell}) \preceq A^{\otimes \ell} \textit{ for all } \ell \in \mathbb{N}.$$

Proof. As observed above, $M((A - xx^*) \otimes L) \succeq 0$ if and only if $M^{\mathrm{nc}}((A - xx^*) \otimes L) \succeq 0$. Note that for each $\ell \in \mathbb{N}$, the matrix $M^{\mathrm{nc}}((A - xx^*) \otimes L)$ contains $L((A - xx^*) \otimes (xx^*)^{\otimes \ell - 1})$ as a principal submatrix. Indeed, this follows from the fact that the vector $\langle x \rangle$ can be written as

$$1 \oplus_{\ell \in \mathbb{N}} x^{\otimes \ell}.$$

If we partition the matrix $M^{\mathrm{nc}}((A - xx^*) \otimes L)$ according to the degree of $\langle x \rangle$, we thus have that the matrices $L((A - xx^*) \otimes (xx^*)^{\otimes \ell - 1})$ form the diagonal blocks. Using that $M^{\mathrm{nc}}((A - xx^*) \otimes L) \succeq 0$, we thus see that

$$A \otimes L((xx^T)^{\otimes \ell - 1}) \succeq L((xx^T)^{\otimes \ell} \qquad \text{for all } \ell \in \mathbb{N}.$$

Combined with $L(xx^*) = A$ this shows Eq. (42):

$$L((xx^T)^{\otimes \ell} \preceq A \otimes L((xx^T)^{\otimes \ell - 1} \preceq A^{\otimes 2} \otimes L((xx^T)^{\otimes \ell - 2} \preceq \cdots \preceq A^{\otimes \ell - 1} \otimes L(xx^T) = A^{\otimes \ell}. \qquad \square$$

4.3 The Doherty-Parrilo-Spedalieri hierarchy as a moment problem on the bi-sphere

SG: TODO: copy-paste from the other pdf + improve!

5 Discussion / open questions

A Deriving the complex convergence results from their real version

References

[BSM03] A. Berman and N. Shaked-Monderer. *Completely Positive Matrices.* World Scienfic, 2003.

[FP16] H. Fawzi and P. A. Parrilo. Self-scaled bounds for atomic cone ranks: applications to nonnegative rank and cp-rank. *Mathematical Programming*, 158(1):417–465, 2016.

[GdLL19] S. Gribling, D. de Laat, and M. Laurent. Lower bounds on matrix factorization ranks via noncommutative polynomial optimization. *Foundations of Computational Mathematics*, Jan 2019.

[Put93] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42:969–984, 1993.

[Shi16] Y. Shitov. A short proof that NMF is NP-hard. *arXiv:1605.04000*, 2016.

[Tch57] V. Tchakaloff. Formules de cubature mécaniques à coéfficients non négatifs. *Bulletin des Sciences Mathématiques*, 81:123–134, 1957.

[Vav09] S. A. Vavasis. On the complexity of nonnegative matrix factorization. *SIAM Journal on Optimization*, 20(3):1364–1377, 2009.