Bounding the separable rank via polynomial optimization

Sander Gribling [a,1], Monique Laurent [b,*], Andries Steenkamp [c,2]

[a] *IRIF, Université de Paris, France*
[b] *CWI and Tilburg University, Netherlands*
[c] *CWI, Netherlands*

A R T I C L E   I N F O

A B S T R A C T

We investigate questions related to the set $\mathcal{SEP}_d$ consisting of the linear maps ρ acting on $\mathbb{C}^d \otimes \mathbb{C}^d$ that can be written as a convex combination of rank one matrices of the form $xx^* \otimes yy^*$. Such maps are known in quantum information theory as the separable bipartite states, while nonseparable states are called entangled. In particular we introduce bounds for the separable rank $\mathrm{rank}_{\mathrm{sep}}(\rho)$, defined as the smallest number of rank one states $xx^* \otimes yy^*$ entering the decomposition of a separable state ρ. Our approach relies on the moment method and yields a hierarchy of semidefinite-based lower bounds, that converges to a parameter $\tau_{\mathrm{sep}}(\rho)$, a natural convexification of the combinatorial parameter $\mathrm{rank}_{\mathrm{sep}}(\rho)$. A distinguishing feature is exploiting the positivity constraint $\rho - xx^* \otimes yy^* \succeq 0$ to impose positivity of a polynomial matrix localizing map, the dual notion of the notion of sum-of-squares polynomial matrices. Our approach extends naturally to the multipartite setting and to the real separable rank, and it permits strengthening some known bounds for the completely positive rank. In addition, we indicate how the moment approach also applies to define hierarchies of semidefinite relaxations for the set $\mathcal{SEP}_d$ and permits to give new proofs,

using only tools from moment theory, for convergence results
on the DPS hierarchy from Doherty et al. (2002) [16].

1. Introduction

The main object of study in this paper is the following matrix cone

$$\mathcal{SEP}_d := \text{cone}\{xx^* \otimes yy^* \colon x \in \mathbb{C}^d, \ y \in \mathbb{C}^d, \ \|x\| = \|y\| = 1\} \subseteq \mathcal{H}^d \otimes \mathcal{H}^d \simeq \mathcal{H}^{d^2}, \quad (1)$$

sometimes also denoted as $\mathcal{SEP}$ when the dimension d is not important. Throughout $\mathcal{H}^d$ denotes the cone of complex Hermitian $d \times d$ matrices and $\mathcal{H}^d_+$ is the subcone of Hermitian positive semidefinite matrices. Matrices in $\mathcal{H}^d_+$ are also known as *unnormalized states* and matrices in $\mathcal{H}^d_+$ with trace 1 are called *normalized states*. The cone $\mathcal{SEP}_d$ is of particular interest in the area of quantum information theory: its elements are known as the *(unnormalized, bipartite) 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 $\mathcal{SEP}_d$ is said to be *entangled*. Entangled states can be used to observe quantum, non-classical behaviors that may be displayed by two physically separated quantum systems, as already pointed out in the early work [1]. Entanglement is now recognized as an additional important resource that can be used in quantum information processing to carry out a great variety of tasks such as quantum computation, quantum communication, quantum cryptography and teleportation (see, e.g., [44,58] and references therein). Therefore, deciding whether a state is separable or entangled is a question of fundamental interest in quantum information theory. Gurvits [27] has shown that the (weak) membership problem for the set $\mathcal{SEP}_d \cap \{\rho : \text{Tr}(\rho) = 1\}$ is an NP-hard problem. In addition, the problem was shown to be strongly NP-hard in [24]. Hence it is important to have tractable criteria for separability or entanglement of quantum states. Throughout, we restrict for simplicity to the case of bipartite states, acting on two copies of $\mathbb{C}^d$, but the treatment extends naturally to the case of m-partite states that act on $\mathbb{C}^{d_1} \otimes \ldots \otimes \mathbb{C}^{d_m}$ with $m \geq 2$ and $d_1, \ldots, d_m$ possibly distinct. We will return below to the question of testing separability, but first we introduce the relevant notion of separable rank, which plays a central role in this paper.

The separable rank

In this work we consider the following problem: given a state $\rho \in \mathcal{SEP}_d$, 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 $\mathrm{rank}_{\mathrm{sep}}(\rho)$. One sets $\mathrm{rank}_{\mathrm{sep}}(\rho) = \infty$ when ρ is entangled. The separable rank has been previously studied, e.g., in [55,14,8] (where it is called the *optimal ensemble cardinality* or the *length* of ρ) and it can be seen as a 'complexity measure' of the state (with an infinite rank for entangled states). Easy bounds on the separable rank are $\mathrm{rank}(\rho) \leq \mathrm{rank}_{\mathrm{sep}}(\rho) \leq \mathrm{rank}(\rho)^2$, where the left most inequality can be strict (see [14]) and the right most one follows using Caratheodory's theorem [55]. We approach the problem of determining the separable rank from the moment perspective. We use the observation that, if $\mathrm{rank}_{\mathrm{sep}}(\rho) = r$ 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, as we will see later, this measure may be assumed to be supported on the semi-algebraic 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. Then we obtain a lower bound on the separable rank of ρ, denoted $\tau_{\mathrm{sep}}(\rho)$, 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 entries of ρ (see Eq. (38)). Hence, here we view the separable rank as a moment problem over the product of two balls. This view will enable us to design a hierarchy of tractable semidefinite based parameters, denoted $\xi_t^{\mathrm{sep}}(\rho)$. These parameters provide lower bounds on the separable rank and converge to $\tau_{\mathrm{sep}}(\rho)$ (see Section 3.2).

In view of the definition of $\mathcal{SEP}_d$ in Eq. (1), one may also view separability of a state ρ as a moment problem on the bi-sphere $\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}$, where $\mathbb{S}^{d-1} = \{x \in \mathbb{C}^d : \|x\| = 1\}$ denotes the (complex) unit sphere. However, this approach does *not* (straightforwardly) lead to bounds on the separable rank. Indeed, for a measure μ on the bi-sphere whose fourth-degree moments correspond to entries of ρ, we necessarily have $\int 1 d\mu = \mathrm{Tr}(\rho)$. To get bounds on the separable rank, it is thus crucial to use another scaling for the points (a_ℓ, b_ℓ) entering a separable decomposition of ρ as, for instance, the scaling used in Eq. (3), but other scalings are possible as indicated in Section 3.1.

Our approach extends to several other settings, in particular, to the case of multi-partite separable states (when ρ acts on the tensor product of more than two spaces) and to the case of real states (instead of complex valued ones). It 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. In [13] it was shown that, if ρ is a diagonal matrix, 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. Vavasis [56] has shown that computing the nonnegative rank of a matrix is an NP-hard problem and, more recently, Shitov [49] showed $\exists \mathbb{R}$-hardness of this problem. Hence computing the mixed separable rank has the same hardness complexity status as the nonnegative rank. Determining the complexity status of the separable rank

remains open, but there is no reason to expect that it should be any easier than the mixed separable rank.

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 $\mathcal{V}_\rho$ in Eq. (3) 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 however lead to a description of the set $\mathcal{V}_\rho$ with a number of polynomial constraints that is exponential in d. Instead, we will directly exploit the constraint $\rho - xx^* \otimes yy^* \succeq 0$, which is of the form $G(x) \succeq 0$ for some polynomial matrix $G(x)$ (i.e., with entries polynomials in $x, \overline{x}$). This constraint enables us to impose positivity constraints on polynomial matrix localizing maps, a matrix analog of the usual scalar localizing maps used in the moment method (see Section 2.2). Such polynomial matrix localizing constraints can also be used to bound the completely positive rank of a completely positive matrix, and we will show that this permits to strengthen some known bounds on the completely positive rank from [26] (see Section 4.3).

Our hierarchy of bounds $\xi_t^{\mathrm{sep}}(\rho)$ on the separable rank can also be used to detect entanglement. Indeed, as mentioned above, by Caratheodory's theorem, the separable rank of a state $\rho \in \mathcal{SEP}_d$ can be upper bounded, e.g., by $\mathrm{rank}(\rho)^2 \leq d^4$. We can leverage this fact and the asymptotic convergence of our hierarchy of lower bounds to detect entanglement: a state $\rho \in \mathcal{H}^d \otimes \mathcal{H}^d$ is entangled if and only if $\xi_t^{\mathrm{sep}}(\rho) > \mathrm{rank}(\rho)^2$ for some integer $t \geq 1$, i.e., there is a level of our hierarchy which is infeasible or provides a lower bound on $\mathrm{rank}_{\mathrm{sep}}(\rho)$ which is strictly larger than Caratheodory's bound. In addition, a certificate of entanglement is then provided by the dual semidefinite program. Hence our hierarchy of semidefinite parameters $\xi_t^{\mathrm{sep}}(\rho)$ can also be used to provide a type of entanglement witnesses (see Section 5.1).

The Doherty-Parrilo-Spedalieri (DPS) hierarchy for $\mathcal{SEP}$

As mentioned above a fundamental problem in quantum information theory is to have efficient criteria for checking separability or entanglement of quantum states. A second main contribution of our work concerns a hierarchy of outer approximations to the set $\mathcal{SEP}$ that we describe now. Doherty, Parrilo, and Spedalieri [16] designed what is now known as the *DPS hierarchy*, a hierarchy of outer approximations $\mathcal{DPS}_{1,t}$ $(t \geq 1)$ for the set $\mathcal{SEP}$. It is based on the principle of *state extension*: if $\rho := \sum_\ell \lambda_\ell x_\ell x_\ell^* \otimes y_\ell y_\ell^* \in \mathcal{SEP}_d$ with $\lambda_\ell \geq 0$ then, for any integer $t \geq 1$, ρ admits an extension $\rho_{1,t} := \sum_\ell \lambda_\ell x_\ell x_\ell^* \otimes (y_\ell y_\ell^*)^{\otimes t}$ acting on $\mathbb{C}^d \otimes (\mathbb{C}^d)^{\otimes t}$. The state ρ can be recovered from its extension $\rho_{1,t}$ by tracing out $t - 1$ of the copies of the second space and the extension $\rho_{1,t}$ satisfies several natural conditions such as symmetry (under permuting the t copies of the second register) and the so-called *positive partial transpose (PPT)* criterion from [31] (which states that taking the transpose of some of the copies preserves positive semidefiniteness). The relaxation $\mathcal{DPS}_{1,t}$ consists of those ρ for which a state $\rho_{1,t}$ exists satisfying these necessary conditions. Here the state extension is one-sided (since one extends only in the y-direction); the two-sided analog (in both x- and y-directions)

has also been considered, leading to the hierarchy $\mathcal{DPS}_{t,t} \subseteq \mathcal{DPS}_{1,t}$ (see Section 5.4 for details). For fixed t, deciding membership in $\mathcal{DPS}_{1,t}$ (or $\mathcal{DPS}_{t,t}$) boils down to testing feasibility of a semidefinite program of size polynomial in d. The DPS hierarchy is *complete*, in the sense that we have equality: $\bigcap_{t \geq 1} \mathcal{DPS}_{1,t} = \mathcal{SEP}_d$ [16].

One can also interpret the set $\mathcal{SEP}$ in the language of moments of distributions on the bi-sphere: ρ is separable if there exists an atomic measure on the bi-sphere whose fourth-degree moments agree with ρ (see, e.g., [17,29,38]). Another main contribution in this paper will be to make the links between this moment approach and the DPS hierarchy more apparent. These links enable us to give an alternative proof of completeness for the DPS hierarchy that is based on the theory of positive-operator valued measures. In contrast, existing proofs rely on other tools such as quantum de Finetti theorems or sums of squares. Indeed one can also design approximation hierarchies for $\mathcal{SEP}_d$, starting from its definition in Eq. (1) and applying the moment approach to the bi-sphere $\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}$. Depending on the degrees that are allowed in the $x, \overline{x}$ variables and the $y, \overline{y}$ variables, this leads to several possible variants of relaxations for $\mathcal{SEP}_d$ that we explore in Section 5.2, denoted there as $\mathcal{R}_t$ (when the full degree is at most $2t$), $\mathcal{R}_{t,t}$ (when the degree in $x, \overline{x}$ is at most $2t$ and the same for the degree in $y, \overline{y}$) and $\mathcal{R}_{1,t}$ (when the degree in $x, \overline{x}$ is at most 2 and the degree in $y, \overline{y}$ is at most $2t$). We provide a convergence proof for each of these hierarchies (i.e., show their completeness) using tools from the moment method (i.e., existence of an atomic representing measure under certain positivity conditions), which we apply to the setting of matrix polynomials for the hierarchy $\mathcal{R}_{1,t}$ (see Section 5.3). In addition we show that the hierarchy $\mathcal{R}_{1,t}$ (resp., $\mathcal{R}_{t,t}$) coincides with the DPS hierarchy $\mathcal{DPS}_{1,t}$ (resp., $\mathcal{DPS}_{t,t}$). Therefore we offer a new convergence proof for the DPS hierarchy that is based on the moment method.

Related literature on approximation hierarchies for $\mathcal{SEP}$

There is a vast literature about the set $\mathcal{SEP}$ of separable states and approximations thereof (such as the DPS hierarchy), so we only mention here some of the results that are most relevant to this paper. The PPT criterion, introduced in [46,31], is a necessary condition for separability. While it was shown to be sufficient to ensure separability of bipartite states acting on $\mathbb{C}^2 \otimes \mathbb{C}^3$ [59], it is in general not sufficient for separability of states acting on larger dimensional spaces (see, e.g., [32,59]). In fact it has been shown that no semidefinite representation exists for $\mathcal{SEP}_d$ when $d \geq 3$ [20]. As mentioned above, the authors in [16] use symmetric state extensions and the PPT conditions to define the hierarchy $\mathcal{DPS}_{1,t}$ ($t \geq 1$). They show it to be complete (i.e., $\cap_{t \geq 1} \mathcal{DPS}_{1,t} = \mathcal{SEP}$) using the quantum de Finetti theorem from [6] (note that this completeness proof in fact does not use the PPT conditions).

Navascues, Owari and Plenio [40] show a quantitative result on the convergence of the sets $\mathcal{DPS}_{1,t}$ to $\mathcal{SEP}_d$. Consider $\rho \in \mathcal{DPS}_{1,t}$, whose membership is certified by the extended state $\rho_{1,t}$ acting on $\mathbb{C}^d \otimes (\mathbb{C}^d)^{\otimes t}$, and let $\rho_1 \in \mathcal{H}^d$ be obtained by tracing out the part of $\rho_{1,t}$ that acts on $(\mathbb{C}^d)^{\otimes t}$; then $\rho_1 \otimes I_d$ is clearly separable. In [40] it is shown that

$$\tilde{\rho} := (1 - \epsilon)\rho + \epsilon\big(\rho_1 \otimes \frac{I_d}{d}\big) \in \mathcal{SEP}_d \quad \text{where } \epsilon = O\big(\big(\frac{d}{t}\big)^2\big); \tag{4}$$

that is, by moving ρ in the direction of $\rho_1 \otimes I_d/d$ by $\epsilon = O\big(\big(\frac{d}{t}\big)^2\big)$, one finds a separable state.

An *entanglement witness* for a state ρ is any certificate that certifies $\rho \notin \mathcal{SEP}_d$. One way to obtain such an entanglement witness is to exhibit one of the constraints defining a relaxation of $\mathcal{SEP}_d$ (such as $\mathcal{DPS}_{1,t}$) that is violated by ρ, like for example, one of the PPT conditions. More generally one can obtain an entanglement witness for $\rho \notin \mathcal{SEP}_d$ by finding a hyperplane separating ρ and $\mathcal{SEP}_d$, i.e., a matrix $W \in \mathcal{H}^d \otimes \mathcal{H}^d$ such that

$$\text{Tr}(W\rho) > h_{\mathcal{SEP}}(W) := \max\{\text{Tr}(W\sigma) \colon \sigma \in \mathcal{SEP}_d\}, \tag{5}$$

which shows again the importance of linear optimization over the set $\mathcal{SEP}_d$ and of designing tractable relaxations for $\mathcal{SEP}_d$. The function $h_{\mathcal{SEP}}(W)$ in Eq. (5) is known as the *support function* of $\mathcal{SEP}_d$ in the direction W. Analogously define the support function of $\mathcal{DPS}_{1,t}$ as

$$h_{\mathcal{DPS}_{1,t}}(W) := \max\{\text{Tr}(W\rho) : \rho \in \mathcal{DPS}_{1,t}\}.$$

As an application of the quantitative result in Eq. (4) the following is shown in [40]:

$$h_{\mathcal{SEP}}(W) \le h_{\mathcal{DPS}_{1,t}}(W) \le \big(1 + O\big(\big(\frac{d}{t}\big)^2\big)\big)h_{\mathcal{SEP}}(W). \tag{6}$$

Clearly, either Eq. (4) or Eq. (6) implies equality $\bigcap_{t \ge 1} \mathcal{DPS}_{1,t} = \mathcal{SEP}_d$, i.e., completeness of the DPS hierarchy.

Fang and Fawzi [19] investigate the DPS hierarchy from the dual sum-of-squares perspective. In particular, they show a representation result for matrix polynomials that are nonnegative on the sphere, which they use to give an alternative proof for Eq. (6). Namely, they show that, if F is a polynomial matrix in d variables and degree $2k$ such that $0 \preceq F(x) \preceq I$ on $\mathbb{S}^{d-1}$ then, for all $t \ge C_k d$, $F(x) + C_k'\big(\frac{d}{t}\big)^2 I$ is a Hermitian sum-of-squares matrix polynomial of degree $2t$ on $\mathbb{S}^{d-1}$, where C_k, C_k' are constants depending only k. In addition, detailed proofs are given in [19] for the description of the dual cones of the cones $\mathcal{DPS}_{1,t}$: while the dual cone of $\mathcal{SEP}_d$ consists of the matrices W for which the polynomial $p_W := \langle W, xx^* \otimes yy^* \rangle$ is nonnegative on the bi-sphere $\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}$, the dual cone of $\mathcal{DPS}_{1,t}$ consists of the W's for which the polynomial $\|y\|^{2(t-1)} p_W$ is a sum of Hermitian squares.

For the problem of approximating the support function $h_{\mathcal{SEP}}(W)$, Harrow, Natarayan and Wu [29] propose to strengthen the set $\mathcal{DPS}_{1,t}$ by adding equality constraints arising from the classical optimality conditions. In this way they obtain a hierarchy of bounds for $h_{\mathcal{SEP}}(W)$, stronger than $h_{\mathcal{DPS}_{1,t}}(W)$, that converges in finitely many steps to $h_{\mathcal{SEP}}(W)$.

Li and Ni [38] use the moment approach on the bi-sphere for testing separability of a state ρ (in the general multipartite setting). For this, given a generic sum-of-squares

polynomial F, they consider the problem of minimizing the expectation $\int F d\mu$ over the probability measures μ on the bi-sphere whose degree-4 moments correspond to the entries of ρ, and the corresponding moment relaxations (whose constraints are essentially those in the program defining the set $\mathcal{R}_t$ in Eq. (60)). Then a separability certificate can be obtained at a finite relaxation level when the optimal solution satisfies the so-called flatness condition. Note that the separability problem only asks for the existence of such a measure μ, thus, it is a feasibility problem. The optimization approach in [38], based on optimizing a generic polynomial F, relies on the fact that this 'encourages' flatness of an optimal solution (which then permits to get a separable decomposition and thus a certificate of separability). Indeed Nie [41] shows that if both the objective and constraints of a polynomial optimization problem are generic, then flatness occurs at some finite relaxation level. Dressler, Nie, and Yang [17] strengthen the approach in [38]: they use a symmetry argument which permits to replace the bi-sphere by its subset consisting of the points $(x, y) \in \mathbb{C}^d \times \mathbb{C}^d$ that have x_1, y_1 real and nonnegative. This provides a formulation that uses less real variables ($2(2d-1)$ instead of $4d$) and leads to stronger and more economical moment relaxations. Separability of real states is considered in [43], where a similar reduction is applied, namely by restricting to the vectors (x, y) in the (real) bi-sphere satisfying $\sum_{i=1}^d x_i \geq 0$ and $\sum_{i=1}^d y_i \geq 0$.

Related literature on factorization ranks

Various notions of "factorization ranks" have been studied extensively in the literature such as (versions of) tensor ranks [34], nonnegative matrix factorization (NMF) rank [25], positive semidefinite matrix factorization rank [21], completely positive matrix factorization rank [2]; we refer to these references and further references therein for details. Given the importance of factorizations for applications, designing algorithmic methods for finding a factorization of a given type (when it exists) is a topic of ongoing research (see, e.g., [25,51,52] and references therein). The above mentioned factorization ranks are often hard to compute (see [56,48,49] for nonnegative rank, [50] for positive semidefinite rank, [30] for tensor rank), which motivates the search for good bounds for a given factorization rank. Such bounds can be obtained using a variety of techniques. For example, using dedicated combinatorial methods (see, e.g., [21] and references therein), optimization methods (see, e.g., [22]), or using a moment-based approach as we do here. A moment-based approach has previously been used to derive hierarchies of bounds for the rank of tensors [53], for the symmetric nuclear norm of tensors [42], for the nonnegative rank, the completely positive rank, the positive semidefinite rank, and the completely positive semidefinite rank of matrices [26]. In this paper, we consider the separable rank, a notion which has been present in the (quantum information theory) literature, although no systematic study of bounds for it has been carried out so far to the best of our knowledge.

Contents of the paper

The paper is organized as follows. In Section 2 we introduce the preliminaries on polynomial optimization that we will need in the rest of the paper. In particular, in

Section 2.2, we introduce some of the main notions in the general setting of sum-of-squares matrix polynomials and matrix-valued linear maps. In Section 2.3 we recall the moment method and present the main underlying results from real algebraic geometry and moment theory. Since some of these results are presented in the literature in the real setting while we need the complex setting, we give arguments on how to extend the results from real to complex in Appendix A. Section 3 is devoted to the new hierarchy of bounds for the separable rank. In Section 4 we indicate several extensions of our approach, in particular for the real separable rank of real states and for getting improved bounds on the completely positive rank. We also present numerical results on examples to illustrate the behavior of the bounds in Section 4.2. Finally, in Section 5 we revisit the Doherty-Parrilo-Spedalieri hierarchy of relaxations for the set $\mathcal{SEP}$ of separable states. In particular, we provide a new, alternative proof for their completeness, that uses the tools from the moment approach previously developed.

2. Preliminaries on polynomial optimization

In this section, we group some preliminaries about polynomial optimization that we need in the rest of the paper; for a general reference we refer, e.g., to [35–37] and further references therein. We will deal with polynomial optimization in real and complex variables, which is the setting needed for the application to the set of separable states and the separable rank treated in this paper, and we will also need to deal with polynomial matrices and matrix-valued linear maps.

2.1. Polynomials, linear functionals and moment matrices

We first fix some notation that we use throughout the paper. $\mathbb{N}$ denotes the set of nonnegative integers. We set $[n] = \{1, 2, ..., n\}$ for an integer $n \geq 1$, $[k, n] = \{k, k + 1, \ldots, n - 1, n\}$ for integers $k \leq n$, 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 $\langle X, Y \rangle = \text{Tr}(X^*Y) = \sum_{i,j=1}^{n} \overline{X}_{ij} Y_{ij}$ 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 $\text{cone}(S)$ and $\text{conv}(S)$ 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:

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

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. 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 $\Sigma[\mathbf{x}, \overline{\mathbf{x}}]$ (or simply Σ) denotes the conic hull of Hermitian squares. For any integer $t \in \mathbb{N}$ we let $\Sigma[\mathbf{x}, \overline{\mathbf{x}}]_{2t} = \operatorname{cone}\{p\overline{p} \mid p \in [\mathbf{x}, \overline{\mathbf{x}}]_t\} = \Sigma[\mathbf{x}, \overline{\mathbf{x}}] \cap \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_{2t}$ (or simply Σ_{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

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

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 matrices.   It will also be useful to apply linear functionals to polynomial matrices, i.e., matrices whose entries are polynomials, by considering an entrywise action. That is, for a polynomial matrix $G = (G_{ij})_{i,j=1}^m \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^{m \times m}$ 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, 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

$$M_t(L) := L([\mathbf{x}, \overline{\mathbf{x}}]_t[\mathbf{x}, \overline{\mathbf{x}}]_t^*) = \big(L(m\overline{m'}) \big)_{m,m' \in [\mathbf{x}, \overline{\mathbf{x}}]_t}. \tag{7}$$

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:

$$L(p\overline{p}) \geq 0 \ \forall p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \iff M_t(L) \succeq 0. \tag{8}$$

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

$$L(p\overline{p}) = L(\mathbf{a}^*[\mathbf{x}, \overline{\mathbf{x}}]_t[\mathbf{x}, \overline{\mathbf{x}}]_t^* \mathbf{a}) = \mathbf{a}^* L([\mathbf{x}, \overline{\mathbf{x}}]_t[\mathbf{x}, \overline{\mathbf{x}}]_t^*)\mathbf{a} = \mathbf{a}^* M_t(L)\mathbf{a}. \tag{9}$$

More generally, if $p = \mathbf{a}^*[\mathbf{x}, \overline{\mathbf{x}}]_t$ and $q = \mathbf{b}^*[\mathbf{x}, \overline{\mathbf{x}}]_t$ with $\mathbf{a}, \mathbf{b} \in \mathbb{C}_0^{\mathbb{N}^n \times \mathbb{N}^n}$, then $L(p\overline{q}) = \mathbf{a}^* M_t(L)\mathbf{b}$. If $t = \infty$ we write $M(L)$ instead of $M_\infty(L)$.

Observe that the moment matrix of an evaluation functional L_a at $a \in \mathbb{C}^d$ satisfies $M(L_a) = [a, \overline{a}]_t[a, \overline{a}]_t^*$ 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

$$gL : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathbb{C}$$

$$p \mapsto L(gp).$$

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

$$gL \text{ is positive} \iff L(g \cdot [\mathbf{x}, \overline{\mathbf{x}}][\mathbf{x}, \overline{\mathbf{x}}]^*) = M(gL) \succeq 0. \tag{10}$$

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*.

2.2. SoS-polynomial matrices and matrix-valued linear maps

There is a natural extension of the previously defined concepts to the matrix-valued setting. This extension will be useful, in particular, to define a matrix analog of localizing moment constraints and to provide a moment approach to the hierarchy by Doherty, Parrilo and Spedalieri [16].

SoS-polynomial matrices. A polynomial matrix $S \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^{m \times m}$ is called an *SoS-polynomial matrix* if $S = UU^*$ for some polynomial matrix $U \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^{m \times k}$ and some integer $k \in \mathbb{N}$, or, equivalently, if $S \in \text{cone}\{\vec{\mathbf{p}}\vec{\mathbf{p}}^* : \vec{\mathbf{p}} = (p_1, \ldots, p_m) \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^m\}$.

Matrix-valued linear functionals. Consider a matrix-valued linear functional

$$\mathcal{L} : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathbb{C}^{m \times m}$$
$$p \mapsto \mathcal{L}(p) = \big(L_{ij}(p)\big)_{i,j \in [m]},$$

where $\mathcal{L} = (L_{ij})_{i,j=1}^m$ and each $L_{ij} \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^*$ is a scalar-valued linear functional. Then $\mathcal{L}$ is *Hermitian* if $\mathcal{L}(\overline{p}) = \mathcal{L}(p)^*$, i.e., $L_{ij}(\overline{p}) = \overline{L_{ji}(p)}$ for all $i, j \in [m]$, for all $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$. In addition $\mathcal{L}$ is said to be *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:

$$\mathcal{L}(p\overline{p}) = (L_{ij}(p\overline{p}))_{i,j=1}^m \succeq 0 \text{ for all } p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]. \tag{11}$$

In analogy to Eq. (7) it is natural to define the *moment matrix* $M(\mathcal{L})$ as

$$M(\mathcal{L}) := \mathcal{L}([\mathbf{x}, \overline{\mathbf{x}}][\mathbf{x}, \overline{\mathbf{x}}]^*) = (L_{ij}([\mathbf{x}, \overline{\mathbf{x}}][\mathbf{x}, \overline{\mathbf{x}}]^*))_{i,j=1}^m = (M(L_{ij}))_{i,j=1}^m, \tag{12}$$

which thus acts on $\mathbb{C}^m \otimes \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$. Clearly, $M(\mathcal{L})$ is a Hermitian matrix if $\mathcal{L}$ is Hermitian. Note that $M(\mathcal{L})$ can be viewed as an $m \times m$ block-matrix whose (i, j)th block is the moment matrix $M(L_{ij})$.

When $\mathcal{L}$ acts on a truncated polynomial space $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_{2t}$ its moment matrix $M_t(\mathcal{L})$, *truncated at order t*, is defined in the obvious way by

$$M_t(\mathcal{L}) := \mathcal{L}([\mathbf{x}, \overline{\mathbf{x}}]_t [\mathbf{x}, \overline{\mathbf{x}}]_t^*) = (M_t(L_{ij}))_{i,j=1}^m,$$

with $M_t(\mathcal{L}) = M(\mathcal{L})$ if $t = \infty$.

One may also define the action of $\mathcal{L}$ on a polynomial matrix $S = (S_{ij})_{i,j=1}^m \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^{m \times m}$ by

$$\langle \mathcal{L}, S \rangle := \sum_{i,j=1}^m L_{ij}(S_{ij}). \tag{13}$$

If $\mathcal{L}$ and S are both Hermitian then $\langle \mathcal{L}, S \rangle \in \mathbb{R}$. As before, given $g \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ we may define a new (localizing) matrix-valued linear map $g\mathcal{L}$ by:

$$g\mathcal{L} : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathbb{C}^{m \times m}$$
$$p \mapsto (g\mathcal{L})(p) = \mathcal{L}(gp) = \big(L_{ij}(gp)\big)_{i,j \in [m]}.$$

Positivity of $\mathcal{L}$ and its moment matrix $M(\mathcal{L})$.   The analog of Eq. (10) does not extend to the matrix-valued case: If $M(\mathcal{L})$ is positive semidefinite, then $\mathcal{L}$ is positive, but the reverse implication may not hold in general. In the next two lemmas, we present alternative characterizations for positivity of a matrix-valued map $\mathcal{L}$ and positivity of its moment matrix $M(\mathcal{L})$ that make this more apparent.

Lemma 1. *$\mathcal{L}$ is positive, i.e., Eq. (11) holds, if and only if any of the following equivalent conditions holds:*

$$v^* \mathcal{L}(p\overline{p})v = \Big(\sum_{i,j=1}^m \overline{v_i} v_j L_{ij} \Big)(p\overline{p}) = (v^* \mathcal{L} v)(p\overline{p}) \geq 0$$

$$\textit{for all } v \in \mathbb{C}^m \textit{ and } p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}], \tag{14}$$

$$M(v^* \mathcal{L} v) \succeq 0 \textit{ for all } v \in \mathbb{C}^m, \tag{15}$$

$$(v \otimes \mathbf{a})^* M(\mathcal{L}) (v \otimes \mathbf{a}) \geq 0 \textit{ for all } v \in \mathbb{C}^m \textit{ and } \mathbf{a} \in \mathbb{C}_0^{\mathbb{N}^n \times \mathbb{N}^n}. \tag{16}$$

Proof. The equivalence of Eq. (11) and Eq. (14) is clear. The equivalence of Eq. (14) and Eq. (15) follows using Eq. (8) applied to each (scalar-valued) map $v^* \mathcal{L} v$. To see the equivalence of Eq. (14) and Eq. (16), write a polynomial $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ as $p = \mathbf{a}^*[\mathbf{x}, \overline{\mathbf{x}}]$ with $\mathbf{a} = (a_{\alpha, \beta}) \in \mathbb{C}_0^{\mathbb{N}^n \times \mathbb{N}^n}$. Then, for any $v \in \mathbb{C}^m$, following Eq. (9), we have:

$$v^* \mathcal{L}(p\overline{p})v = v^*(L_{ij}(p\overline{p}))_{i,j=1}^m v = v^*(\mathbf{a}^* M(L_{ij})\mathbf{a})_{i,j=1}^m v = (v \otimes \mathbf{a})^* M(\mathcal{L}) v \otimes \mathbf{a},$$

using the definition of $M(\mathcal{L})$ from Eq. (12). $\square$

Lemma 2. $M(\mathcal{L}) \succeq 0$ *if and only if any of the following equivalent conditions holds:*

$$w^* M(\mathcal{L}) w \geq 0 \ \text{for all} \ w \in \mathbb{C}^m \otimes \mathbb{C}_0^{\mathbb{N}^n \times \mathbb{N}^n}, \tag{17}$$

$$\langle \mathcal{L}, \vec{\mathbf{p}} \vec{\mathbf{p}}^* \rangle = \sum_{i,j=1}^{m} L_{ij}(p_i \overline{p}_j) \geq 0 \ \text{for all} \ \vec{\mathbf{p}} = (p_1, \ldots, p_m) \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^m, \tag{18}$$

$$\langle \mathcal{L}, S \rangle \geq 0 \ \text{for all SoS-polynomial matrices} \ S \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^{m \times m}. \tag{19}$$

Proof. Eq. (17) is clear. To see the equivalence with Eq. (18) consider a vector $w = (w_{i,(\alpha,\beta)})_{i,(\alpha,\beta)}$ in $\mathbb{C}^m \otimes \mathbb{C}_0^{\mathbb{N}^n \times \mathbb{N}^n}$ and, for each $i \in [m]$, define the vector $\mathbf{a}_i = (w_{i,(\alpha,\beta)})_{(\alpha,\beta)} \in \mathbb{C}_0^{\mathbb{N}^n \times \mathbb{N}^n}$, the corresponding polynomial $p_i = \mathbf{a}_i^*[\mathbf{x}, \overline{\mathbf{x}}]$, and define the polynomial vector $\vec{\mathbf{p}} = (p_1, \ldots, p_m) \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^m$. Then

$$w^* M(\mathcal{L}) w = w^* (M(L_{ij}))_{i,j=1}^m w = \sum_{i,j=1}^{m} \mathbf{a}_i^* (L_{ij}([\mathbf{x}, \overline{\mathbf{x}}][\mathbf{x}, \overline{\mathbf{x}}]^*))_{i,j=1}^m \mathbf{a}_j = \sum_{i,j=1}^{m} L_{ij}(p_i \overline{p}_j),$$

implying the equivalence of Eq. (17) and Eq. (18). The equivalence with Eq. (19) follows since SoS-polynomial matrices are conic combinations of terms of the form $\vec{\mathbf{p}} \vec{\mathbf{p}}^*$. $\square$

Note that Eq. (16) is the restriction of Eq. (17), where we restrict to vectors w in tensor product form $w = v \otimes \mathbf{a}$. In addition, we recover Eq. (14) if, in Eq. (18), we restrict to polynomials $p_1, \ldots, p_m$ of the form $p_i = v_i p$ (for $i \in [m]$) for some $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ and $v = (v_1, \ldots, v_m) \in \mathbb{C}^m$. This shows again that Eq. (14) is more restrictive than Eq. (18). Summarizing, we have the following implication.

Lemma 3. *If* $M(\mathcal{L}) \succeq 0$ *then* $\mathcal{L}$ *is positive.*

Remark 4. Note that requiring positivity of the moment matrix $M(\mathcal{L})$ not only provides a stronger condition than requiring positivity of $\mathcal{L}$, but it is also a condition that is computationally easier to check. To make this concrete we consider the truncated case when $\mathcal{L}$ is restricted to the subspace $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_{2t}$. Then, the condition $M_t(\mathcal{L}) \succeq 0$ asks whether a single matrix is positive semidefinite, which can be efficiently done. On the other hand, asking whether $\mathcal{L}$ is positive on sums of squares of degree at most $2t$ amounts to checking whether $M_t(v^* \mathcal{L} v) \succeq 0$ for all $v \in \mathbb{C}^m$, i.e., positive semidefiniteness of infinitely many matrices.

Note also that Eq. (19) highlights the duality relationship which exists between $m \times m$ SoS-polynomial matrices and matrix-valued linear maps $\mathcal{L}$ with $M(\mathcal{L}) \succeq 0$.

Link to complete positivity of $\mathcal{L}$. We now point out a link to the notion of complete positivity. Given a linear map $\mathcal{L} : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathbb{C}^{m \times m}$ and an integer $k \in \mathbb{N}$ one can define a new linear map

$$I_k \otimes \mathcal{L} : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^{k \times k} \to \mathbb{C}^{k \times k} \otimes \mathbb{C}^{m \times m}$$

$$(p_{i'j'})_{i',j'=1}^{k} \mapsto (\mathcal{L}(p_{i'j'}))_{i',j'=1}^{k}.$$

Then $\mathcal{L}$ is said to be *completely positive* if $I_k \otimes \mathcal{L}$ is positive for all $k \in \mathbb{N}$. (See, e.g., [45] for a general reference about completely positive maps.)

Lemma 5. $\mathcal{L}$ *completely positive* $\implies I_m \otimes \mathcal{L}$ *positive* $\implies M(\mathcal{L}) \succeq 0$.

Proof. The first implication is obvious. Assume $I_m \otimes \mathcal{L}$ is positive, we show that $M(\mathcal{L}) \succeq 0$. In view of Eq. (18) it suffices to show that $\sum_{i,j=1}^{m} L_{ij}(p_i \overline{p}_j) \geq 0$ for all $\vec{\mathbf{p}} = (p_1, \ldots, p_m) \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^m$. As $\vec{\mathbf{p}} \vec{\mathbf{p}}^*$ is a SoS-polynomial matrix (and thus a positive element), it follows that $(I_m \otimes \mathcal{L})(\vec{\mathbf{p}} \vec{\mathbf{p}}^*) = (\mathcal{L}(p_{i'} \overline{p}_{j'}))_{i',j'=1}^{m} \succeq 0$. Consider the vector $w = (w_{ii'})_{i,i' \in [m]}$ with entries $w_{ii'} = 1$ if $i = i'$ and $w_{ii'} = 0$ otherwise. Then, $\sum_{i,j=1}^{m} L_{ij}(p_i \overline{p}_j) = w^*(\mathcal{L}(p_{i'} \overline{p}_{j'}))_{i',j'=1}^{m} w \geq 0$, as desired. $\square$

Polynomial matrix localizing maps $G \otimes L$. Given a (scalar-valued) linear map $L \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^*$ there is a natural generalization of the above notion of localizing map gL, where, instead of considering a scalar polynomial g, we consider a polynomial matrix $G = (G_{ij})_{i,j=1}^{m} \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^{m \times m}$. Then, we can define the matrix-valued linear map $\mathcal{L} := (G_{ij}L)_{i,j=1}^{m}$, that we denote by $G \otimes L$, by

$$G \otimes L : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathbb{C}^{m \times m}$$

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

Following Eq. (12) the moment matrix of $G \otimes L$ is

$$M(G \otimes L) = (G \otimes L)([\mathbf{x}, \overline{\mathbf{x}}][\mathbf{x}, \overline{\mathbf{x}}]^*) = ((G_{ij}L)([\mathbf{x}, \overline{\mathbf{x}}][\mathbf{x}, \overline{\mathbf{x}}]^*))_{i,j=1}^{m} = L(G \otimes [\mathbf{x}, \overline{\mathbf{x}}][\mathbf{x}, \overline{\mathbf{x}}]^*). \tag{20}$$

Remark 6. When $L = L_a$ is the (scalar-valued) evaluation map at a 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 at which G is positive semidefinite. This property motivates using such a positivity constraint in defining our bounds for the separable rank and the completely positive rank.

As observed above, $M(G \otimes L) \succeq 0$ implies that $G \otimes L$ is positive. Note that, by Eq. (15), $G \otimes L$ is positive if and only if $M((v^*Gv)L) \succeq 0$ for all $v \in \mathbb{C}^m$, while, by Eq. (18), $M(G \otimes L) \succeq 0$ if and only if $L(\vec{\mathbf{p}}^* G \vec{\mathbf{p}}) \geq 0$ for all $\vec{\mathbf{p}} \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^m$. In particular,

for a truncated linear map $L \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_{2t}^*$, the condition $M_t(G \otimes L) \succeq 0$ implies any of the following two equivalent conditions (the truncated analogs of (14) and (15)), which characterize positivity of $G \otimes L$ on Σ_{2t}:

$$L(v^*Gv \cdot p\overline{p}) \geq 0 \text{ for all } v \in \mathbb{C}^m \text{ and } p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_t, \qquad (21)$$

$$M_t((v^*Gv)L) \succeq 0 \text{ for all } v \in \mathbb{C}^m. \qquad (22)$$

While it is computationally easy to check whether $M_t(G \otimes L) \succeq 0$, it is not clear how to check the above conditions efficiently. For this reason, we will select the stronger moment matrix positivity condition when defining our new hierarchy of bounds for the separable rank. However, we note that the weaker positivity condition of the localizing map will be sufficient to establish convergence properties of the bounds.

2.3. 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 and for the moment approach to the DPS approximation hierarchy of the set $\mathcal{SEP}$ of separable states.

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

$$\mathscr{D}(S) := \{u \in \mathbb{C}^n \mid g(u) \geq 0 \text{ for every } g \in S\}. \qquad (23)$$

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

$$S_G := \{v^*Gv : v \in \mathbb{C}^d, \|v\| = 1\} \subseteq \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h, \qquad (24)$$

so that the set

$$\mathscr{D}(S_G) = \{u \in \mathbb{C}^n \mid G(u) \succeq 0\} \qquad (25)$$

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

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

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 $R > 0$,

$$R - \sum_{i=1}^n x_i\overline{x_i} \in \mathcal{M}(S). \qquad (26)$$

Hence a quadratic module is Archimedean if it contains an algebraic certificate of bound-edness of the associated positivity domain. The next lemma shows that, in the case when the algebraic certificate in (26) belongs to the quadratic module $\mathcal{M}(S)_2$, the linear func-tionals that are nonnegative on $\mathcal{M}(S)$ are bounded. Its proof is standard (and easy) and thus omitted.

Lemma 7. *Let $S \subseteq \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h$ be such that $R - \sum_{i=1}^n x_i \overline{x_i} \in \mathcal{M}(S)_2$ 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 $\mathcal{M}(S)_{2t}$. 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

$$\sup_{t \in \mathbb{N}} L_t(1) < \infty, \tag{27}$$

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 for approximating matrix factorization ranks. It is a complex analog of results by Putinar [47] and Tchakaloff [54]. For completeness, we will indicate in Appendix A.2 how to derive from these results the following complex analog.

Theorem 8. *Let $S \subseteq \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h$ be a set of Hermitian polynomials such that the quadratic module $\mathcal{M}(S)$ is Archimedean and consider a Hermitian linear map $L : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathbb{C}$. Assume that L is nonnegative on $\mathcal{M}(S)$. Then the following holds.*

(i) *(based on [47]) L has a representing measure μ that is supported by $\mathscr{D}(S)$, i.e., we have $L(p) = \int_{\mathscr{D}(S)} p \, d\mu$ for all $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$.*

(ii) *(based on [54]) For any integer $k \in \mathbb{N}$, there exists a linear functional $\widehat{L} : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathbb{C}$ which coincides with L on $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_k$ and has a finite atomic representing measure supported by $\mathscr{D}(S)$, i.e., we have*

$$\widehat{L}(p) = L(p) \text{ for every } p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_k, \tag{28}$$

$$\widehat{L} = \sum_{\ell=1}^K \lambda_\ell L_{v_\ell}, \tag{29}$$

for some integer $K \geq 1$, scalars $\lambda_1, \lambda_2, ..., \lambda_K > 0$ and vectors $v_1, v_2, ..., v_K \in \mathscr{D}(S)$.

We will often apply the above theorem to a linear functional $L \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^*$ that additionally satisfies the positivity condition: $(G \otimes L)(p\overline{p}) \succeq 0$ for all $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$, for some Hermitian polynomial matrix $G \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^{m \times m}$. Then, in view of Lemma 1

(combined with Eq. (21) and Eq. (22)), one may still apply Theorem 8 after replacing the set S by the set $S \cup S_G$ so that the resulting measure μ will be supported by $\mathscr{D}(S \cup S_G) \subseteq \{x : G(x) \succeq 0\}$, thus within the positivity domain of G.

Matrix-valued linear functionals and matrix-valued measures. We now mention extensions of the previous results in Theorem 8 from the scalar-valued case to the matrix-valued case, that we will use for the moment approach to the DPS hierarchy.

For the next result we use (a specification of) a result of Cimpric and Zalar [10, Theorem 5], which shows an operator-valued version of Theorem 8 (i). Since the latter is stated in the real case we indicate in Appendix A.3 how to derive from it its complex analog that we need for the implication (ii) $\Rightarrow$ (i) in Theorem 9 below. In a nutshell, this implication relies on a version of Riesz' representation theorem for positive operator valued linear maps (see, e.g., [28]) combined with a density argument (for going from polynomials to continuous functions) and Putinar's Positivstellensatz.

Theorem 9 (*based on [10]*). *Let* $S \subseteq \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h$ *be a set of Hermitian polynomials such that the quadratic module* $\mathcal{M}(S)$ *is Archimedean and let* $\mathcal{L} : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathcal{H}^m$ *be a Hermitian matrix-valued linear map. The following assertions are equivalent.*

(i) $\mathcal{L}$ *has a representing measure* μ *that is supported by* $\mathscr{D}(S)$ *and takes its values in the cone* $\mathcal{H}_+^m$ *of* $m \times m$ *Hermitian positive semidefinite matrices.*

(ii) $\mathcal{L}$ *is nonnegative on* $\mathcal{M}(S)$, *i.e.,* $\mathcal{L}(gp\overline{p}) \succeq 0$ *for all* $g \in S \cup \{1\}$ *and* $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$.

(iii) $M(g\mathcal{L}) \succeq 0$ *for all* $g \in S \cup \{1\}$.

(iv) $g\mathcal{L}$ *is completely positive for all* $g \in S \cup \{1\}$.

Proof. First we show that (i) implies (iv). Let $k \in \mathbb{N}$, let $P = (p_{i'j'})_{i',j'=1}^k \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^{k \times k}$ be a polynomial matrix such that $P(x) \succeq 0$ for all $x \in \mathbb{C}^d$, and let $g \in S \cup \{1\}$; we show that $(I_k \otimes g\mathcal{L})(P) \succeq 0$. For this note that

$$(I_k \otimes g\mathcal{L})(P) = (g\mathcal{L})(p_{i'j'})_{i',j'=1}^k = (\mathcal{L}(gp_{i'j'}))_{i',j'=1}^k = \left(\int_{\mathscr{D}(S)} gp_{i'j'} d\mu \right)_{i',j'=1}^k \succeq 0.$$

Here, the last inequality follows (for example) from Theorem 10 below, using the fact that $g(x) \geq 0$ on $\mathscr{D}(S)$, $P(x) = (p_{i'j'}(x))_{i',j'=1}^k \succeq 0$ for all x, and μ takes its values in $\mathcal{H}_+^m$. Indeed, say D is an upper bound on the degree of $g(x)P(x)$. Then, by Theorem 10 applied to $\mathcal{L}$ restricted to $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_D$, there exist an integer $K \in \mathbb{N}$, matrices $\Lambda_\ell \succeq 0$ and vectors $v_\ell \in \mathscr{D}(S)$ (for $\ell \in [K]$)) such that $(\mathcal{L}(gp_{i'j'}))_{i',j'=1}^k = \sum_{\ell=1}^K g(v_\ell)\Lambda_\ell \otimes P(v_\ell)$, which proves it is a positive semidefinite matrix. The implication (iv) $\Longrightarrow$ (iii) follows from Lemma 5 and (iii) $\Longrightarrow$ (ii) follows from Lemma 3.

Finally, for the implication (ii) $\Longrightarrow$ (i) we refer to the arguments in Appendix A.3. $\square$

What the above result shows is that, while in general the notions of complete positivity, positivity and having a positive semidefinite moment matrix are not equivalent, these properties become equivalent when considering a linear map $\mathcal{L}$ acting on an Archimedean quadratic module. We will apply these results to the case of the quadratic module of the unit sphere (with $S = \{1 - \sum_i x_i \overline{x_i}\}$) for the moment approach to the DPS hierarchy in Section 5.2.

Finally, there is also an analog of Theorem 8 (ii) for the matrix-valued case.

Theorem 10 *(Kimsey [33]). Let $S \subseteq \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h$ be a set of Hermitian polynomials and let $\mathcal{L} : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathcal{H}^m$ be a Hermitian matrix-valued linear map. Assume $\mathcal{L}$ has a representing measure supported by $\mathscr{D}(S)$ and taking values in the cone $\mathcal{H}_+^m$. Then, for any integer $k \in \mathbb{N}$, the restriction of $\mathcal{L}$ to $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_k$ has another representing measure that is finitely atomic; that is, there exists $K \in \mathbb{N}$, matrices $\Lambda_1, \ldots, \Lambda_K \in \mathcal{H}_+^m$ and vectors $v_1, \ldots, v_K \in \mathscr{D}(S)$ such that $\mathcal{L}(p) = \sum_{\ell=1}^K \Lambda_\ell p(v_\ell)$ for all polynomials $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_k$.*

3. A hierarchy of lower bounds on the separable rank

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

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

$$\rho = \sum_{\ell=1}^r a_\ell a_\ell^* \otimes b_\ell b_\ell^*. \tag{30}$$

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

$$\rho - a_\ell a_\ell^* \otimes b_\ell b_\ell^* \succeq 0 \quad \text{for all } \ell \in [r]. \tag{31}$$

Let

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

denote the maximum diagonal entry of ρ. Then, in view of (31), the vectors a_ℓ, b_ℓ also satisfy $|(a_\ell)_i|^2 |(b_\ell)_j|^2 \leq \rho_{ij,ij}$ for all $i, j \in [d]$, which implies the following boundedness conditions

$$\|a_\ell\|_\infty^2 \cdot \|b_\ell\|_\infty^2 \leq \rho_{\max} \quad \text{and} \quad \|a_\ell\|_2^2 \cdot \|b_\ell\|_2^2 \leq \mathrm{Tr}(\rho) \quad \text{for all } \ell \in [r]. \tag{32}$$

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

$$\|a_\ell\|_\infty^2, \|b_\ell\|_\infty^2 \leq \sqrt{\rho_{\max}} \quad \text{for all } \ell \in [r]. \tag{33}$$

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

$$\|a_\ell\|_2^2 = \|b_\ell\|_2^2 \leq \sqrt{\mathrm{Tr}(\rho)} \quad \text{for all } \ell \in [r]. \tag{34}$$

Yet another possibility would be to rescale so that $\|b_\ell\|_2 = \sqrt{\mathrm{Tr}(\rho)}$ for all ℓ, in which case we would have

$$\|a_\ell\|_2^2 \leq \sqrt{\mathrm{Tr}(\rho)}, \ \|b_\ell\|_2 = \sqrt{\mathrm{Tr}(\rho)} \ \text{ for all } \ell \in [r] \tag{35}$$

or, equivalently (up to rescaling), we may assume that

$$\|a_\ell\|_2^2 \leq \mathrm{Tr}(\rho), \ \|b_\ell\|_2 = 1 \quad \text{for all } \ell \in [r]. \tag{36}$$

To fix ideas we will now apply the first rescaling (33), so that each (a_ℓ, b_ℓ) belongs to the set

$$\mathcal{V}_\rho := \left\{ (x, y) \in \mathbb{C}^d \times \mathbb{C}^d \mid xx^* \otimes yy^* \preceq \rho, \ \|x\|_\infty, \|y\|_\infty \leq \rho_{\max}^{1/4} \right\}. \tag{37}$$

We will consider the impact of doing other rescalings as in Eq. (34), Eq. (35) or Eq. (36) later on in the paper in numerical examples. However, as will be noted in Remark 12, the localizing constraints corresponding to the scaling (33) already imply the localizing constraints corresponding to the inequalities in (32).

From Eq. (30) 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

$$\tau_{\mathrm{sep}}(\rho) := \inf\left\{ \lambda : \lambda > 0, \frac{1}{\lambda}\rho \in \mathrm{conv}\{xx^* \otimes yy^* : (x, y) \in \mathcal{V}_\rho\} \right\}. \tag{38}$$

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

Lemma 11. *For any $\rho \in \mathcal{SEP}_d$, we have $\tau_{\mathrm{sep}}(\rho) \leq \mathrm{rank}_{\mathrm{sep}}(\rho)$. Moreover, if $\rho \notin \mathcal{SEP}_d$ then $\tau_{\mathrm{sep}}(\rho) = \mathrm{rank}_{\mathrm{sep}}(\rho) = \infty$.*

The parameter $\tau_{\mathrm{sep}}(\rho)$ 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. These bounds arise from convex (semidefinite) programs, they in fact also lower bound the (weaker) parameter $\tau_{\mathrm{sep}}(\rho)$ and will be shown to asymptotically converge to it.

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 (30) 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 (37). We explain how to define bounds for $\mathrm{rank}_{\mathrm{sep}}(\rho)$ by using the moment method from Section 2.3.

For this let us consider the linear functional

$$L = \sum_{\ell=1}^{r} L_{(a_\ell, b_\ell)}, \tag{39}$$

the sum of the evaluation functionals at the points entering the decomposition (30). 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. (30). 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 Hermitian polynomial matrix

$$G_\rho(\mathbf{x}, \mathbf{y}) := \rho - \mathbf{x}\mathbf{x}^* \otimes \mathbf{y}\mathbf{y}^* \in \mathbb{C}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_4^{d^2 \times d^2} \tag{40}$$

and the localizing set of Hermitian polynomials

$$S_\rho = \left\{ \sqrt{\rho_{\max}} - x_i \overline{x}_i, \sqrt{\rho_{\max}} - y_i \overline{y}_i : i \in [d] \right\} \subseteq \mathbb{C}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_2^h. \tag{41}$$

To see that $M(G_\rho \otimes L) \succeq 0$ we use Remark 6. Recall also the definition of the set S_{G_ρ} of localizing polynomials corresponding to the polynomial matrix G_ρ in (40):

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

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

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

Remark 12. Note that the localizing constraints for the inequalities in (32) are implied by the localizing constraints for $S_\rho \cup S_{G_\rho}$. This follows from the following two identities:

$$\rho_{\max} - x_i \overline{x}_i y_j \overline{y}_j = (\sqrt{\rho_{\max}} - x_i \overline{x}_i) y_j \overline{y}_j + \sqrt{\rho_{\max}}(\sqrt{\rho_{\max}} - y_j \overline{y}_j) \in \mathcal{M}(S_\rho)_4,$$
$$\mathrm{Tr}(\rho) - (\sum_i x_i \overline{x}_i)(\sum_j y_j \overline{y}_j) = \sum_{i,j}(\rho_{ij,ij} - x_i \overline{x}_i y_j \overline{y}_j) \in \mathcal{M}(S_{G_\rho})_4.$$

Moreover, let us recall for future reference that

$$M(G_\rho \otimes L) \succeq 0 \implies L \ge 0 \text{ on } \mathcal{M}(S_{G_\rho}), \tag{42}$$

which follows from Lemma 1 and the characterization of positivity of $G_\rho \otimes L$ from Eq. (22). The above observations motivate introducing the following parameters. For $t \in \mathbb{N} \cup \{\infty\}$ with $t \ge 2$, define the parameter

$$\begin{aligned}
\xi_t^{\mathrm{sep}}(\rho) := \inf \Big\{ L(1) \mid\ & L : \mathbb{C}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_{2t} \to \mathbb{C} \ \ \text{Hermitian} \ \ \text{s.t.} \\
& L(\mathbf{x}\mathbf{x}^* \otimes \mathbf{y}\mathbf{y}^*) = \rho, \\
& L \ge 0 \text{ on } \mathcal{M}(S_\rho)_{2t}, \\
& M_{t-2}(G_\rho \otimes L) = L(G_\rho \otimes [\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_{t-2}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]_{t-2}^*) \succeq 0 \Big\}.
\end{aligned} \tag{43}$$

For $t = \infty$ the parameter $\xi_\infty^{\mathrm{sep}}(\rho)$ 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 $\xi_*^{\mathrm{sep}}(\rho)$ denote the parameter obtained by adding the constraint $\mathrm{rank}(M(L)) < \infty$ to the definition of $\xi_\infty^{\mathrm{sep}}(\rho)$. One can show that the function $\rho \mapsto \xi_t^{\mathrm{sep}}(\rho)$ is lower semicontinuous, the proof is analogous to that of [26, Lemma 7] and thus omitted. In addition, as we will see in Remark 30, if the program defining $\xi_t^{\mathrm{sep}}(\rho)$ ($t \ge 2$) is feasible then ρ satisfies the PPT criterion, i.e., $\rho^{T_B} \succeq 0$, where ρ^{T_B} is obtained by taking the partial transpose of ρ on the second register (see (70)).

As is well-known, for finite $t \in \mathbb{N}$, the bound $\xi_t^{\mathrm{sep}}(\rho)$ can be expressed as a semidefinite program since nonnegativity of L on the truncated quadratic module $\mathcal{M}(S_\rho)_{2t}$ 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 of 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:

$$\xi_2^{\mathrm{sep}}(\rho) \le \xi_3^{\mathrm{sep}}(\rho) \le \cdots \le \xi_\infty^{\mathrm{sep}}(\rho) \le \xi_*^{\mathrm{sep}}(\rho) \le \mathrm{rank}_{\mathrm{sep}}(\rho) < \infty. \tag{44}$$

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

Lemma 13. *Let $\rho \in \mathcal{SEP}_d$. The infimum is attained in problem* (43) *for any integer $t \geq 2$ or $t = \infty$, and we have $\lim_{t\to\infty} \xi_t^{\mathrm{sep}}(\rho) = \xi_\infty^{\mathrm{sep}}(\rho)$.*

Proof. First we show that problem (43) attains its optimum. For this note that, in view of Eq. (44), we may restrict the optimization to linear functionals L satisfying $L(1) \leq \mathrm{rank}_{\mathrm{sep}}(\rho)$. By the definition of S_ρ in (41), the quadratic module $\mathcal{M}(S_\rho)$ is Archimedean since, with $R = 2d\sqrt{\rho_{\max}}$, $R - \sum_{i=1}^{d}(x_i\overline{x}_i + y_i\overline{y}_i) \in \mathcal{M}(S_\rho)_2$. As L is nonnegative on $\mathcal{M}(S_\rho)_{2t}$, we can apply Lemma 7 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 (43). As $\sup_t L_t(1) \leq \mathrm{rank}_{\mathrm{sep}}(\rho) < \infty$, we can conclude from Lemma 7 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 $\xi_\infty^{\mathrm{sep}}(\rho)$, which implies $\xi_\infty^{\mathrm{sep}}(\rho) \leq L(1) = \lim_{t\to\infty} L_t(1) = \lim_t \xi_t^{\mathrm{sep}}(\rho)$. Note that this L is optimal for $\xi_\infty^{\mathrm{sep}}(\rho)$. $\quad\square$

Lemma 14. *For any $\rho \in \mathcal{H}^d \otimes \mathcal{H}^d$ we have $\xi_\infty^{\mathrm{sep}}(\rho) = \xi_*^{\mathrm{sep}}(\rho) = \tau_{\mathrm{sep}}(\rho)$.*

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

First we show $\xi_*^{\mathrm{sep}}(\rho) \leq \tau_{\mathrm{sep}}(\rho)$. If $\tau_{\mathrm{sep}}(\rho) = \infty$ 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 $\lambda > 0$, $(a_\ell, b_\ell) \in \mathcal{V}_\rho$, $\mu_\ell > 0$ 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 $\xi_*^{\mathrm{sep}}(\rho)$ with $L(1) = \lambda$. Hence, $\xi_*^{\mathrm{sep}}(\rho) \leq L(1) = \lambda$, which shows $\xi_*^{\mathrm{sep}}(\rho) \leq \tau_{\mathrm{sep}}(\rho)$.

Now we show $\tau_{\mathrm{sep}}(\rho) \leq \xi_\infty^{\mathrm{sep}}(\rho)$. If $\xi_\infty^{\mathrm{sep}}(\rho) = \infty$ there is nothing to prove. So assume L is a feasible solution to $\xi_\infty^{\mathrm{sep}}(\rho)$. Then, in view of Eq. (42), $L \geq 0$ on $\mathcal{M}(S_\rho \cup S_{G_\rho})$. As $\mathcal{M}(S_\rho)$ is Archimedean we can apply Theorem 8 (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 $\mathscr{D}(S_\rho \cup S_{G_\rho}) = \mathcal{V}_\rho$. 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 $\tau_{\mathrm{sep}}(\rho) \leq L(1)$, showing $\tau_{\mathrm{sep}}(\rho) \leq \xi_\infty^{\mathrm{sep}}(\rho)$. $\quad\square$

As observed earlier already, since $\mathcal{SEP}_d$ is a d^4-dimensional cone, by Carathéodory theorem we have $\mathrm{rank}_{\mathrm{sep}}(\rho) \leq d^4$ for any $\rho \in \mathcal{SEP}_d$ (or, even stronger, $\mathrm{rank}_{\mathrm{sep}}(\rho) \leq \mathrm{rank}(\rho)^2$). Based on this one can also use the bounds $\xi_t^{\mathrm{sep}}(\rho)$ to test (non-)membership in $\mathcal{SEP}_d$. The bound $\mathrm{rank}(\rho)^2$ in Lemma 15 below can of course be replaced by any

other valid upper bound on the separable rank. Such a valid bound can be obtained, e.g., using the *birank* of ρ, defined as the pair $(\mathrm{rank}(\rho), \mathrm{rank}(\rho^{T_B}))$. Indeed, as $\mathrm{rank}_{\mathrm{sep}}(\rho) = \mathrm{rank}_{\mathrm{sep}}(\rho^{T_B})$, we have

$$\max\{\mathrm{rank}(\rho), \mathrm{rank}(\rho^{T_B})\} \leq \mathrm{rank}_{\mathrm{sep}}(\rho) \leq (\min\{\mathrm{rank}(\rho), \mathrm{rank}(\rho^{T_B})\})^2. \qquad (45)$$

Lemma 15. *Let $\rho \in \mathcal{H}^d \otimes \mathcal{H}^d$. Then, $\rho \in \mathcal{SEP}_d$ if and only if $\xi_t^{\mathrm{sep}}(\rho) \leq \mathrm{rank}(\rho)^2$ for all integers $t \geq 2$.*

Proof. The 'only if' part follows from $\xi_t^{\mathrm{sep}}(\rho) \leq \mathrm{rank}_{\mathrm{sep}}(\rho) \leq \mathrm{rank}(\rho)^2$ when $\rho \in \mathcal{SEP}_d$. Conversely, assume $\xi_t^{\mathrm{sep}}(\rho) \leq \mathrm{rank}(\rho)^2$ for all integers $t \geq 2$. Then, we can use the same argument as in the proof of Lemma 13 and conclude the existence of $L \in \mathbb{C}[\mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}}]^*$ feasible for $\xi_\infty^{\mathrm{sep}}(\rho)$, so that $\xi_\infty^{\mathrm{sep}}(\rho) \leq L(1) < \infty$. Then, by Lemma 14, we have $\tau_{\mathrm{sep}}(\rho) < \infty$, which shows ρ is separable. $\square$

Remark 16. Note that all the results in this section remain valid if, in the definition (43) of the parameter $\xi_t^{\mathrm{sep}}(\rho)$, we omit the 'tensor-type' constraint $M_{t-2}(G_\rho \otimes L) \succeq 0$. Using this additional constraint permits however to define stronger bounds on the separable rank. The results also remain valid if, instead of the polynomials in the set S_ρ, we use either of the following sets of polynomials: $\{\pm(\|x\|^2 - \|y\|^2), \sqrt{\mathrm{Tr}(\rho)} - \|y\|^2\}$ corresponding to (34), or $\{\sqrt{\mathrm{Tr}(\rho)} - \|x\|^2, \pm 1(\sqrt{\mathrm{Tr}(\rho)} - \|y\|^2)\}$ corresponding to (35) (or, equivalently, $\{\mathrm{Tr}(\rho) - \|x\|^2, \pm(1 - \|y\|^2)\}$ corresponding to (36)).

3.3. Block-diagonal reduction for the parameter $\xi_t^{\mathrm{sep}}(\cdot)$

In this section we indicate how to rewrite the program (43) defining $\xi_t^{\mathrm{sep}}(\rho)$ in a more economical way. Observe that all the terms of each of the localizing polynomials $g \in S_\rho$ and the matrix G_ρ have the same degree in $\mathbf{x}$ and in $\overline{\mathbf{x}}$, and also the same degree in $\mathbf{y}$ and in $\overline{\mathbf{y}}$. This enables us to show (see Lemma 17) that we may restrict the optimization in (43) to linear functionals L that satisfy the condition

$$L(\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'}) = 0 \quad \text{if } |\alpha| \neq |\alpha'| \text{ or } |\beta| \neq |\beta'|. \qquad (46)$$

Note that this implies in particular that $L(\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'}) = 0$ if $|\alpha + \alpha'|$ or $|\beta + \beta'|$ is odd.

The computational advantage is that, if L satisfies (46), then the moment matrix $M_t(L)$ and the localizing moment matrices $M_{t-1}(gL)$ and $M_{t-2}(G_\rho \otimes L)$ have a block-diagonal form. To see this consider first the matrix $M_t(L)$, which is indexed by the set

$$I^t := \{(\alpha, \alpha', \beta, \beta') \in (\mathbb{N}^d)^4 : |\alpha + \beta + \alpha' + \beta'| \leq t\} \qquad (47)$$

(where the tuple $(\alpha, \alpha', \beta, \beta')$ corresponds to the monomial $\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'}$). Let us partition I^t into sets depending on two integers $r = |\alpha| - |\alpha'|$ and $s = |\beta| - |\beta'|$. For $r, s \in \{-t, -t+1, \ldots, t\}$ let

$$I_{r,s}^t := \{(\alpha, \alpha', \beta, \beta') \in I^t : |\alpha| - |\alpha'| = r, \ |\beta| - |\beta'| = s\}, \tag{48}$$

then we have

$$I^t = \bigcup_{r,s=-t}^{t} I_{r,s}^t.$$

Then, with respect to this partition of its index set, the matrix $M_t(L)$ is block-diagonal and thus $M_t(L) \succeq 0$ if and only if its principal submatrices indexed by the sets $I_{r,s}^t$ are positive semidefinite. The analogous reasoning applies to each localizing moment matrix $M_{t-1}(gL)$ for $g \in S_\rho$ (indexed by I^{t-1}) and to $M_{t-2}(G_\rho \otimes L)$ (indexed by I^{t-2}).

Lemma 17. *In the definition of the parameter $\xi_t^{\mathrm{sep}}(\rho)$ we may restrict the optimization to linear functionals satisfying the additional condition (46).*

Proof. Assume L is feasible for $\xi_t^{\mathrm{sep}}(\rho)$; we construct another feasible solution $\tilde{L}$ with the same objective value: $\tilde{L}(1) = L(1)$, and satisfying (46). For this define $\tilde{L}(\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'}) = L(\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'})$ if $|\alpha| = |\alpha'|$ and $|\beta| = |\beta'|$, and $\tilde{L}(\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'}) = 0$ otherwise. Then, $\tilde{L}(1) = L(1)$ and, by construction, $\tilde{L}$ satisfies (46). We claim that $\tilde{L}$ is feasible for program (43). Clearly, we have $\tilde{L}(\mathbf{x}\mathbf{x}^* \otimes \mathbf{y}\mathbf{y}^*) = \rho$. We now show that $M_t(\tilde{L}) \succeq 0$, $M_{t-1}(g\tilde{L}) \succeq 0$ for $g \in S_\rho$, and $M_{t-2}(G_\rho \otimes \tilde{L}) \succeq 0$.

We first show that $M_t(\tilde{L}) \succeq 0$. We use the partitioning $I^t = \cup_{r,s=-t}^t I_{r,s}^t$ of the row/column indices.

As the principal submatrix of $M_t(\tilde{L})$ indexed by $I_{r,s}^t$ only involves evaluations of L at monomials of the form $\mathbf{x}^\gamma \overline{\mathbf{x}}^{\gamma'} \mathbf{y}^\delta \overline{\mathbf{y}}^{\delta'}$ with $|\gamma| = |\gamma'|$ and $|\delta| = |\delta'|$, it coincides with the principal submatrix of $M_t(L)$ indexed by $I_{r,s}^t$ and thus it is positive semidefinite. Hence, by construction, the matrix $M_t(\tilde{L})$ is block-diagonal with respect to the partition $I_t = \cup_{r,s=0}^t I_{r,s}$ of its index set, with positive semidefinite diagonal blocks, which implies $M_t(\tilde{L}) \succeq 0$.

Consider now a localizing polynomial $g \in S_\rho$. Note that all its terms have the same degree in $\mathbf{x}$ and $\overline{\mathbf{x}}$ and also the same degree in $\mathbf{y}$ and $\overline{\mathbf{y}}$ (equal to 0 or 1). We consider the partition of the index set of $M_{t-1}(gL)$ as $I^{t-1} = \cup_{r,s=-t+1}^{t-1} I_{r,s}^{t-1}$. Again, the principal submatrix of $M_{t-1}(gL)$ indexed by $I_{r,s}^{t-1}$ involves only values $\tilde{L}(\mathbf{x}^\gamma \overline{\mathbf{x}}^{\gamma'} \mathbf{y}^\delta \overline{\mathbf{y}}^{\delta'})$ with $|\gamma| = |\gamma'|$ and $|\delta| = |\delta'|$ and thus it coincides with the principal submatrix of $M_{t-1}(gL)$ indexed by $I_{r,s}^{t-1}$. Hence, the matrix $M_{t-1}(g\tilde{L})$ is block-diagonal with respect to the partition $I_{t-1} = \cup I_{r,s}^{t-1}$ of its index set, with positive semidefinite diagonal blocks, which implies $M_{t-1}(g\tilde{L}) \succeq 0$.

The analogous reasoning applies to showing that $M_{t-2}(G_\rho \otimes \tilde{L}) \succeq 0$. For this we consider the partition of its index set $[d]^2 \times I_{t-2}$ into $\cup_{r,s=-t+2}^{t-2}([d]^2 \times I_{r,s}^{t-2})$ and observe that $M_{t-2}(G_\rho \otimes \tilde{L})$ is block-diagonal with respect to this partition, with positive semidefinite diagonal blocks. $\square$

An alternative way to arrive at Eq. (46) by exploiting sign symmetries. Let $\mathbb{T}$ be the *circle group*, the multiplicative group of all complex numbers of modulus 1:

$$\mathbb{T} = \{z \in \mathbb{C} \colon |z| = 1\}.$$

The set $\mathcal{SEP}_d$ is naturally invariant under the action of $(w_x, w_y) \in \mathbb{T} \times \mathbb{T}$ on vectors $(x, y) \in \mathbb{C}^d \times \mathbb{C}^d$ given by $(w_x, w_y)(x, y) = (w_x x, w_y y)$ (and its extension to states). Indeed, we have

$$(w_x, w_y) \cdot (xx^* \otimes yy^*) = (w_x x)(w_x x)^* \otimes (w_y y)(w_y y)^* = xx^* \otimes yy^*.$$

Likewise, the localizing constraints are invariant under this group action, and this group action extends to the linear functionals L used as variables in the definition of $\xi_t^{\mathrm{sep}}(\rho)$. Since $\mathbb{T} \times \mathbb{T}$ admits a Haar measure, in the derivation of $\xi_t^{\mathrm{sep}}(\rho)$ we may therefore restrict to linear functionals that are invariant under this group action. That is, we may assume that

$$L(\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'}) = w_x^{|\alpha|-|\alpha'|} w_y^{|\beta|-|\beta'|} L(\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'}) \quad \text{for all } (w_x, w_y) \in \mathbb{T} \times \mathbb{T}.$$

This implies that

$$L(\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'}) = 0 \text{ if } |\alpha| \neq |\alpha'| \text{ or } |\beta| \neq |\beta'|.$$

Indeed, suppose for example that $|\alpha| - |\alpha'| =: r \neq 0$. Then using the above with $w_x = e^{\mathbf{i}\,\pi/r}$ shows that $L(\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'}) = -L(\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'})$ and hence $L(\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'}) = 0$.

Note that Dressler, Nie and Yang [17] used this same group action to argue that, alternatively, one may restrict to $(x, y) \in \mathbb{C}^d \times \mathbb{C}^d$ having leading coordinates that are real nonnegative: $x_1, y_1 \geq 0$. While this permits to eliminate variables (and work with $2(2d-1)$ instead of $4d$ real variables), this reduction does not permit to block-diagonalize the moment matrices as indicated above. We also refer to [23] and the recent paper [57] for more details about exploiting sign symmetries.

Block-diagonal reduction example. To illustrate the effect of the block-diagonalization we consider an example with $\rho \in \mathcal{H}^3 \otimes \mathcal{H}^3 \simeq \mathcal{H}^9$ (i.e., $d_1 = d_2 = 3$) and relaxation order $t = 3$. In Table 1 we indicate the respective sizes of the matrices involved in the program for $\xi_3^{\mathrm{sep}}(\rho)$ with and without block-diagonalization (in column 'block' and 'non-block', respectively). There, '# entries' stands for $\sum_i m_i^2$, where m_i are the sizes of the matrices involved in the program, and '# variables' indicates the total number of variables in each case. The last line indicates the typical run time for such an instance, we collect the computational details later in Section 4.2. Note that the full program cannot be solved and thus block-diagonalization is crucial to enable computation. For the next case $(d_1, d_2) = (2, 6)$ or $(4, 4)$ one can compute the bound of order $t = 2$ but not the bound of order $t = 3$ even after block-diagonalization.

Table 1

Matrix sizes block- vs. non-block-diagonalized.

Matrix	Block	Non-block
$M_3(L)$	$25 \times (12 \times 12$ to $96 \times 96)$	455×455
$M_2(gL)$	$78 \times (6 \times 6$ to $38 \times 38)$	$6 \times (91 \times 91)$
$M_1(G_\rho \otimes L)$	$5 \times (36 \times 36$ to $108 \times 108)$	234×234
# entries	110480	286624
# variables	6952	18564
run time	4.6 min	memory error

Remark 18. As observed above, using the block-diagonalized version of the program for $\xi_3^{\mathrm{sep}}(\rho)$ is crucial to be able to compute the bounds for some larger matrix sizes. We note however that the optimal solution to this program will not satisfy the flatness condition rank $M_t(L) = $ rank $M_{t-1}(L)$ (with $t = 2, 3$). Indeed one can check that this flatness condition can hold only in the trivial case $\rho = 0$. Intuitively this can be (roughly) explained by noting that, due to its symmetric structure, L tends to lie within the interior of the feasible region. Hence our approach, which produces lower bounds on $\mathrm{rank}_{\mathrm{sep}}(\rho)$, can be viewed as being complementary to the approach in, e.g., [17,38,43], which uses flatness to produce separable decompositions of ρ and thus upper bounds on $\mathrm{rank}_{\mathrm{sep}}(\rho)$.

4. Extensions and connections to other matrix factorization ranks

Here we explain some simple extensions of the approach given in the previous section to related notions of factorization ranks.

Without going into details let us mention that the approach generalizes in a straight-forward way to the separable rank of *multipartite* separable quantum states. In that case we have an n-partite quantum state ρ acting on $(\mathbb{C}^d)^{\otimes n}$, and separability means that ρ belongs to the set

$$\mathrm{cone}\{x_1 x_1^* \otimes x_2 x_2^* \otimes \ldots \otimes x_n x_n^* : x_1, \ldots, x_n \in \mathbb{C}^d, \|x_i\| = 1 \ (i \in [n])\}.$$

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

The approach also extends to an alternative (but equivalent) definition of separability, which uses mixed states instead of pure states, i.e., where one requires ρ to be of the form $\rho = \sum_{\ell=1}^r A_\ell \otimes B_\ell$ with $A_\ell, B_\ell \in \mathcal{H}_+^d$. Analogously, the smallest such integer r is called the *mixed separable rank* of ρ. This notion has been considered, e.g., in [12,13,17] and mixed separable decompositions are called S-decompositions in [43] (which deals with real states). To define bounds on the mixed separable rank one can follow the same approach as in Section 3 but one here has to introduce more variables. Indeed, we now need variables $\mathbf{x} = (x_{ij})_{1 \leq i \leq j \leq d}$ and $\mathbf{y} = (y_{ij})_{1 \leq i \leq j \leq d}$ to model the entries of the matrices $A_\ell \in \mathcal{H}_+^d$ and $B_\ell \in \mathcal{H}_+^d$ (while we previously only needed variables $(x_i)_{i \in [d]}$ and $(y_i)_{i \in [d]}$ to model the vectors $a_\ell \in \mathbb{C}^d$ and $b_\ell \in \mathbb{C}^d$) and one should assume that the corresponding Hermitian matrices $X = (x_{ij})$ and $Y = (y_{ij})$ are positive semidefinite. One may again scale the variables so that they satisfy a boundedness condition $|x_{ij}|, |y_{ij}| \leq \sqrt{\rho_{\max}}$.

This enables to design hierarchies of lower bounds that converge to the mixed separable analog of the parameter $\tau_{\text{sep}}(\rho)$. The details are analogous and thus omitted.

In what follows we mention two other possible extensions, for the real separable rank and for the completely positive rank, where we give some more details as well as some numerical results.

4.1. Specialization to bipartite real states

The treatment in Section 3 for the separable rank can be adapted in an obvious manner to the case of the *real* separable rank. Here we are given a real symmetric bipartite state $\rho \in \mathcal{S}^d \otimes \mathcal{S}^d$, where $\mathcal{S}^d$ is the set of real symmetric $d \times d$ matrices. Then ρ is called *real separable* if it admits a decomposition (30) with all vectors $a_\ell, b_\ell \in \mathbb{R}^d$ real valued, and the smallest r for which such a decomposition exists is the *real separable rank*, denoted $\text{rank}_{\text{sep}}^{\mathbb{R}}(\rho)$. Note that it can be that a real state is separable but not real separable; this is the case for the state Sep3 discussed in Section 4.2. One can define in an analogous manner the corresponding parameter $\tau_{sep}^{\mathbb{R}}(\rho)$ and the hierarchy of bounds $\xi_t^{\text{sep},\mathbb{R}}(\rho)$ that converge asymptotically to $\tau_{sep}^{\mathbb{R}}(\rho)$. The difference in the formulation of these parameters is that we now replace the complex conjugate by the real transpose operation and work with linear functionals L acting on the real polynomial space $\mathbb{R}[\mathbf{x}, \mathbf{y}]_{2t}$. So the parameter $\xi_t^{\text{sep},\mathbb{R}}(\rho)$ reads

$$
\begin{aligned}
\xi_t^{\text{sep},\mathbb{R}}(\rho) := \inf \Big\{ L(1) \mid & L : \mathbb{R}[\mathbf{x}, \mathbf{y}]_{2t} \to \mathbb{R} \text{ s.t.} \\
& L(\mathbf{x}\mathbf{x}^T \otimes \mathbf{y}\mathbf{y}^T) = \rho, \\
& L \geq 0 \text{ on } \mathcal{M}(S_\rho)_{2t}, \\
& M_{t-2}(G_\rho \otimes L) = L(G_\rho \otimes [\mathbf{x}, \mathbf{y}]_{t-2}[\mathbf{x}, \mathbf{y}]_{t-2}^T) \succeq 0 \Big\}.
\end{aligned}
\tag{49}
$$

Again we may impose an additional block-diagonal structure on the positive semidefinite matrices entering this program. Indeed, since the polynomials involved in the constraints leading to the above program have the property that all their terms have an even degree in $\mathbf{x}$ and an even degree in $\mathbf{y}$, we may assume that the variable L satisfies the condition

$$
L(\mathbf{x}^\alpha \mathbf{y}^\beta) = 0 \text{ if } |\alpha| \text{ or } |\beta| \text{ is odd.}
\tag{50}
$$

Note that this is the real analog of condition (46) in the complex case. The additional constraint (50) permits to replace each of the positive semidefinite constraints for the matrices $M_t(L)$, $M_{t-1}(gL)$ for $g \in S_\rho$, and $M_{t-2}(G_\rho \otimes L)$ by four smaller positive semidefinite constraints, each of size roughly $1/4$ of the original size. For this let I^t denote the index set of the matrix $M_t(L)$ which we partition into $I^t = \cup_{a,b \in \{0,1\}} I_{a,b}^t$, where $I_{a,b}^t$ consists of the pairs $(\alpha, \beta) \in I^t$ with given parity $|\alpha| \equiv a$, $|\beta| \equiv b$ modulo 2. Then, with respect to this partition of its index set, the matrix $M_t(L)$ is block-diagonal and thus $M_t(L) \succeq 0$ if and only if $M_t(L)[I_{a,b}^t] \succeq 0$ for $a, b \in \{0, 1\}$. The same block-diagonalization

Table 2
Examples and numerical bounds level $t = 2$.

ρ	(d_1, d_2)	birank(ρ)	$\xi_2^{\text{sep}}(\rho)$			$\xi_2^{\text{sep},\mathbb{R}}(\rho)$			rank$_{\text{sep}}(\rho)$	time
			S1	S2	S3	S1	S2	S3		
Sep1 [7]	(2, 2)	(2, 2)	**2.0**	1.0	1.0	**2.0**	1.0	1.0	2	< 1
Sep2 [7]	(2, 2)	(3, 3)	1.421	1.0	1.0	1.421	1.0	1.0	3	< 1
Sep3 [7]	(2, 3)	(4, 6)	1.333	1.0	1.0	*	*	*	6	< 1
Sep4 [17]	(3,3)	(2,2)	1.0	1.0	1.0	1.0	**1.953**	1.0	2	< 1
Sep5 [17]	(2,2)	(4,4)	1.069	1.0	1.0	N/A	N/A	N/A	≤ 7	< 1
Sep6 [17]	(3,3)	(7,7)	1.053	1.0	1.0	N/A	N/A	N/A	≤ 9	< 1
Ent1 [9]	(3, 3)	(4, 4)	2.069	*	1.525	2.069	*	1.525	∞	< 1
Ent2 [17]	(2,2)	(2,4)	*	*	*	N/A	N/A	N/A	∞	< 1

Table 3
Examples and numerical bounds level $t = 3$.

ρ	(d_1, d_2)	birank(ρ)	$\xi_3^{\text{sep}}(\rho)$			$\xi_3^{\text{sep},\mathbb{R}}(\rho)$			rank$_{\text{sep}}(\rho)$	time
			S1	S2	S3	S1	S2	S3		
Sep1	(2, 2)	(2, 2)	**2.0**	**2.0**	**2.0**	**2.0**	**2.0**	**2.0**	2	< 1
Sep2	(2, 2)	(3, 3)	1.909	2.0	**2.178**	1.909	2.0	**2.178**	3	2
Sep3	(2, 3)	(4, 6)	2.423	3.0	2.790	*	*	*	6	25
Sep4	(3,3)	(2,2)	**1.652**	**2.0**	-	**1.65**	**2.0**	1.0	2	261
Sep5	(2,2)	(4,4)	1.988	2.048	2.079	N/A	N/A	N/A	≤ 7	4
Sep6	(3,3)	(7,7)	2.715	3.326	-	N/A	N/A	N/A	≤ 9	290
Ent1	(3, 3)	(4, 4)	-	-	*	-	*	*	∞	67

Table 4
Examples and numerical bounds level $t = 4$.

ρ	(d_1, d_2)	birank(ρ)	$\xi_4^{\text{sep}}(\rho)$			$\xi_4^{\text{sep},\mathbb{R}}(\rho)$			rank$_{\text{sep}}(\rho)$	time
			S1	S2	S3	S1	S2	S3		
Sep1	(2, 2)	(2, 2)	**2.0**	**2.0**	**2.0**	**2.0**	**2.0**	**2.0**	2	105
Sep2	(2, 2)	(3, 3)	**3.0**	**3.0**	**3.0**	**3.0**	**3.0**	**3.0**	3	332
Sep5	(2,2)	(4,4)	4.0	4.0	4.0	N/A	N/A	N/A	≤ 7	161

Run time given in seconds
*: Infeasibility certificate returned
-: Solver could not reach a conclusion (not a memory error)
N/A: Not Applicable
We indicate using boldface when a bound (after rounding up) equals the separable rank.

applies to the matrices $M_{t-1}(gL)$ for $g \in S_\rho$. For the matrix $M_{t-2}(G_\rho \otimes L)$ we consider the block-diagonalization obtained by partitioning its index set as $\cup_{a,b \in \{0,1\}}([d^2] \times I_{a,b}^{t-2})$.

Some numerical results on the behavior of the bounds will be given in the next section.

4.2. Numerical results for bipartite complex and real states

Here we collect some numerical results that illustrate the behavior of the bounds $\xi_t^{\text{sep}}(\rho)$ and $\xi_t^{\text{sep},\mathbb{R}}(\rho)$ for different choices of localizing constraints, see Tables 2 to 4 for examples at order $t = 2, 3, 4$ respectively. Computations were made in Windows using Julia [3], JuMP [18] and MOSEK [15] with hardware specifications: i7-8750 CPU with 32 Gb Memory.[3] For our examples we will use the separable states Sep1, Sep2, and

---

[3] The code is available at: https://github.com/JAndriesJ/sep-rank.

Sep3, and the entangled state Ent1 that we describe now. For numerical stability we do the computations with a scaling of these states so that they have trace equal to 1. We present the examples in matrix form with lines drawn to indicate the block structure $\rho = \big((\rho_{ij,i'j'})_{j,j'\in[d_2]}\big)_{i,i'\in[d_1]}$. Zero-valued entries are left blank.

$$
\mathrm{Sep1} := \left[\begin{array}{cc|cc}
1 & & & \\
 & & & \\\hline
 & & & \\
 & & & 1
\end{array}\right]
\quad;\quad
\mathrm{Sep2} := \left[\begin{array}{cc|cc}
2 & 1 & 1 & 1\\
1 & 1 & 1 & 1\\\hline
1 & 1 & 1 & 1\\
1 & 1 & 1 & 2
\end{array}\right]
$$

$$
\mathrm{Sep3} := \left[\begin{array}{ccc|ccc}
4 & & & & & \\
 & 4 & 2 & & & 2\\
 & 2 & 2 & 1 & -1 & \\\hline
 & & 1 & 2 & 1 & -1\\
 & & -1 & 1 & 5 & 1\\
 & 2 & & -1 & 1 & 2
\end{array}\right]
\quad;\quad
\mathrm{Ent1} = \left[\begin{array}{ccc|ccc|ccc}
1 & & & 1 & & & & & 1\\
 & 2 & & & 1 & & & & \\
 & & \tfrac12 & & & 1 & & & \\\hline
1 & & & \tfrac12 & & & 1 & & \\
1 & & & & 1 & & & & 1\\
 & & & & & 2 & & 1 & \\\hline
1 & & & & & 2 & & 1 & \\
 & & 1 & & & & 2 & \tfrac12 & \\
1 & & & 1 & & & & & 1
\end{array}\right].
$$

The separable states Sep1, Sep2, and Sep3 were previously studied for example in [7], where it was moreover shown that for a separable state ρ with local dimensions $(d_1, d_2) = (2, 3)$ and birank (r, s) one has $\mathrm{rank}_{\mathrm{sep}}(\rho) = \max\{r, s\}$. The entangled state Ent1 was constructed by Choi in [9] as the first example in dimension $(d_1, d_2) = (3, 3)$ of an entangled state ρ that satisfies the PPT condition. In addition we revisit four examples taken from [17]. They include three separable states, named here Sep4, Sep5 and Sep6 and corresponding to Examples 3.8, 3.10 and 3.11 in [17], and one entangled state, named here Ent2 and corresponding to Example 3.9 in [17].

In Section 3.1 we provided three different choices of localizing constraints in (33), (34) and (36), that we denote here as S1, S2 and S3, respectively. The examples show that the different choices lead to incomparable bounds. Indeed, let us the notation S1 < S2 as short hand for "there exists a ρ such that $\xi_t^{\mathrm{sep}}(\rho)$ (using scaling S1) $< \xi_t^{\mathrm{sep}}(\rho)$ (using scaling S2)". Then at level $t = 2$ the state Sep1 demonstrates both S3 < S1 and S2 < S1, and at level $t = 3$ Sep2 demonstrates both S2 < S3 and S1 < S3 and Sep3 demonstrates both S1 < S2 and S3 < S2. A case where the various constraints differ in ability to detect entanglement is provided by the state Ent1 at order $t = 2$. On the other hand, for the state Ent2, all three scalings detect entanglement at order $t = 2$ (thus at the same order as for the approach in [17]).

As mentioned in Section 4.1, there exist real states $\rho \in \mathcal{S}^d \otimes \mathcal{S}^d$ that are separable but do not admit a decomposition using real vectors $a_\ell, b_\ell \in \mathbb{R}^d$. Our bound $\xi_2^{\mathrm{sep},\mathbb{R}}(\rho)$ provides a proof of the latter for the state Sep3: its real separable rank is infinity since our lower bound is infeasible (i.e., there exists a dual certificate that proves $\mathrm{rank}_{\mathrm{sep}}^{\mathbb{R}}(\mathrm{Sep3}) = \infty$).
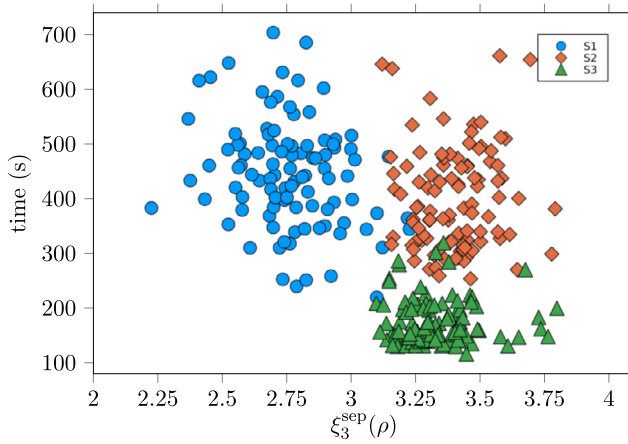
Fig. 1. Scatter plot of $\xi_3^{\text{sep}}(\rho)$ vs computation time (seconds) for 100 random matrices, grouped and colored by rescalings S1, S2 and S3.

Finally we note that one sometimes needs to go beyond level $t = 2$ (and thus beyond the PPT criterion) to reveal entanglement: with the localizing constraints S3 the bound for Ent1 is feasible at $t = 2$, but infeasible at $t = 3$. Going to a higher level naturally increases the size of the SDP. For the examples Sep3, Sep4, Sep6, and Ent1 this prevented us from computing level $t = 4$.

In addition, we show in Fig. 1 a scatter plot of the bound $\xi_3^{\text{sep}}(\rho)$ vs. its computation time in seconds for 100 random complex matrices ρ grouped and colored by the respective scalings S1, S2 and S3. These matrices are defined by $\rho = \sum_{j=1}^{5} a^{(j)}(a^{(j)})^* \otimes b^{(j)}(b^{(j)})^*$, where $a^{(j)}, b^{(j)} \in \mathbb{C}^3$ are random vectors whose entries are of the form $x + \mathbf{i}\,y$ with $x, y \in \mathcal{N}(0, 1)$. (We also normalize the trace here for numerical stability.) This construction guarantees separability and provides the upper bound $\text{rank}_{\text{sep}}(\rho) \leq 5$. Such states also satisfy the reverse inequality $\text{rank}_{\text{sep}}(\rho) \geq 5$ almost surely (since $\text{rank}(\rho) = 5$ almost surely). We use this class of examples merely to test the quality of the bounds. From the figure we can draw the following observations. First, the bounds are concentrated around the means 2.7, 3.4 and 3.3 for the scalings S1, S2 and S3, respectively. Second, in this class of examples the S1 rescaling yields inferior bounds as compared to S2 and S3. Third, out of the hundred examples and for the three different scalings considered, no bound exceeded the value 4.

4.3. Stronger bounds for the completely positive rank

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

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

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 [11,5]. We refer to [2] 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

$$A = \sum_{\ell=1}^{r} a_\ell a_\ell^T \tag{51}$$

for entrywise nonnegative vectors $a_\ell \in \mathbb{R}_+^d$ ($\ell \in [r]$). The smallest such r is called the *completely positive rank* of A and denoted as $\text{rank}_{\text{cp}}(A)$. In [22] the authors defined the parameter $\tau_{\text{cp}}(A)$ as

$$\tau_{\text{cp}}(A) := \inf \left\{ \lambda : \lambda > 0, \frac{1}{\lambda} \rho \in \text{conv}\{xx^T : x \in \mathbb{R}_+^d, \ xx^T \leq A, \ xx^T \preceq A\} \right\} \tag{52}$$

to lower bound the completely positive rank (as well as an SDP-based bound $\tau_{\text{cp}}^{\text{sos}}(A)$). In [26] the authors studied (among others) the completely positive rank from the polynomial optimization perspective and derived a hierarchy of semidefinite programming bounds, denoted here as $\xi_{t,(2019)}^{\text{cp}}(A)$. There the fact was 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

$$L((\mathbf{x}\mathbf{x}^T)^{\otimes \ell}) \preceq A^{\otimes \ell} \qquad \text{for all } \ell \in \mathbb{N} \tag{53}$$

for the linear functional arising from the atomic decomposition (51). Based on this the following bounds are defined in [26] and shown to converge to $\tau_{\text{cp}}(A)$ as $t \to \infty$:

$$\begin{aligned}
\xi_{t,(2019)}^{\text{cp}}(A) := \inf \Big\{ & L(1) : L \in [\mathbf{x}]_{2t}^*, \\
& L(\mathbf{x}\mathbf{x}^T) = A, \\
& L \geq 0 \text{ on } \mathcal{M}(\{\sqrt{A_{ii}} x_i - x_i^2 : i \in [d]\})_{2t}, \\
& L \geq 0 \text{ on } \mathcal{M}(\{A_{ij} - x_i x_j : i, j \in [d], i \neq j\})_{2t}, \\
& L((\mathbf{x}\mathbf{x}^T)^{\otimes \ell}) \preceq A^{\otimes \ell} \text{ for all } \ell \in [t] \Big\}.
\end{aligned} \tag{54}$$

The same convergence result holds if we replace the last constraint in (54) with the constraint

$$L \geq 0 \qquad \text{on } \mathcal{M}(\{v^T(A - \mathbf{x}\mathbf{x}^T)v : v \in \mathbb{R}^d\})_{2t}. \tag{55}$$

Using the same reasoning as in Section 3.2, we see that we can strengthen the parameter $\xi_{t,(2019)}^{\text{cp}}(A)$ by adding the constraint

$$M_{t-1}((A - \mathbf{x}\mathbf{x}^T) \otimes L) = L((A - \mathbf{x}\mathbf{x}^T) \otimes [\mathbf{x}]_{t-1}[\mathbf{x}]_{t-1}^T) \succeq 0. \tag{56}$$

Let $\xi_t^{\mathrm{cp}}(A)$ denote the parameter defined in this way, so that $\xi_{t,(2019)}^{\mathrm{cp}}(A) \leq \xi_t^{\mathrm{cp}}(A)$. Note that Lemmas 1 and 3 show that Eq. (56) implies Eq. (55). We now show that Eq. (56) in fact implies Eq. (53), which means that adding Eq. (56) strengthens both approaches provided in [26]; we present below numerical examples that illustrate this. 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*

$$M^{\mathrm{nc}}((A - xx^T) \otimes L) := L((A - xx^T) \otimes \langle x \rangle \langle x \rangle^T). \tag{57}$$

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 19. *Consider $A \in \mathbb{R}^{d \times d}$ and $L \in \mathbb{R}[\mathbf{x}]^*$. If $L(\mathbf{x}\mathbf{x}^T) = A$ and $M((A - \mathbf{x}\mathbf{x}^T) \otimes L) \succeq 0$, then Eq. (53) holds, i.e.,*

$$L((\mathbf{x}\mathbf{x}^T)^{\otimes \ell}) \preceq A^{\otimes \ell} \text{ for all } \ell \in \mathbb{N}.$$

Proof. As observed above, $M((A - \mathbf{x}\mathbf{x}^T) \otimes L) \succeq 0$ if and only if $M^{\mathrm{nc}}((A - \mathbf{x}\mathbf{x}^T) \otimes L) \succeq 0$. Note that for each $\ell \in \mathbb{N}$, the matrix $M^{\mathrm{nc}}((A - \mathbf{x}\mathbf{x}^T) \otimes L)$ contains $L((A - \mathbf{x}\mathbf{x}^T) \otimes (\mathbf{x}\mathbf{x}^T)^{\otimes \ell - 1})$ as a principal submatrix. To see this write the vector $\langle x \rangle$ of noncommutative monomials as $1 \oplus_{\ell \in \mathbb{N}} x^{\otimes \ell}$ by grouping the monomials according to their degree. With respect to this partition of its index set the matrix $M^{\mathrm{nc}}((A - xx^T) \otimes L)$ has the matrices $L((A - \mathbf{x}\mathbf{x}^*) \otimes (\mathbf{x}\mathbf{x}^*)^{\otimes \ell - 1})$ as its diagonal blocks. Since $M^{\mathrm{nc}}((A - \mathbf{x}\mathbf{x}^*) \otimes L) \succeq 0$, we obtain

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

Combined with $L(\mathbf{x}\mathbf{x}^*) = A$ this permits to show Eq. (53):

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

We conclude this section with some numerical results. To demonstrate the impact of the constraints (56) we compare our bounds $\xi_3^{\mathrm{cp}}(A)$ to the bounds $\xi_{3,(2019)}^{\mathrm{cp}}(A)$ from [26] on the cp-rank of some matrices A known to have a high cp-rank, taken from [4]. The boldface entries in Table 5 show a strict improvement in the bounds. For these computations we used the high precision solver SDPA-GMP [39] because MOSEK [15] and SDPA [61,60] could not certify solutions.[4]

---

[4] The code is available at: https://github.com/JAndriesJ/ju-cp-rank.

Table 5
Bounds for completely positive rank at level $t = 3$.

A	rank(A)	n	$\lfloor \frac{n^2}{4} \rfloor$	$\xi^{\mathrm{cp}}_{3,(2019)}(A)$	$\xi^{\mathrm{cp}}_3(A)$	rank$_{\mathrm{cp}}(A)$
M_7	7	7	12	10.5	**11.4**	14
$\widetilde{M}_7$	7	7	12	10.5	10.5	14
$\widetilde{M}_8$	8	8	16	13.82	**14.5**	18
$\widetilde{M}_9$	9	9	20	17.74	**18.4**	26

5. Entanglement witnesses

The moment approach we have developed in the previous section for bounding the separable rank of a state ρ can be viewed as searching for a (non-normalized) measure on the product of two balls, with the additional property that, for any point (x, y) in its support, we have $\rho - xx^* \otimes yy^* \succeq 0$. We will first observe in Section 5.1 how this approach can also be used to detect entanglement, i.e., non-membership in the set $\mathcal{SEP}$.

As mentioned earlier one can also capture the set $\mathcal{SEP}$ by viewing it as a moment problem on the bi-sphere (the product of two unit spheres). In the rest of this section we will show that this second moment approach corresponds exactly to the well-known state extension perspective that leads to the Doherty-Parrilo-Spedalieri hierarchy of approximations of $\mathcal{SEP}$ from [16].

5.1. Entanglement witnesses based on the hierarchy of parameters $\xi^{\mathrm{sep}}_t(\rho)$

Our approach to design lower bounds on the separable rank also directly leads to a way to detect non-membership in the set $\mathcal{SEP}$ or, in other words, to a way to witness entanglement of a state. Indeed, as shown in Lemma 15, a state ρ is separable if and only if $\xi^{\mathrm{sep}}_t(\rho) \leq \mathrm{rank}(\rho)^2$ for all $t \geq 2$. In other words, ρ is entangled if and only if $\xi^{\mathrm{sep}}_t(\rho) > \mathrm{rank}(\rho)^2$ for some integer $t \geq 2$ (which includes $\xi^{\mathrm{sep}}_t(\rho) = \infty$ in case the program defining $\xi^{\mathrm{sep}}_t(\rho)$ is infeasible).

In order to get a certificate of entanglement it is therefore convenient to consider the dual semidefinite program to the program (43) defining the parameter $\xi^{\mathrm{sep}}_t(\rho)$, which reads:

$$\sup \Big\{ \langle \rho, \Lambda \rangle \mid \Lambda \in \mathbb{C}^{d \times d} \otimes \mathbb{C}^{d \times d} \text{ Hermitian s.t. } 1 - \langle \Lambda, xx^* \otimes yy^* \rangle \in \mathcal{M}(S_\rho)_{2t}$$
$$+ \operatorname{cone}\{ \langle G_\rho, \vec{p}\vec{p}^{\,*} \rangle : \vec{p} \in (\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{t-2})^{d^2} \} \Big\}. \tag{58}$$

Lemma 20. *For any integer $t \geq 2$, the matrix $\Lambda = 0$ is a strictly feasible solution for* (58).

Proof. First we observe that, for small $\lambda > 0$, the matrix $\Lambda = \lambda \cdot I_{d^2}$ is a feasible solution for (58). For this we show that the polynomial $1 - \lambda \langle I, xx^* \otimes yy^* \rangle = 1 - \lambda \sum_i \sum_j x_i \overline{x}_i y_j \overline{y}_j$ lies in the quadratic module $\mathcal{M}(S_\rho)_{2t}$ for small $\lambda > 0$. We know that there exists a scalar $R > 0$ such that $R - \sum_i x_i \overline{x}_i - \sum_j y_j \overline{y}_j \in \mathcal{M}(\mathcal{S}_\rho)$. Then also $R - \sum_i x_i \overline{x}_i$ and $R - \sum_i y_i \overline{y}_i$

lie in $\mathcal{M}(S_\rho)_{2t}$, as well as $(R - \sum_i x_i \overline{x}_i)(R + \sum_i y_i \overline{y}_i)$ and $(R + \sum_i x_i \overline{x}_i)(R - \sum_i y_i \overline{y}_i)$. Adding up the latter two polynomials we obtain that the polynomial $R^2 - \sum_i \sum_j x_i \overline{x}_i y_j \overline{y}_j$ belongs to $\mathcal{M}(S_\rho)_{2t}$, which shows that $\Lambda = \lambda \cdot I_{d^2}$ is feasible for all $0 < \lambda \leq R^{-2}$.

We now show that any Λ satisfying $\|\Lambda\| \leq R^{-2}$ is feasible (which shows there is a ball contained in the feasible region of (58)). For this write

$$1 - \langle \Lambda, xx^* \otimes yy^* \rangle = \underbrace{1 - \langle \|\Lambda\| I_{d^2}, xx^* \otimes yy^* \rangle}_{(a)} + \underbrace{\langle \|\Lambda\| I_{d^2} - \Lambda, xx^* \otimes yy^* \rangle}_{(b)}.$$

In the first part of the proof we have shown that term (a) belongs to $\mathcal{M}(S_\rho)_{2t}$ if $\|\Lambda\| \leq R^{-2}$. In addition, term (b) is a sum of squares since $\|\Lambda\| I_{d^2} - \Lambda$ is positive semidefinite. Together, this shows $1 - \langle \Lambda, xx^* \otimes yy^* \rangle \in \mathcal{M}(S_\rho)_{2t}$ and therefore Λ is feasible. □

As a consequence, strong duality holds between the program (43) defining $\xi_t^{\text{sep}}(\rho)$ and its dual (58). That is, if the program (58) is bounded then its optimal value is finite and equal to $\xi_t^{\text{sep}}(\rho)$ and, otherwise, its optimal value is equal to ∞ and thus $\xi_t^{\text{sep}}(\rho)$ is infeasible. Therefore, we obtain that ρ is entangled if and only if, for some integer $t \geq 2$, there exists a matrix $\Lambda \in \mathbb{C}^{d \times d} \otimes \mathbb{C}^{d \times d}$ which is feasible for (58) and satisfies $\langle \rho, \Lambda \rangle > \text{rank}(\rho)^2$. In that case such matrix Λ provides a certificate that the state ρ is entangled.

5.2. The Doherty-Parrilo-Spedalieri hierarchy: moment perspective

Recall definition (1) of the set of separable states $\mathcal{SEP}_d$, so $\rho \in \mathcal{SEP}_d$ if and only if it is of the form

$$\rho = \sum_{\ell=1}^r \lambda_\ell \, a_\ell a_\ell^* \otimes b_\ell b_\ell^*, \tag{59}$$

where $\lambda_\ell > 0$, $a_\ell, b_\ell \in \mathbb{C}^d$ with $\|a_\ell\| = 1 = \|b_\ell\|$. To this decomposition we can associate a linear functional on $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]$ that is a conic combination of evaluation functionals at points on the bi-sphere: $L = \sum_{\ell=1}^r \lambda_\ell L_{(a_\ell, b_\ell)}$. By construction, this linear functional is positive on Hermitian squares, it vanishes on the ideal generated by $1 - \|\mathbf{x}\|^2$ and $1 - \|\mathbf{y}\|^2$ (called the *bi-sphere ideal* for short) and it satisfies $L(xx^* \otimes yy^*) = \rho$. This naturally suggests a hierarchy of outer approximations to the set $\mathcal{SEP}$: a state ρ belongs to the t-th level of this hierarchy if there exists an L that satisfies these constraints for polynomials of degree at most $2t$. Formally, we consider the set

$$\begin{aligned} \mathcal{R}_t := \{ \rho \in \mathcal{H}^d \otimes \mathcal{H}^d : \exists L : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{2t} \to \mathbb{C} \text{ Hermitian s.t.} \\ L(\mathbf{xx}^* \otimes \mathbf{yy}^*) = \rho, \\ L = 0 \text{ on } \mathcal{I}(1 - \|\mathbf{x}\|^2, 1 - \|\mathbf{y}\|^2)_{2t}, \\ M_t(L) \succeq 0 \}. \end{aligned} \tag{60}$$

We will show in Section 5.4 that this set is in fact closely related to the DPS hierarchy of outer approximations to the set $\mathcal{SEP}_d$: if we introduce separate degree-bounds on the $\mathbf{x}, \overline{\mathbf{x}}$ variables and the $\mathbf{y}, \overline{\mathbf{y}}$ variables, then we recover the original formulation from [16].

First we note that we can easily show that the sets $\mathcal{R}_t$ converge to $\mathcal{SEP}$, i.e., $\mathcal{SEP}_d = \bigcap_{t \geq 2} \mathcal{R}_t$, using the tools from polynomial optimization (Theorem 8).

Proposition 21. *We have:* $\mathcal{SEP}_d = \bigcap_{t \geq 2} \mathcal{R}_t$.

Proof. Assume $\rho \in \bigcap_{t \geq 2} \mathcal{R}_t$, we show $\rho \in \mathcal{SEP}_d$. For any $t \geq 2$ let L_t be an associated certificate of membership in $\mathcal{R}_t$. Then we have $L_t(1) = L_t(\|\mathbf{x}\|^2 \|\mathbf{y}\|^2) = \text{Tr}(\rho)$. Hence it follows from Lemma 7 that the sequence $(L_t)_t$ has a pointwise converging subsequence, with limit $L \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]^*$. Then $L \geq 0$ on Σ and $L = 0$ on $\mathcal{I}(1 - \|\mathbf{x}\|^2, 1 - \|\mathbf{y}\|^2)$. Using Theorem 8 we can conclude that there exist scalars $\mu_\ell > 0$ and points $(a_\ell, b_\ell) \in \mathbb{C}^d \times \mathbb{C}^d$ with $\|a_\ell\| = \|b_\ell\| = 1$ such that $L(p) = \sum_{\ell=1}^K \mu_\ell p(a_\ell, b_\ell)$ when p has degree at most 4. In particular, we obtain

$$\rho = L(\mathbf{x}\mathbf{x}^* \otimes \mathbf{y}\mathbf{y}^*) = \sum_{\ell=1}^K \mu_\ell \, a_\ell a_\ell^* \otimes b_\ell b_\ell^*,$$

which shows that $\rho \in \mathcal{SEP}_d$. $\quad\square$

Next, we reformulate the positivity condition $M_t(L) \succeq 0$ in a way that will be useful for making the link to the DPS hierarchy. As observed in Section 3.3, we may additionally require the linear functionals in Eq. (60) to satisfy the constraint Eq. (46), which we repeat here for convenience:

$$L(\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'}) = 0 \ \text{ if } |\alpha| \neq |\alpha'| \text{ or } |\beta| \neq |\beta'|. \tag{61}$$

This permits to block-diagonalize the associated moment matrix $M_t(L)$ according to the partition given in Eqs. (47) and (48), thus permitting to replace the constraint $M_t(L) \succeq 0$ by $M_t(L)[I_{r,s}^t] \succeq 0$ for $r, s \in [-t, t]$. In fact, using the bi-sphere ideal constraint, one can reduce the size of these matrices even further and replace the matrices $M_t(L)[I_{r,s}^t]$ by their submatrices $M_t(L)[I_{r,s}^{=t}]$, where the sets $I_{r,s}^{=t} \subseteq I_{r,s}^t$ are defined by

$$I_{r,s}^{=t} := \left\{ (\alpha, \alpha', \beta, \beta') \in (\mathbb{N}^d)^4 : |\alpha + \alpha' + \beta + \beta'| = t, \ |\alpha| - |\alpha'| = r, \ |\beta| - |\beta'| = s \right\}. \tag{62}$$

In other words we can show the following reformulation of the set $\mathcal{R}_t$:

$$\begin{aligned}
\mathcal{R}_t = \big\{ \rho \in \mathcal{H}^d \otimes \mathcal{H}^d : \exists L : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{2t} &\to \mathbb{C} \text{ Hermitian s.t.} \\
L(\mathbf{x}\mathbf{x}^* \otimes \mathbf{y}\mathbf{y}^*) &= \rho, \\
L = 0 \text{ on } \mathcal{I}(1 - \|\mathbf{x}\|^2, 1 - \|\mathbf{y}\|^2)_{2t}, & \\
M_t(L)[I_{r,s}^{=t}] \succeq 0 \text{ for all } r, s &\in \{-t, -t+1, \ldots, t\} \big\}.
\end{aligned} \tag{63}$$

We will show this result in a slightly different setting (closer to that of the original formulation of the DPS hierarchy). Similar arguments as those used in the proof of Lemma 22 below can be used to show the equivalence between (60) and (63).

In order to connect the moment approach on the bi-sphere to the original formulation of the DPS hierarchy we need to introduce a separate degree bound on the $\mathbf{x}, \overline{\mathbf{x}}$ variables and the $\mathbf{y}, \overline{\mathbf{y}}$ variables. For integers $k, t \geq 1$ we let $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{k,t}$ (resp., $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{=k,=t}$) denote the set of polynomials that have degree at most k (resp., equal to k) in $\mathbf{x}, \overline{\mathbf{x}}$ and degree at most t (resp., equal to t) in $\mathbf{y}, \overline{\mathbf{y}}$, and we set

$$\Sigma_{2k,2t} = \operatorname{cone}\{p\overline{p} : p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{k,t}\} = \Sigma \cap \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{2k,2t},$$

$$\Sigma_{=2k,=2t} = \Sigma \cap \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{=2k,=2t}.$$

We define the sets

$$
\begin{aligned}
\mathcal{R}_{k,t} := \{\rho \in \mathcal{H}^d \otimes \mathcal{H}^d : &\exists L : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{2k,2t} \to \mathbb{C} \text{ Hermitian s.t.} \\
& L(\mathbf{x}\mathbf{x}^* \otimes \mathbf{y}\mathbf{y}^*) = \rho, \\
& L = 0 \text{ on } \mathcal{I}(1 - \|\mathbf{x}\|^2, 1 - \|\mathbf{y}\|^2)_{2k,2t}, \\
& L \geq 0 \text{ on } \Sigma_{2k,2t}\}.
\end{aligned}
\tag{64}
$$

Note the inclusion $\mathcal{R}_{k,t} \subseteq \mathcal{R}_{k+t}$. The two regimes that we will be interested in are $k = 1$ and $k = t$ since, as we will show in Section 5.4, the sets $\mathcal{R}_{1,t}$ and $\mathcal{R}_{t,t}$ (for $t \in \mathbb{N}$) coincide with the approximation hierarchies $\mathcal{DPS}_{1,t}$ and $\mathcal{DPS}_{t,t}$ from [16].

We will give a more economical reformulation for the positivity condition on L in Proposition 23. For this we first show that for linear functionals L that vanish on the bi-sphere ideal and satisfy (61) the following two positivity conditions are equivalent: $L \geq 0$ on $\Sigma_{2k,2t}$ and $L \geq 0$ on $\Sigma_{=2k,=2t}$. That is, we only need to require positivity on *homogeneous* polynomials.

Lemma 22. *Let $L \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{2k,2t}^*$ be such that $L = 0$ on $\mathcal{I}(1 - \|\mathbf{x}\|^2, 1 - \|\mathbf{y}\|^2)_{2k,2t}$ and L satisfies Eq. (61). Then we have $L \geq 0$ on $\Sigma_{2k,2t}$ if and only if $L \geq 0$ on $\Sigma_{=2k,=2t}$.*

Proof. Assume $L \geq 0$ on $\Sigma_{=2k,=2t}$. We show that L is positive on Hermitian squares in $\Sigma_{2k,2t}$. Let $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{k,t}$, we want to show that $L(p\overline{p}) \geq 0$. For this we decompose p as $p = p_{00} + p_{01} + p_{10} + p_{11}$, where, for $a, b \in \{0, 1\}$, we group in p_{ab} the terms of p that involve a monomial $\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'}$ with $|\alpha + \alpha'| \equiv k - a$ modulo 2 and $|\beta + \beta'| \equiv t - b$ modulo 2. Then we have

$$L(p\overline{p}) = \sum_{a,b \in \{0,1\}} L(p_{ab}\overline{p}_{ab}),$$

where we use Eq. (61) to see that $L(p_{ab}\overline{p}_{a'b'}) = 0$ if $(a, b) \neq (a', b')$. Hence it remains to show that $L(p_{ab}\overline{p}_{ab}) \geq 0$ for $a, b \in \{0, 1\}$. Write $p_{ab} = \sum c_{\alpha\alpha'\beta\beta'}\mathbf{x}^\alpha\overline{\mathbf{x}}^{\alpha'}\mathbf{y}^\beta\overline{\mathbf{y}}^{\beta'}$. We define the polynomial

$$q_{ab} = \sum c_{\alpha\alpha'\beta\beta'}\mathbf{x}^\alpha\overline{\mathbf{x}}^{\alpha'}\mathbf{y}^\beta\overline{\mathbf{y}}^{\beta'}\|\mathbf{x}\|^{k-a-|\alpha+\alpha'|}\|\mathbf{y}\|^{t-b-|\beta+\beta'|}.$$

Note that in each term the powers of $\|\mathbf{x}\|$ and $\|\mathbf{y}\|$ are by construction both even and nonnegative and therefore the polynomial q_{ab} is homogeneous of degree $k - a$ in $\mathbf{x}, \overline{\mathbf{x}}$ and of degree $t - b$ in $\mathbf{y}, \overline{\mathbf{y}}$. Since L vanishes on the truncated ideal generated by $1 - \|\mathbf{x}\|^2$ and $1 - \|\mathbf{y}\|^2$ we have

$$L(p_{ab}\overline{p}_{ab}) = L(q_{ab}\overline{q}_{ab}) = L(q_{ab}\overline{q}_{ab}\|\mathbf{x}\|^{2a}\|\mathbf{y}\|^{2b}) \geq 0,$$

where the last inequality follows from the fact that $q_{ab}\overline{q}_{ab}\|\mathbf{x}\|^{2a}\|\mathbf{y}\|^{2b} \in \Sigma_{=2k,=2t}$. □

We now proceed to define the analog of Eq. (62) for the (k, t)-setting. Given two integers $r \in \{-k, -k+2, -k+4, \ldots, k\}$ and $s \in \{-t, -t+2, -t+4, \ldots, t\}$ define the set of (exponents of) monomials

$$I_{r,s}^{=k,=t} := \Big\{(\alpha, \alpha', \beta, \beta') \in (\mathbb{N}^d)^4 : |\alpha + \alpha'| = k, \ |\alpha| - |\alpha'| = r, \tag{65}$$
$$|\beta + \beta'| = t, \ |\beta| - |\beta'| = s\Big\}.$$

Note that we restrict our attention to $r \equiv k \mod 2$ and $s \equiv t \mod 2$. If r, s do not satisfy these conditions then $I_{r,s}^{=k,=t} = \emptyset$. We then have the following semidefinite representation of $\mathcal{R}_{k,t}$.

Proposition 23. *For $k, t \in \mathbb{N}$ we have*

$$\mathcal{R}_{k,t} = \Big\{\rho \in \mathcal{H}^d \otimes \mathcal{H}^d : \exists L : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{2k,2t} \to \mathbb{C} \text{ Hermitian s.t.}$$
$$L(\mathbf{x}\mathbf{x}^* \otimes \mathbf{y}\mathbf{y}^*) = \rho,$$
$$L = 0 \text{ on } \mathcal{I}(1 - \|\mathbf{x}\|^2, 1 - \|\mathbf{y}\|^2)_{2k,2t}, \tag{66}$$
$$M_{k,\ell}(L)[I_{r,s}^{=k,=t}] \succeq 0 \text{ for all } r \in [-k, k], s \in [-t, t]\Big\}.$$

Proof. As mentioned above, we may add the constraint (61) to the program (64). It then follows from Lemma 22 that we may replace the condition $L \geq 0$ on $\Sigma_{2k,2t}$ with $L \geq 0$ on $\Sigma_{=2k,=2t}$. Finally we observe that the index sets $I_{r,s}^{=k,=t}$ block-diagonalize $M_{=k,=t}(L)$. Indeed, let $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{=k,=t}$ and write $p = \sum_{r,s} p_{r,s}$ where $p_{r,s}$ is the polynomial corresponding to the terms of p with exponents in $I_{r,s}^{=k,=t}$. Then $p_{r,s}\overline{p}_{r',s'}$ is a linear combination of monomials of the form

$$\mathbf{x}^\alpha\overline{\mathbf{x}}^{\alpha'}\mathbf{y}^\beta\overline{\mathbf{y}}^{\beta'}\overline{\mathbf{x}}^\gamma\mathbf{x}^{\gamma'}\overline{\mathbf{y}}^\delta\mathbf{y}^{\delta'},$$

where we have the following for the degrees in $\mathbf{x}, \overline{\mathbf{x}}$. By assumption $|\alpha| - |\alpha'| = r$ and $|\gamma| - |\gamma'| = r'$, and therefore the degree in $\mathbf{x}$ minus the degree in $\overline{\mathbf{x}}$ is $(|\alpha| + |\gamma'|) - (|\alpha'| + |\gamma|) = r - r'$. Similarly, the degree in $\mathbf{y}$ minus the degree in $\overline{\mathbf{y}}$ equals $s - s'$. Hence, if $(r, s) \neq (r', s')$, then Eq. (61) shows that $L(p_{r,s} \overline{p}_{r',s'}) = 0$. $\square$

Finally, we observe the following alternative formulation of the positivity conditions in Eq. (66) in terms of the noncommutative moment matrices (cf. Eq. (57)):

$$M_{k,t}(L)[I_{r,s}^{=k,=t}] \succeq 0$$
$$\Longleftrightarrow L\big((\mathbf{x}\mathbf{x}^*)^{\otimes(k+r)/2} \otimes (\overline{\mathbf{x}}\,\overline{\mathbf{x}}^*)^{\otimes(k-r)/2} \otimes (\mathbf{y}\mathbf{y}^*)^{\otimes(t+s)/2} \otimes (\overline{\mathbf{y}}\,\overline{\mathbf{y}}^*)^{\otimes(t-s)/2}\big) \succeq 0.$$
$$(67)$$

Although less efficient, this reformulation will permit to connect the program (66) to the original formulation of the DPS hierarchy $\mathcal{DPS}_{1,t}$ (see the proof of Proposition 29).

The analog of Proposition 21 holds for the sets $\mathcal{R}_{k,t}$:

$$\bigcap_{k,t \geq 1} \mathcal{R}_{k,t} = \bigcap_{t \geq 2} \mathcal{R}_{t,t} = \mathcal{SEP}_d;$$

the argument is similar, based on standard tools from polynomial optimization (Theorem 8). In fact, even the (weaker) sets $\mathcal{R}_{1,t}$ already converge to $\mathcal{SEP}$, i.e., we have

$$\bigcap_{t \geq 2} \mathcal{R}_{1,t} = \mathcal{SEP}_d; \tag{68}$$

in other words, in the moment approach it is sufficient to let only the degree in $\mathbf{y}, \overline{\mathbf{y}}$ grow. We will show this in Theorem 25 below, using the tools about matrix-valued polynomial optimization (Theorem 9).

5.3. Convergence of the sets $\mathcal{R}_{1,t}$ to $\mathcal{SEP}$

We first reformulate the set $\mathcal{R}_{1,t}$ from Eqs. (64) and (66) (case $k = 1$) in terms of matrix-valued linear functionals $\mathcal{L}$ on the polynomial space $\mathbb{C}[\mathbf{y}, \overline{\mathbf{y}}]$.

Lemma 24. *For $t \in \mathbb{N}$ we have*

$$\mathcal{R}_{1,t} = \big\{\rho \in \mathcal{H}^d \otimes \mathcal{H}^d : \exists \mathcal{L} : \mathbb{C}[\mathbf{y}, \overline{\mathbf{y}}]_{2t} \to \mathcal{H}^d \ \text{Hermitian s.t.}$$
$$\mathcal{L}(\mathbf{y}\mathbf{y}^*) = \rho,$$
$$\mathcal{L} = 0 \ \text{on} \ \mathcal{I}(1 - \|\mathbf{y}\|^2)_{2t},$$
$$M_t(\mathcal{L}) \succeq 0\big\}. \tag{69}$$

Proof. Let us use $\widehat{\mathcal{R}}_{1,t}$ to denote the set defined in (69). We show that $\mathcal{R}_{1,t} = \widehat{\mathcal{R}}_{1,t}$ using the formulation of $\mathcal{R}_{1,t}$ given in Eq. (64).

First consider $\rho \in \mathcal{R}_{1,t}$ and let $L : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{2,2t} \to \mathbb{C}$ be an associated certificate. Define $\mathcal{L} : \mathbb{C}[\mathbf{y}, \overline{\mathbf{y}}]_{2t} \to \mathcal{H}^d$ by

$$\mathcal{L}(p) = L(\mathbf{x}\mathbf{x}^* p) = \left(L(x_i \overline{x}_j p) \right)_{i,j=1}^d \qquad \text{for all } p \in \mathbb{C}[\mathbf{y}, \overline{\mathbf{y}}]_{2t}.$$

So $\mathcal{L} = (L_{ij})_{i,j=1}^d$ with $L_{ij}(p) = L(x_i \overline{x}_j p)$. By construction $\mathcal{L}(\mathbf{y}\mathbf{y}^*) = \rho$. To see that $\mathcal{L} = 0$ on $\mathcal{I}(1 - \|\mathbf{y}\|^2)_{2t}$, it suffices to observe that, for any $p \in \mathcal{I}(1 - \|\mathbf{y}\|^2)_{2t}$ and any $i, j \in [d]$, the polynomial $x_i \overline{x}_j p$ lies in $\mathcal{I}(1 - \|\mathbf{x}\|^2, 1 - \|\mathbf{y}\|^2)_{2,2t}$. To show that $M_t(\mathcal{L}) \succeq 0$ we use (the degree truncated version of) Lemma 2. That is, we use that $M_t(\mathcal{L}) \succeq 0$ is equivalent to $\sum_{i,j \in [d]} L_{ij}(p_i \overline{p}_j) \geq 0$ for all $(p_1, \ldots, p_d) \in (\mathbb{C}[\mathbf{y}, \overline{\mathbf{y}}]_t)^d$. We have $\sum_{i,j \in [d]} L_{ij}(p_i \overline{p}_j) = L((\sum_i x_i p_i)(\sum_i x_i p_i)^*) \geq 0$, where the last inequality follows from the fact that $(\sum_i x_i p_i)(\sum_i x_i p_i)^* \in \Sigma_{2,2t}$. This shows that if $\rho \in \mathcal{R}_{1,t}$, then $\rho \in \widehat{\mathcal{R}}_{1,t}$.

Conversely, let $\rho \in \widehat{\mathcal{R}}_{1,t}$ and let $\mathcal{L} : \mathbb{C}[\mathbf{y}, \overline{\mathbf{y}}]_{2t} \to \mathcal{H}^d$ be an associated certificate. We write $\mathcal{L}(p) = \left(L_{ij}(p) \right)_{i,j \in [d]}$ with $L_{ij} \in \mathbb{C}[\mathbf{y}, \overline{\mathbf{y}}]_{2t}^*$ for all $i, j \in [d]$. We define a linear functional L on $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{2,2t}$ as follows. For a polynomial $p \in \mathbb{C}[\mathbf{y}, \overline{\mathbf{y}}]_{2t}$ we set $L(p) = \sum_{i \in [d]} L_{ii}(p)$ and, for each $i, j \in [d]$, we set $L(x_i \overline{x}_j p) = L_{ij}(p)$. We extend L to $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{2,2t}$ by setting $L(x^\alpha \overline{x}^\beta p) = 0$ for all $\alpha, \beta \in \mathbb{N}^d$ with $(|\alpha|, |\beta|) \notin \{(0,0), (1,1)\}$, and extending by linearity. We show that L is a certificate for $\rho \in \mathcal{R}_{1,t}$. First observe that $L(\mathbf{x}\mathbf{x}^* \otimes \mathbf{y}\mathbf{y}^*) = \left(L(x_i \overline{x}_j \mathbf{y}\mathbf{y}^*) \right)_{i,j \in [d]} = \mathcal{L}(\mathbf{y}\mathbf{y}^*) = \rho$. By construction we have $L((1 - \sum_i x_i \overline{x}_i) p) = 0$ for all $p \in \mathbb{C}[\mathbf{y}, \overline{\mathbf{y}}]_{2t}$. Moreover, by assumption, $\mathcal{L}((1 - \sum_i y_i \overline{y}_i) p) = 0$ for all $p \in \mathbb{C}[\mathbf{y}, \overline{\mathbf{y}}]_{2t-2}$. Using the construction of L, this implies that $L((1 - \sum_i y_i \overline{y}_i) p) = 0$ for all $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{2,2t-2}$. Together, this shows that $L = 0$ on $\mathcal{I}(1 - \|\mathbf{x}\|^2, 1 - \|\mathbf{y}\|^2)_{2,2t}$.

It remains to show that $L \geq 0$ on $\Sigma_{2,2t}$. Let $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{1,t}$, we show that $L(p\overline{p}) \geq 0$. For this, write $p = p_0 + p_1 + p_2$, where p_0 has degree 0 in $\mathbf{x}, \overline{\mathbf{x}}$, p_1 has degree $(1,0)$ in $(\mathbf{x}, \overline{\mathbf{x}})$, and p_2 has degree $(0,1)$ in $(\mathbf{x}, \overline{\mathbf{x}})$. By definition, $L(p_a \overline{p}_b) = 0$ if $a \neq b$ and thus $L(p\overline{p}) = L(p_0\overline{p}_0) + L(p_1\overline{p}_1) + L(p_2\overline{p}_2)$. We have $L(p_0\overline{p}_0) = \sum_{i=1}^d L_{ii}(p_0\overline{p}_0) \geq 0$, since $M_t(L_{ii}) \succeq 0$ for each $i \in [d]$ as $M_t(\mathcal{L}) \succeq 0$. Next we show that $L(p_1\overline{p}_1) \geq 0$. To do so, write $p_1 = \sum_{i=1}^d x_i q_i$ where $q_i \in \mathbb{C}[\mathbf{y}, \overline{\mathbf{y}}]_t$ for $i \in [d]$. It then follows that $L(p_1\overline{p}_1) = \sum_{i,j=1}^d L(x_i \overline{x}_j q_i \overline{q}_j) = \sum_{i,j=1}^d L_{ij}(q_i \overline{q}_j) = \langle \mathcal{L}, \vec{q}\vec{q}^* \rangle \geq 0$ where $\vec{q} = (q_1, \ldots, q_d) \in (\mathbb{C}[\mathbf{y}, \overline{\mathbf{y}}]_t)^d$ and the last inequality follows from $M_t(\mathcal{L}) \succeq 0$ (using Lemma 2). This also directly implies that $L(p_2\overline{p}_2) \geq 0$. It follows that $L \geq 0$ on $\Sigma_{2,2t}$ and thus $\rho \in \mathcal{R}_{1,t}$. $\square$

We can now show the convergence of the sets $\mathcal{R}_{1,t}$ to $\mathcal{SEP}$. The proof is analogous to that of Proposition 21, except that it now relies on the results for matrix-valued polynomial optimization (Theorems 9 and 10).

Theorem 25. *We have $\mathcal{SEP}_d = \bigcap_{t \geq 2} \mathcal{R}_{1,t}$.*

Proof. Assume $\rho \in \bigcap_{t \geq 2} \mathcal{R}_{1,t}$ and for each $t \geq 2$ let $\mathcal{L}_t$ be a corresponding certificate for membership in $\mathcal{R}_{1,t}$. Using Lemma 7 one can show that the sequence $(\mathcal{L}_t)_t$ has a

pointwise converging subsequence. Let $\mathcal{L}$ be its limit. It then follows from Theorems 9 and 10 that there exists a $K \in \mathbb{N}$, matrices $\Lambda_1, \ldots, \Lambda_K \in \mathcal{S}_+^d$, and vectors $v_1, \ldots, v_K \in \mathbb{C}^d$ with $\|v_i\| = 1$ such that

$$\mathcal{L}(p) = \sum_{k=1}^{K} \Lambda_k p(v_k) \qquad \text{for all } p \in \mathbb{C}[\mathbf{y}, \overline{\mathbf{y}}]_2.$$

In particular,

$$\rho = \mathcal{L}(\mathbf{y}\mathbf{y}^*) = \sum_{k=1}^{K} \Lambda_k \otimes v_k v_k^*,$$

which shows that $\rho \in \mathcal{SEP}_d$. $\square$

5.4. The Doherty-Parrilo-Spedalieri hierarchy: state extension perspective

In the previous section we introduced the sets $\mathcal{R}_{k,t}$ for integers $k, t \geq 1$ and we mentioned that there are two regimes of interest: $k = 1$ and $k = t$, leading to the two hierarchies $\mathcal{R}_{1,t}$ and $\mathcal{R}_{t,t}$. We now show that these hierarchies in fact coincide with the DPS hierarchies, denoted here as $\mathcal{DPS}_{1,t}$ and $\mathcal{DPS}_{t,t}$, that are defined in terms of (one-sided and two-sided) state extensions [16]. For ease of notation, we will mostly focus on the first regime and show the equality $\mathcal{R}_{1,t} = \mathcal{DPS}_{1,t}$; the arguments naturally adapt to the second regime to show $\mathcal{R}_{t,t} = \mathcal{DPS}_{t,t}$. This permits to recover the convergence of the DPS hierarchy to $\mathcal{SEP}$ from the corresponding convergence result for the sets $\mathcal{R}_{1,t}$ (Theorem 25) obtained via the moment approach.

We begin with giving the original formulation of the DPS hierarchy $\mathcal{DPS}_{1,t}$ in terms of (one-sided) state extensions. To do so, we require a few definitions.

Given a bipartite state $\rho \in \mathcal{H}^d \otimes \mathcal{H}^d$ it is convenient to denote the two vector spaces (aka registers) composing the tensor product space on which ρ acts as A and B. Then we may also denote ρ as ρ_{AB}. The *partial trace* of ρ_{AB} with respect to the second register is the operator $\rho_A = \mathrm{Tr}_B(\rho)$ that acts on the first register and is defined by tracing out the second register. In the same way, $\rho_B = \mathrm{Tr}_A(\rho)$ is the second partial trace, which acts on the second register and is obtained by tracing out the first one. Concretely, say $\rho = (\rho_{ij,i'j'})_{i,j,i',j' \in [d]}$ after fixing a basis of $\mathbb{C}^d \otimes \mathbb{C}^d$. Then we have

$$\rho_A = \mathrm{Tr}_B(\rho) = \Big(\sum_{j=1}^{d} \rho_{ij,i'j} \Big)_{i,i'=1}^{d} \quad \text{and} \quad \rho_B = \mathrm{Tr}_A(\rho) = \Big(\sum_{i=1}^{d} \rho_{ij,ij'} \Big)_{j,j'=1}^{d}.$$

The *partial transpose* ρ^{T_B} of ρ with respect to the second register B is defined by

$$(\rho^{T_B})_{ij,i'j'} = \rho_{ij',i'j} \quad \text{for all } i, i' \in [d], \ j, j' \in [d] \tag{70}$$

and T_{B} denotes the corresponding transpose operator that acts on $\mathcal{H}^d \otimes \mathcal{H}^d$ by taking the partial transpose on the second register, so that $\mathsf{T}_{\mathsf{B}}(\rho) = \rho^{T_B}$. The partial transpose ρ^{T_A} with respect to the first register is defined analogously by $(\rho^{T_A})_{ij,i'j'} = \rho_{i'j,ij'}$ for all $i, i', j, j' \in [d]$. Note that $\rho^{T_A} = (\rho^{T_B})^T = \overline{\rho^{T_B}}$ if ρ is Hermitian, and thus $\rho^{T_B} \succeq 0$ implies $\rho^{T_A} \succeq 0$.

Given an integer $t \geq 2$ the construction of the relaxation $\mathcal{DPS}_{1,t}$ relies on the following observation: If ρ_{AB} has a decomposition as in (59) then one may introduce t copies of the second register and define the following extended state $\rho_{AB_{[t]}}$ acting on $\mathbb{C}^d \otimes (\mathbb{C}^d)^{\otimes t}$:

$$\rho_{AB_{[t]}} := \sum_{\ell=1}^{r} \lambda_l \, x_\ell x_\ell^* \otimes (y_\ell y_\ell^*)^{\otimes t}. \tag{71}$$

There is a natural action of the symmetric group $\mathrm{Sym}(t)$ on $(\mathbb{C}^d)^{\otimes t}$, defined by $\sigma(v_1 \otimes \ldots \otimes v_t) = v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(t)}$ for $v_1, \ldots, v_t \in \mathbb{C}^d$ and $\sigma \in \mathrm{Sym}(t)$, and extended to the space $(\mathbb{C}^d)^{\otimes t}$ by linearity. Let $\mathrm{Sym}((\mathbb{C}^d)^{\otimes t})$ denote the invariant subspace of $(\mathbb{C}^d)^{\otimes t}$ under this action and let Π_t denote the projection from $(\mathbb{C}^d)^{\otimes t}$ onto its invariant subspace $\mathrm{Sym}((\mathbb{C}^d)^{\otimes t})$, defined by

$$\Pi_t(w) = \frac{1}{t!} \sum_{\sigma \in \mathrm{Sym}(t)} \sigma(w) \quad \text{for } w \in (\mathbb{C}^d)^{\otimes t}.$$

Then, $I_A \otimes \Pi_t$ acts onto $\mathbb{C}^d \otimes (\mathbb{C}^d)^{\otimes t}$.

We now present some natural properties that the extended state $\rho_{AB_{[t]}}$ from (71) satisfies:

(1) $\rho_{AB_{[t]}}$ is positive semidefinite.
(2) $\rho_{AB} = \mathrm{Tr}_{B_{[2:t]}}(\rho_{AB_{[t]}})$, where, in $\mathrm{Tr}_{B_{[2:t]}}(\rho(AB_{[t]}))$, we trace out the last $t-1$ copies of the second register B.
(3) $(I_A \otimes \Pi_t)\rho_{AB_{[t]}}(I_A \otimes \Pi_t) = \rho_{AB_{[t]}}$, i.e., $\rho_{AB_{[t]}}$ is symmetric in the last t registers.
(4) $I_A \otimes \mathsf{T}_B^{\otimes s} \otimes I_B^{\otimes(t-s)}(\rho_{AB_{[t]}}) \succeq 0$ for any $1 \leq s \leq t$.

For property (2) we use the fact that each vector y_ℓ lies in the unit sphere and the last property (4) follows from the fact that $\mathsf{T}_B(yy^*) = (yy^*)^T = \overline{y}\,\overline{y}^*$ and thus

$$I_A \otimes \mathsf{T}_B^{\otimes s} \otimes I_B^{\otimes(t-s)}(\rho_{AB_{[t]}}) = \sum_{\ell=1}^{r} \lambda_\ell x_\ell x_\ell^* \otimes (\overline{y}_\ell \, \overline{y}_\ell^*)^{\otimes s} \otimes (y_\ell y_\ell^*)^{\otimes(t-s)} \succeq 0$$

if $\rho_{AB_{[t]}}$ satisfies (71). Property (4) is known as the *positive partial transpose* (PPT) criterion or as the Peres-Horodecki criterion [31]. Clearly, in view of the symmetry property (3), taking the partial transpose of *any* s copies (thus not only the first s ones) among the t copies of the second register preserves positivity. The above properties are used to define the hierarchy $\mathcal{DPS}_{1,t}$.

Definition 26. For an integer $t \geq 2$ the DPS relaxation of order t is defined as

$$\mathcal{DPS}_{1,t} := \left\{ \rho_{AB} \in \mathcal{H}^d \otimes \mathcal{H}^d : \exists \rho_{1,t} \text{ Hermitian linear map acting on } \mathbb{C}^d \otimes (\mathbb{C}^d)^{\otimes t} \text{ s.t.} \right.$$
(72)

$$\operatorname{Tr}_{B_{[2:t]}}(\rho_{1,t}) = \rho_{AB}, \tag{73}$$

$$(I_A \otimes \Pi_t)\rho_{1,t}(I_A \otimes \Pi_t) = \rho_{1,t}, \tag{74}$$

$$\left. I_A \otimes \mathsf{T}_B^{\otimes s} \otimes I_B^{\otimes (t-s)}(\rho_{1,t}) \succeq 0 \text{ for all } s \in \{0\} \cup [t] \right\}. \tag{75}$$

Remark 27. In the definition of the set $\mathcal{DPS}_{1,t}$ only one part of the system is extended, which is why we refer to this as a *one-sided* state extension. One can define a stronger relaxation of $\mathcal{SEP}$ by considering a *two-sided* state extension. Given two integers $k, t \geq 2$ one can define $\mathcal{DPS}_{k,t}$ as the set of states ρ_{AB} that have an extension $\rho_{k,t}$ acting on $(\mathbb{C}^d)^{\otimes k} \otimes (\mathbb{C}^d)^{\otimes t}$, which satisfies the appropriate analogs of the above properties (1)-(4). One may consider in particular the case $k = t$, leading to the sets $\mathcal{DPS}_{t,t}$ that satisfy

$$\mathcal{SEP}_d \subseteq \mathcal{DPS}_{t,t} \subseteq \mathcal{DPS}_{1,t}.$$

Doherty, Parrilo and Spedaglieri [16] show that the relaxations $\mathcal{DPS}_{1,t}$ converge to $\mathcal{SEP}$.

Theorem 28 *([16]). We have $\mathcal{SEP}_d \subseteq \mathcal{DPS}_{1,t+1} \subseteq \mathcal{DPS}_{1,t}$ and $\mathcal{SEP}_d = \bigcap_{t \geq 1} \mathcal{DPS}_{1,t}$. As a consequence, we also have $\mathcal{SEP}_d = \bigcap_{t \geq 1} \mathcal{DPS}_{t,t}$.*

We now show that equality $\mathcal{R}_{1,t} = \mathcal{DPS}_{1,t}$ holds for all $t \in \mathbb{N}$. Therefore, Theorem 28 follows directly from Theorem 25. Using similar arguments one can also show that $\mathcal{DPS}_{k,t} = \mathcal{R}_{k,t}$ and thus $\mathcal{DPS}_{t,t} = \mathcal{R}_{t,t}$.

Proposition 29. *For any integer $t \geq 2$ we have $\mathcal{R}_{1,t} = \mathcal{DPS}_{1,t}$.*

Proof. Assume first $\rho_{AB} \in \mathcal{R}_{1,t}$, with certificate L satisfying Eq. (64) (with $k = 1, \ell = t$). We claim that $\rho_{1,t} := L(\mathbf{xx}^* \otimes (\mathbf{yy}^*)^{\otimes t})$ is a certificate for membership of ρ_{AB} in $\mathcal{DPS}_{1,t}$. Indeed, Eq. (73) holds since $\operatorname{Tr}_{B_{[2:t]}}(\rho_{1,t}) = L(\mathbf{x\overline{x}}^* \otimes \mathbf{y\overline{y}}^*) = \rho_{AB}$ follows using the bisphere ideal condition on L. The symmetry condition in Eq. (74) holds for $\rho_{1,t}$ since L acts on commutative polynomials, and the PPT condition in Eq. (75) holds for $\rho_{1,t}$ as a consequence of the positivity condition: $L \geq 0$ on $\Sigma_{2,2t}$.

Conversely, assume that $\rho_{AB} \in \mathcal{DPS}_{1,t}$, with state $\rho_{1,t}$ as certificate satisfying (72)-(75). We construct a linear functional L acting on $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{2,2t}$ that certifies membership of ρ_{AB} in $\mathcal{R}_{1,t}$, i.e., satisfies the program (64) (with $k = 1$). In a first step we set

$$L(\mathbf{xx}^* \otimes (\mathbf{yy}^*)^{\otimes t}) := \rho_{1,t}. \tag{76}$$

In other words we set

$$L(x_i \overline{x}_{i'} y_{j_1} \overline{y}_{j'_1} \cdots y_{j_t} \overline{y}_{j'_t}) := (\rho_{1,t})_{i\vec{\mathbf{j}}, i'\vec{\mathbf{j}}'}$$

for any $i, i' \in [d]$ and $\vec{\mathbf{j}} = (j_1, \ldots, j_t), \vec{\mathbf{j}}' = (j'_1, \ldots, j'_t) \in [d]^t$. Using the symmetry condition (74), it follows that this definition does not depend on the order of the variables y_j (or $\overline{y}_j$).

Indeed, by Eq. (74) we know that

$$(\rho_{1,t})_{i\vec{\mathbf{j}}, i'\vec{\mathbf{j}}'} = (\rho_{1,t})_{i\sigma(\vec{\mathbf{j}}), i'\tau(\vec{\mathbf{j}}')} \text{ for all } i, i' \in [d], \vec{\mathbf{j}}, \vec{\mathbf{j}}' \in [d]^t$$

for all permutations $\sigma, \tau \in \text{Sym}(t)$, where $\sigma(\vec{\mathbf{j}}) = (j_{\sigma(1)}, j_{\sigma(2)}, \ldots, j_{\sigma(t)})$ for $\vec{\mathbf{j}} = (j_1, \ldots, j_t)$, which shows that

$$L(x_i \overline{x}_{i'} y_{j_1} \overline{y}_{j'_1} \cdots y_{j_t} \overline{y}_{j'_t}) = L(x_i \overline{x}_{i'} y_{j_{\sigma(1)}} \overline{y}_{j'_{\tau(1)}} \cdots y_{j_{\sigma(t)}} \overline{y}_{j'_{\tau(t)}}).$$

This shows that $\rho_{1,t}$ defines a linear functional L acting on polynomials with degree 1 in $\mathbf{x}$, degree 1 in $\overline{\mathbf{x}}$, degree t in $\mathbf{y}$, and degree t in $\overline{\mathbf{y}}$. We now show how to extend this linear functional L to $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}, \mathbf{y}, \overline{\mathbf{y}}]_{2,2t}$ in such a way that it becomes a certificate for $\rho_{AB} \in \mathcal{R}_{1,t}$.

First we extend L to all monomials $\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'}$ with degree at most 2 in $\mathbf{x}, \overline{\mathbf{x}}$ and degree at most $2t$ in $\mathbf{y}, \overline{\mathbf{y}}$. For this we set

$$L(\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'}) := 0 \text{ if } |\alpha| \neq |\alpha'| \text{ or } |\beta| \neq |\beta'|. \tag{77}$$

Otherwise, $|\alpha + \alpha'|, |\beta + \beta'|$ are even and we set

$$L(\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'}) := L(\|\mathbf{x}\|^{2-|\alpha+\alpha'|} \|\mathbf{y}\|^{2t-|\beta+\beta'|} \mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'}).$$

By construction, L is Hermitian (since $\rho_{1,t}$ is Hermitian) and L vanishes on $\mathcal{I}(1-\|\mathbf{x}\|^2, 1-\|\mathbf{y}\|^2)_{2,2t}$.

It remains to show that $L \geq 0$ on $\Sigma_{2,2t}$. In view of Lemma 22, it suffices to show that $L \geq 0$ on $\Sigma_{=2,=2t}$, or, equivalently, that the moment matrix $M_{=1,=t}(L)$, indexed by monomials $\mathbf{x}^\alpha \overline{\mathbf{x}}^{\alpha'} \mathbf{y}^\beta \overline{\mathbf{y}}^{\beta'}$ with $|\alpha + \alpha'| = 1$ and $|\beta + \beta'| = t$, is positive semidefinite. In view of (77) the matrix $M_{=1,=t}(L)$ is block-diagonal with respect to the partition of its index set according to the value of $(|\alpha|, |\beta|)$, i.e., according to the partition of $I^{=1,=t} = \bigcup_{r,s} I_{r,s}^{=1,=t}$ defined in Eq. (65) with $-1 \leq r \leq 1, -t \leq s \leq t$ and $r \equiv 1, s \equiv t$ modulo 2. So we are left with the task of showing that all diagonal blocks $M_{=1,=t}(L)[I_{r,s}^{=1,=t}]$ are positive semidefinite. For this we use Eq. (67) to obtain that

$$M_{=1,=t}(L)[I_{r,s}^{=1,=t}] \succeq 0 \iff L(\mathbf{x}\mathbf{x}^{*\otimes(\frac{r+1}{2})} \otimes \overline{\mathbf{x}\mathbf{x}}^{*\otimes(\frac{1-r}{2})} \otimes \mathbf{y}\mathbf{y}^{*\otimes(\frac{t+s}{2})} \otimes \overline{\mathbf{y}\mathbf{y}}^{*\otimes(\frac{t-s}{2})}) \succeq 0.$$

This holds for all r, s such that $-1 \leq r \leq 1, -t \leq s \leq t, r \equiv 1, s \equiv t$ modulo 2 if and only if

$$L(\mathbf{x}\mathbf{x}^* \otimes \mathbf{y}\mathbf{y}^{*\otimes(t-s')} \otimes \overline{\mathbf{y}\mathbf{y}}^{*\otimes s'}) \succeq 0, \quad L(\overline{\mathbf{x}\mathbf{x}}^* \otimes \mathbf{y}\mathbf{y}^{*\otimes(t-s')} \otimes \overline{\mathbf{y}\mathbf{y}}^{*\otimes s'}) \succeq 0$$

for all $s' \in \{0\} \cup [t]$ (setting $s' = \frac{t-s}{2}$). In view of Eq. (76) we obtain that

$$L(\mathbf{x}\mathbf{x}^* \otimes \mathbf{y}\mathbf{y}^{*\otimes(t-s')} \otimes \overline{\mathbf{y}\mathbf{y}}^{*\otimes s'}) = I_A \otimes I_B^{\otimes(t-s')} \otimes \mathsf{T}_B^{\otimes s'}(\rho_{1,t}) \tag{78}$$

and, since L is Hermitian,

$$\begin{aligned}
L(\overline{\mathbf{x}\mathbf{x}}^* \otimes \mathbf{y}\mathbf{y}^{*\otimes(t-s')} \otimes \overline{\mathbf{y}\mathbf{y}}^{*\otimes s'}) &= \overline{L(\mathbf{x}\mathbf{x}^* \otimes \overline{\mathbf{y}\mathbf{y}}^{*\otimes(t-s')} \otimes \mathbf{y}\mathbf{y}^{*\otimes s'})} \\
&= \overline{I_A \otimes \mathsf{T}_B^{\otimes(t-s')} \otimes I_B^{\otimes s'}(\rho_{1,t})}.
\end{aligned}$$

Therefore, the positive semidefiniteness of all the diagonal blocks composing the matrix $M_{=1,=t}(L)$ follows from the PPT condition (75) combined with the symmetry condition (74) and the fact that the conjugate of a Hermitian positive semidefinite matrix remains positive semidefinite. □

Remark 30. Note that it follows from relation (78) in the above proof that, if $\rho = L(\mathbf{x}\mathbf{x}^* \otimes \mathbf{y}\mathbf{y}^*)$ where the linear functional L satisfies $L \geq 0$ on $\Sigma_{2,2}$, then ρ satisfies the PPT condition (75). In particular this implies that the PPT condition is contained in the definition of the parameter $\xi_t^{\mathrm{sep}}(\rho)$: if the program (43) defining $\xi_t^{\mathrm{sep}}(\rho)$ is feasible (for $t \geq 2$), then ρ satisfies the PPT condition.

Declaration of competing interest

There is no declaration of competing interest.

Appendix A. Deriving the complex results from their real analogs

In this appendix we show how the proofs of Theorems 8 and 9 can be obtained from their real versions in [47,54,10]. We begin with recalling in Section A.1 the links between the main properties of the complex objects introduced in the paper and their real analogs. Then we give the proof of Theorem 8 in Section A.2 and of Theorem 9 in Section A.3.

A.1. Preliminaries on changing variables from complex to real

Vectors and matrices. Throughout we set $\mathbf{i} = \sqrt{-1} \in \mathbb{C}$. Then any complex scalar $x \in \mathbb{C}$ can be written (uniquely) as $x = x_{\mathrm{Re}} + \mathbf{i}\, x_{\mathrm{Im}}$, where $x_{\mathrm{Re}} := \mathrm{Re}(x)$ and $x_{\mathrm{Im}} := \mathrm{Im}(x)$ denote, respectively, the real and imaginary parts of x. This notation extends to vectors and matrices by letting the maps $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ act entrywise. Any vector $x \in \mathbb{C}^n$ can be written $x = x_{\mathrm{Re}} + \mathbf{i}\, x_{\mathrm{Im}}$ with $x_{\mathrm{Re}} := \mathrm{Re}(x), x_{\mathrm{Im}} := \mathrm{Im}(x) \in \mathbb{R}^n$. This gives a bijection

$$\phi : \mathbb{C}^n \to \mathbb{R}^n \times \mathbb{R}^n \; ; \; \mathbf{x} \mapsto (\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}). \tag{79}$$

Similarly, for a complex matrix $G \in \mathbb{C}^{m \times m'}$ set $G_{\mathrm{Re}} := \mathrm{Re}(G), G_{\mathrm{Im}} := \mathrm{Im}(G) \in \mathbb{R}^{m \times m'}$ and define the $2m \times 2m'$ real matrix

$$G^{\mathbb{R}} := \begin{bmatrix} G_{\mathrm{Re}} & -G_{\mathrm{Im}} \\ G_{\mathrm{Im}} & G_{\mathrm{Re}} \end{bmatrix}. \tag{80}$$

Then $G \in \mathbb{C}^{m \times m}$ is Hermitian, i.e., $G^* = G$, if and only if $G_{\mathrm{Re}} = G_{\mathrm{Re}}^T$ and $G_{\mathrm{Im}}^T = -G_{\mathrm{Im}}$. Moreover, for $G \in \mathbb{C}^{m \times m}$ Hermitian and $w \in \mathbb{C}^m$ we have the identity

$$w^* G w = (w_{\mathrm{Re}} - \mathbf{i}\, w_{\mathrm{Im}})^T (G_{\mathrm{Re}} + \mathbf{i}\, G_{\mathrm{Im}})(w_{\mathrm{Re}} + \mathbf{i}\, w_{\mathrm{Im}}) = \begin{bmatrix} w_{\mathrm{Re}}^T & w_{\mathrm{Im}}^T \end{bmatrix} \begin{bmatrix} G_{\mathrm{Re}} & -G_{\mathrm{Im}} \\ G_{\mathrm{Im}} & G_{\mathrm{Re}} \end{bmatrix} \begin{bmatrix} w_{\mathrm{Re}} \\ w_{\mathrm{Im}} \end{bmatrix}, \tag{81}$$

which implies the well-known equivalence

$$G \succeq 0 \iff G^{\mathbb{R}} = \begin{bmatrix} G_{\mathrm{Re}} & -G_{\mathrm{Im}} \\ G_{\mathrm{Im}} & G_{\mathrm{Re}} \end{bmatrix} \succeq 0.$$

Polynomials. Polynomials in $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ with complex variables $\mathbf{x} \in \mathbb{C}^n$ can be transformed into polynomials in $\mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}]$ with real variables $\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}} \in \mathbb{R}^n$, via the change of variables $\mathbf{x} = \mathbf{x}_{\mathrm{Re}} + \mathbf{i}\, \mathbf{x}_{\mathrm{Im}}$. In this way, any $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ corresponds to a unique pair of real polynomials

$$p_{\mathrm{Re}}(\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}) := \mathrm{Re}(p(\mathbf{x}_{\mathrm{Re}} + \mathbf{i}\, \mathbf{x}_{\mathrm{Im}}, \mathbf{x}_{\mathrm{Re}} - \mathbf{i}\, \mathbf{x}_{\mathrm{Im}})) \in \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}],$$

$$p_{\mathrm{Im}}(\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}) := \mathrm{Im}(p(\mathbf{x}_{\mathrm{Re}} + \mathbf{i}\, \mathbf{x}_{\mathrm{Im}}, \mathbf{x}_{\mathrm{Re}} - \mathbf{i}\, \mathbf{x}_{\mathrm{Im}})) \in \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}]$$

satisfying the identity

$$p(\mathbf{x}, \overline{\mathbf{x}}) = p(\mathbf{x}_{\mathrm{Re}} + \mathbf{i}\, \mathbf{x}_{\mathrm{Im}}, \mathbf{x}_{\mathrm{Re}} - \mathbf{i}\, \mathbf{x}_{\mathrm{Im}}) = p_{\mathrm{Re}}(\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}) + \mathbf{i}\, p_{\mathrm{Im}}(\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}). \tag{82}$$

Note that the degrees are preserved: $\deg_{\mathbf{x}, \overline{\mathbf{x}}}(p) = \max\{\deg_{\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}}(p_{\mathrm{Re}}), \deg_{\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}}(p_{\mathrm{Im}})\}$. A polynomial p is Hermitian, i.e., $\overline{p} = p$, if and only if $p_{\mathrm{Im}} = 0$. Hence, the map

$$\mathrm{Re} : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h \to \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}] \; ; \; p(\mathbf{x}, \overline{\mathbf{x}}) \mapsto p_{\mathrm{Re}}(\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}) \tag{83}$$

is injective. This map is also surjective: take any $f \in \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}]$ and define the polynomial $p(\mathbf{x}, \overline{\mathbf{x}}) := f(\frac{\mathbf{x} + \overline{\mathbf{x}}}{2}, \frac{\mathbf{x} - \overline{\mathbf{x}}}{2\mathbf{i}}) \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$, then p is Hermitian and satisfies $f = p_{\mathrm{Re}}$. Finally, since any $p\overline{p}$ is Hermitian we have

$$\mathrm{Re}(p\overline{p}) = p_{\mathrm{Re}}^2 + p_{\mathrm{Im}}^2.$$

Hence sums of Hermitian squares in $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ are mapped to real sums of squares in $\mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}]$ and vice versa.

Polynomial matrices. For vectors and matrices with polynomial entries in $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$, the maps $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ act entrywise. Additionally, for a polynomial matrix $G \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^{m \times m'}$, we can define the real polynomial matrix $G^{\mathbb{R}} \in \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}]^{2m \times 2m'}$ using relation (80), where $G_{\mathrm{Re}}, G_{\mathrm{Im}}$ are defined entrywise: if $G = (G_{ij})$ then $G_{\mathrm{Re}} = ((G_{ij})_{\mathrm{Re}})$ and $G_{\mathrm{Im}} = ((G_{ij})_{\mathrm{Im}})$. Then G is Hermitian if and only if $G^{\mathbb{R}}$ is symmetric and as we next observe this correspondence extends to sums of squares.

Lemma 31. *Let $G \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^{m \times m}$ be a polynomial matrix and let $G^{\mathbb{R}} \in \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}]^{2m \times 2m}$ be the corresponding real polynomial matrix defined via (80). Then G is a Hermitian SoS-polynomial matrix if and only if $G^{\mathbb{R}}$ is a (real) SoS-polynomial matrix.*

Proof. Assume G is a Hermitian SoS-polynomial matrix. Let $G = UU^*$ with $U \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^{m \times k}$. Applying the change of variables from complex to real we get

$$
\begin{aligned}
G(\mathbf{x}, \overline{\mathbf{x}}) &= G_{\mathrm{Re}}(\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}) + \mathbf{i}\, G_{\mathrm{Im}}(\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}) \\
&= U(\mathbf{x}_{\mathrm{Re}} + \mathbf{i}\, \mathbf{x}_{\mathrm{Im}}, \mathbf{x}_{\mathrm{Re}} - \mathbf{i}\, \mathbf{x}_{\mathrm{Im}}) U^*(\mathbf{x}_{\mathrm{Re}} + \mathbf{i}\, \mathbf{x}_{\mathrm{Im}}, \mathbf{x}_{\mathrm{Re}} - \mathbf{i}\, \mathbf{x}_{\mathrm{Im}}) \\
&= (U_{\mathrm{Re}} + \mathbf{i}\, U_{\mathrm{Im}})(U_{\mathrm{Re}}^T - \mathbf{i}\, U_{\mathrm{Im}}^T) = U_{\mathrm{Re}} U_{\mathrm{Re}}^T + U_{\mathrm{Im}} U_{\mathrm{Im}}^T + \mathbf{i}\left(U_{\mathrm{Im}} U_{\mathrm{Re}}^T - U_{\mathrm{Re}} U_{\mathrm{Im}}^T \right).
\end{aligned}
$$

This implies $G_{\mathrm{Re}} = U_{\mathrm{Re}} U_{\mathrm{Re}}^T + U_{\mathrm{Im}} U_{\mathrm{Im}}^T$ and $G_{\mathrm{Im}} = U_{\mathrm{Im}} U_{\mathrm{Re}}^T - U_{\mathrm{Re}} U_{\mathrm{Im}}^T$, and thus

$$
\begin{aligned}
G^{\mathbb{R}} &:= \begin{bmatrix} G_{\mathrm{Re}} & -G_{\mathrm{Im}} \\ G_{\mathrm{Im}} & G_{\mathrm{Re}} \end{bmatrix} = \begin{bmatrix} U_{\mathrm{Re}} U_{\mathrm{Re}}^T + U_{\mathrm{Im}} U_{\mathrm{Im}}^T & -(U_{\mathrm{Im}} U_{\mathrm{Re}}^T - U_{\mathrm{Re}} U_{\mathrm{Im}}^T) \\ U_{\mathrm{Im}} U_{\mathrm{Re}}^T - U_{\mathrm{Re}} U_{\mathrm{Im}}^T & U_{\mathrm{Re}} U_{\mathrm{Re}}^T + U_{\mathrm{Im}} U_{\mathrm{Im}}^T \end{bmatrix} \\
&= \begin{bmatrix} U_{\mathrm{Re}} & -U_{\mathrm{Im}} \\ U_{\mathrm{Im}} & U_{\mathrm{Re}} \end{bmatrix} \begin{bmatrix} U_{\mathrm{Re}}^T & U_{\mathrm{Im}}^T \\ -U_{\mathrm{Im}}^T & U_{\mathrm{Re}}^T \end{bmatrix} =: U^{\mathbb{R}}(U^{\mathbb{R}})^T,
\end{aligned}
$$

which shows $G^{\mathbb{R}}$ is an SoS-polynomial matrix. The converse result follows from retracing the above steps. $\square$

Quadratic modules. Given a set $S \subseteq \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h$ of Hermitian polynomials we define its real analog by applying the map $\mathrm{Re}(\cdot)$ from (83) elementwise to the set S and set

$$
S_{\mathrm{Re}} := \mathrm{Re}(S) = \{p_{\mathrm{Re}} : p \in S\} \subseteq \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}]. \tag{84}
$$

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

$$
S^G := \{w^* G w : w \in \mathbb{C}^m\} \subseteq \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h
$$

and, for the corresponding real symmetric matrix $G^{\mathbb{R}} \in \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}]^{2m \times 2m}$ defined via (80), we define the set of real polynomials

$$
S^{G^{\mathbb{R}}} := \{(w_{\mathrm{Re}}, w_{\mathrm{Im}})^T G^{\mathbb{R}}(w_{\mathrm{Re}}, w_{\mathrm{Im}}) : w \in \mathbb{C}^m\} \subseteq \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}].
$$

These two sets satisfy the expected correspondence:

$$S^{G^{\mathbb{R}}} = \text{Re}(S^G),$$

since, in view of relation (81), we have $\text{Re}(w^*Gw) = (w_{\text{Re}}, w_{\text{Im}})^T G^{\mathbb{R}}(w_{\text{Re}}, w_{\text{Im}})$ for all $w \in \mathbb{C}^m$.

This correspondence extends to the (real part of the) truncated complex quadratic module $\mathcal{M}(S)_{2t}$ generated by $S \subseteq \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h$ and the truncated real quadratic module generated by the corresponding set $S_{\text{Re}} \subseteq \mathbb{R}[\mathbf{x}_{\text{Re}}, \mathbf{x}_{\text{Im}}]$ via (84), which is denoted here as $\mathcal{M}^{\mathbb{R}}(S_{\text{Re}})_{2t}$ and defined by

$$\mathcal{M}^{\mathbb{R}}(S_{\text{Re}})_{2t} := \text{cone}\{g_{\text{Re}}f^2 : f \in \mathbb{R}[\mathbf{x}_{\text{Re}}, \mathbf{x}_{\text{Im}}] \, , \, g \in S, \quad \deg(g_{\text{Re}}f^2) \leq 2t\}.$$

Namely, we have

$$\text{Re}(\mathcal{M}(S)_{2t}) = \mathcal{M}^{\mathbb{R}}(S_{\text{Re}})_{2t}.$$

Indeed we have $\text{Re}(gp\overline{p}) = g_{\text{Re}}(p_{\text{Re}}^2 + p_{\text{Im}}^2)$ and the next relation, collected for further reference:

$$gp\overline{p} \in \mathcal{M}_{2t}(S) \iff \text{Re}(gp\overline{p}) = g_{\text{Re}}(p_{\text{Re}}^2 + p_{\text{Im}}^2) \in \mathcal{M}_{2t}^{\mathbb{R}}(S_{\text{Re}}) \text{ for all } p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}], \, g \in S.$$
$$(85)$$

Lemma 32. *For any $S \subseteq \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h$, the (complex) quadratic module $\mathcal{M}(S)$ is Archimedean if and only if the real quadratic module $\mathcal{M}^{\mathbb{R}}(S_{\text{Re}})$ is Archimedean.*

Proof. Directly from Eq. (85) since, for any scalar $R \in \mathbb{R}$, $R^2 - \mathbf{x}^*\mathbf{x} \in \mathcal{M}(S)$ if and only if $\text{Re}(R^2 - \mathbf{x}^*\mathbf{x}) = R^2 - \mathbf{x}_{\text{Re}}^T\mathbf{x}_{\text{Re}} + \mathbf{x}_{\text{Im}}^T\mathbf{x}_{\text{Im}} \in \mathcal{M}^{\mathbb{R}}(S_{\text{Re}})$. $\quad\square$

Positivity domains and measures. There is a natural correspondence between the complex positivity domain $\mathscr{D}(S)$ of a set $S \subseteq \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h$ and the real positivity domain of the corresponding set $S_{\text{Re}} \subseteq \mathbb{R}[\mathbf{x}_{\text{Re}}, \mathbf{x}_{\text{Im}}]$, which is denoted $\mathscr{D}^{\mathbb{R}}(S_{\text{Re}})$ and defined by

$$\mathscr{D}^{\mathbb{R}}(S_{\text{Re}}) := \{(w_{\text{Re}}, w_{\text{Im}}) \in \mathbb{R}^{2n} : g_{\text{Re}}(w_{\text{Re}}, w_{\text{Im}}) \geq 0 \, \forall \, g \in S\}.$$

Indeed, in view of Eq. (83) and using the complex/real bijection map ϕ from (79), we have

$$\mathscr{D}^{\mathbb{R}}(S_{\text{Re}}) = \phi(\mathscr{D}(S)).$$

Given a measure $\mu^{\mathbb{R}}$ on $\mathbb{R}^{2n}$ we define the complex measure μ on $\mathbb{C}^n$ as $\mu = \mu^{\mathbb{R}} \circ \phi$, the push-forward of $\mu^{\mathbb{R}}$ by the map ϕ^{-1}, so that

$$\int\limits_{\mathbb{C}^n} p(\mathbf{x})d\mu = \int\limits_{\mathbb{R}^{2n}} p \circ \phi^{-1}(\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}})d\mu^{\mathbb{R}}$$

$$= \int\limits_{\mathbb{R}^{2n}} p_{\mathrm{Re}}(\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}})d\mu^{\mathbb{R}} + \mathbf{i} \int\limits_{\mathbb{R}^{2n}} p_{\mathrm{Im}}(\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}})d\mu^{\mathbb{R}} \tag{86}$$

for any $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ (using (82)).

If $\mu^{\mathbb{R}}$ is supported by $\mathscr{D}^{\mathbb{R}}(S_{\mathrm{Re}})$ (i.e., $\mu^{\mathbb{R}}(\mathbb{R}^{2n} \setminus \mathscr{D}^{\mathbb{R}}(S_{\mathrm{Re}})) = 0$), then μ is supported by $\mathscr{D}(S)$ (i.e., $\mu(\mathbb{C}^n \setminus \mathscr{D}(S)) = 0$). This follows from the fact that $\phi(\mathbb{C}^n \setminus \mathscr{D}(S)) = \mathbb{R}^{2n} \setminus \mathscr{D}^{\mathbb{R}}(S_{\mathrm{Re}})$.

Linear functionals. For a linear functional $L : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathbb{C}$ we have $L(p) = \mathrm{Re}(L(p)) + \mathbf{i}\,\mathrm{Im}(L(p))$ for all $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$. Recall that L is Hermitian if $\overline{L(p)} = L(\overline{p})$. For any Hermitian L, we can define a real linear functional $L^{\mathbb{R}} : \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}] \to \mathbb{R}$ by

$$L^{\mathbb{R}}(f) := L\big(f\big(\frac{\mathbf{x} + \overline{\mathbf{x}}}{2}, \frac{\mathbf{x} - \overline{\mathbf{x}}}{2\mathbf{i}}\big)\big) \quad \text{for any} \quad f \in \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}]. \tag{87}$$

For a Hermitian polynomial $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h$, by Eq. (83) we have $p_{\mathrm{Re}}(\frac{\mathbf{x} + \overline{\mathbf{x}}}{2}, \frac{\mathbf{x} - \overline{\mathbf{x}}}{2\mathbf{i}}) = p(\mathbf{x}, \overline{\mathbf{x}})$ and thus

$$L(p) = L^{\mathbb{R}}(p_{\mathrm{Re}}) \quad \text{for any } p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h. \tag{88}$$

Then for any $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ we have

$$L(p) = L\big(p_{\mathrm{Re}}\big(\frac{\mathbf{x} + \overline{\mathbf{x}}}{2}, \frac{\mathbf{x} - \overline{\mathbf{x}}}{2\mathbf{i}}\big)\big) + \mathbf{i}\, L\big(p_{\mathrm{Im}}\big(\frac{\mathbf{x} + \overline{\mathbf{x}}}{2}, \frac{\mathbf{x} - \overline{\mathbf{x}}}{2\mathbf{i}}\big)\big) = L^{\mathbb{R}}(p_{\mathrm{Re}}) + \mathbf{i}\, L^{\mathbb{R}}(p_{\mathrm{Im}}). \tag{89}$$

In particular, we have $L(p\overline{p}) = L^{\mathbb{R}}(p_{\mathrm{Re}}^2 + p_{\mathrm{Im}}^2)$ for any $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$. This implies that L is positive (on Hermitian sums of squares) if and only if $L^{\mathbb{R}}$ is positive (on real sums of squares). Since $\mathrm{Re}(\cdot)$ preserves degrees, the restriction of $L^{\mathbb{R}}$ to $\mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}]_t$ corresponds to the restriction of L to $\mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_t$. This gives the following correspondence for truncated quadratic modules.

Lemma 33. *Given $S \subseteq \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h$, a Hermitian linear map $L \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathbb{C}$, the corresponding set $S_{\mathrm{Re}} \subseteq \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}]$ and the corresponding real linear map $L^{\mathbb{R}} \in \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}] \to \mathbb{R}$ we have*

$$L \geq 0 \text{ on } \mathcal{M}(S)_{2t} \iff L^{\mathbb{R}} \geq 0 \text{ on } \mathcal{M}^{\mathbb{R}}(S_{\mathrm{Re}})_{2t} \quad \text{for any } t \in \mathbb{N} \cup \{\infty\}.$$

Proof. This follows from the linearity of L and $L^{\mathbb{R}}$ since, by Eq. (85), $gp\overline{p} \in \mathcal{M}(S)$ if and only if $\mathrm{Re}(gp\overline{p}) = g_{\mathrm{Re}}(p_{\mathrm{Re}}^2 + p_{\mathrm{Im}}^2) \in \mathcal{M}^{\mathbb{R}}(S_{\mathrm{Re}})$ and, by Eq. (88), $L(gp\overline{p}) = L^{\mathbb{R}}(g_{\mathrm{Re}}(p_{\mathrm{Re}}^2 + p_{\mathrm{Im}}^2))$. $\square$

Finally note that an evaluation functional L_w at a point $w \in \mathbb{C}^d$ corresponds to the evaluation functional $L_{(w_{\mathrm{Re}}, w_{\mathrm{Im}})}$ at the point $(w_{\mathrm{Re}}, w_{\mathrm{Im}}) \in \mathbb{R}^{2d}$ since, for every $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$, we have

$$L_w(p) = p(w, \overline{w}) = p_{\mathrm{Re}}(w_{\mathrm{Re}}, w_{\mathrm{Im}}) + \mathbf{i}\, p_{\mathrm{Im}}(w_{\mathrm{Re}}, w_{\mathrm{Im}})$$
$$= L^{\mathbb{R}}_{(w_{\mathrm{Re}}, w_{\mathrm{Im}})}(p_{\mathrm{Re}}) + \mathbf{i}\, L^{\mathbb{R}}_{(w_{\mathrm{Re}}, w_{\mathrm{Im}})}(p_{\mathrm{Im}}).$$

Matrix-valued linear functionals. Consider a complex matrix-valued linear map

$$\mathcal{L} : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathbb{C}^{m \times m}, \quad p \mapsto \mathcal{L}(p) := \big(L_{ij}(p)\big)_{i,j \in [m]},$$

where each $L_{ij} : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathbb{C}$ is scalar-valued. Then $\mathcal{L}$ is Hermitian if and only if, for all $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$, we have $\mathcal{L}(\overline{p}) = \mathcal{L}(p)^*$, i.e.,

$$\Big(\mathrm{Re}(L_{ij}(\overline{p})) + \mathbf{i}\, \mathrm{Im}(L_{ij}(\overline{p}))\Big)_{i,j=1}^m = \Big(\mathrm{Re}(L_{ji}(p)) - \mathbf{i}\, \mathrm{Im}(L_{ji}(p))\Big)_{i,j=1}^m$$

or, equivalently, $\mathrm{Re}(L_{i,j}(\overline{p})) = \mathrm{Re}(L_{j,i}(p))$ and $\mathrm{Im}(L_{i,j}(\overline{p})) = -\mathrm{Im}(L_{j,i}(p))$ for all $i, j \in [m]$. This implies that if $\mathcal{L}$ is Hermitian and p is Hermitian then the complex matrix $\mathcal{L}(p)$ is Hermitian.

Assume $\mathcal{L}$ is Hermitian. Then we define the real matrix-valued linear functional

$$\mathcal{L}^{\mathbb{R}} : \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}] \to \mathbb{R}^{2m \times 2m}, \quad f \in \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}] \mapsto \mathcal{L}^{\mathbb{R}}(f)$$

$$\mathcal{L}^{\mathbb{R}}(f) := \left(\mathcal{L}\Big(f\big(\tfrac{\mathbf{x}+\overline{\mathbf{x}}}{2}, \tfrac{\mathbf{x}-\overline{\mathbf{x}}}{2i}\big)\Big)\right)^{\mathbb{R}} = \begin{bmatrix} \mathrm{Re}(\mathcal{L}(f(\tfrac{\mathbf{x}+\overline{\mathbf{x}}}{2}, \tfrac{\mathbf{x}-\overline{\mathbf{x}}}{2\mathbf{i}}))) & -\mathrm{Im}(\mathcal{L}(f(\tfrac{\mathbf{x}+\overline{\mathbf{x}}}{2}, \tfrac{\mathbf{x}-\overline{\mathbf{x}}}{2\mathbf{i}}))) \\ \mathrm{Im}(\mathcal{L}(f(\tfrac{\mathbf{x}+\overline{\mathbf{x}}}{2}, \tfrac{\mathbf{x}-\overline{\mathbf{x}}}{2\mathbf{i}}))) & \mathrm{Re}(\mathcal{L}(f(\tfrac{\mathbf{x}+\overline{\mathbf{x}}}{2}, \tfrac{\mathbf{x}-\overline{\mathbf{x}}}{2\mathbf{i}}))) \end{bmatrix}. \tag{90}$$

Since $f(\tfrac{\mathbf{x}+\overline{\mathbf{x}}}{2}, \tfrac{\mathbf{x}-\overline{\mathbf{x}}}{2\mathbf{i}})$ is Hermitian it follows that $-\mathrm{Im}(\mathcal{L}(f(\tfrac{\mathbf{x}+\overline{\mathbf{x}}}{2}, \tfrac{\mathbf{x}-\overline{\mathbf{x}}}{2\mathbf{i}}))) = \mathrm{Im}(\mathcal{L}(f(\tfrac{\mathbf{x}+\overline{\mathbf{x}}}{2}, \tfrac{\mathbf{x}-\overline{\mathbf{x}}}{2\mathbf{i}})))^T$. Hence $\mathcal{L}^{\mathbb{R}}$ takes its values in the cone $\mathcal{S}^{2m}$ of symmetric matrices.

Lemma 34. *Given a Hermitian linear map $\mathcal{L} : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathbb{C}^{m \times m}$ and the corresponding map $\mathcal{L}^{\mathbb{R}}$ from (90), $g \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h$ and $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ we have the following equivalence*

$$\mathcal{L}(gp\overline{p}) \succeq 0 \iff \mathcal{L}^{\mathbb{R}}(g_{\mathrm{Re}}(p_{\mathrm{Re}}^2 + p_{\mathrm{Im}}^2)) \succeq 0.$$

Proof. From Eqs. (80), (83) and (90) we obtain that

$$0 \preceq \mathcal{L}(gp\overline{p}) \iff 0 \preceq \begin{bmatrix} \mathrm{Re}(\mathcal{L}(gp\overline{p})) & -\mathrm{Im}(\mathcal{L}(gp\overline{p})) \\ \mathrm{Im}(\mathcal{L}(gp\overline{p})) & \mathrm{Re}(\mathcal{L}(gp\overline{p})) \end{bmatrix} = \mathcal{L}^{\mathbb{R}}(g_{\mathrm{Re}}(p_{\mathrm{Re}}^2 + p_{\mathrm{Im}}^2)),$$

because $gp\overline{p}$ is Hermitian. $\square$

Corollary 35. *Given $S \subseteq \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h$, a Hermitian linear map $\mathcal{L} : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathbb{C}^{m \times m}$ is positive on $\mathcal{M}(S)$ if and only if the corresponding real linear map $\mathcal{L}^{\mathbb{R}}$ from (90) is positive on $\mathcal{M}^{\mathbb{R}}(S_{\mathrm{Re}})$.*

A.2. Deriving Theorem 8 from its real analog

We can now derive Theorem 8, which we stated for complex polynomials, from the following well-known results for real polynomials from [47] and [54].

Theorem 36. *Let $S \subseteq \mathbb{R}[\mathbf{x}]$ such that $\mathcal{M}^{\mathbb{R}}(S)$ is Archimedean and let $L : \mathbb{R}[\mathbf{x}] \to \mathbb{R}$ be a linear map that is nonnegative on $\mathcal{M}^{\mathbb{R}}(S)$. Then the following holds.*

(i) *(Putinar [47]) There exists a measure $\mu^{\mathbb{R}}$ that is supported on $\mathscr{D}^{\mathbb{R}}(S)$, the positivity domain of S defined by*

$$\mathscr{D}^{\mathbb{R}}(S) = \{a \in \mathbb{R}^n : g(a) \geq 0 \text{ for all } g \in S\},$$

such that $L(f) = \int f d\mu$ for all $f \in \mathbb{R}[\mathbf{x}]$.

(ii) *(Tchakaloff [54]) For any integer $k \in \mathbb{N}$ there exists a linear map $\widehat{L} : \mathbb{R}[\mathbf{x}] \to \mathbb{R}$ such that*

$$\widehat{L}(f) = L(f) \ \forall f \in \mathbb{R}[\mathbf{x}]_k \quad \text{and} \quad \widehat{L} = \sum_{\ell=1}^{K} \lambda_\ell L_{a^\ell}$$

for some integer $K \geq 1$, scalars $\lambda_1, \dots, \lambda_K > 0$ and vectors $a^1, \dots, a^K \in \mathscr{D}^{\mathbb{R}}(S)$.

We now indicate how to derive Theorem 8 from Theorem 36. For this consider $S \subseteq \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]^h$ and a linear map $L : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathbb{C}$. Assume $\mathcal{M}(S)$ is Archimedean and $L \geq 0$ on $\mathcal{M}(S)$. We consider the set $S_{\mathrm{Re}} \subseteq \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}]$ of real polynomials defined via (84) and the associated linear map $L^{\mathbb{R}} : \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}] \to \mathbb{R}$ defined via (87). By Lemma 32 the quadratic module $\mathcal{M}^{\mathbb{R}}(S_{\mathrm{Re}})$ is Archimedean and, by Lemma 33, $L^{\mathbb{R}} \geq 0$ on $\mathcal{M}^{\mathbb{R}}(S_{\mathrm{Re}})$. Hence we can apply Theorem 36 to S_{Re} and $L^{\mathbb{R}}$.

By Theorem 36 (i), there exists a (real) measure $\mu^{\mathbb{R}}$ that is supported by $\mathscr{D}^{\mathbb{R}}(S_{\mathrm{Re}})$ and satisfies $L^{\mathbb{R}}(f) = \int f d\mu^{\mathbb{R}}$ for all $f \in \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}]$. Consider the (complex) measure μ defined by relation (86), which is therefore supported by the set $\mathscr{D}(S)$. We claim that μ is a representing measure for L. Indeed, for $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$, using (89) we have

$$L(p) = L^{\mathbb{R}}(p_{\mathrm{Re}}) + \mathbf{i}\, L^{\mathbb{R}}(p_{\mathrm{Im}}) = \int p_{\mathrm{Re}} d\mu^{\mathbb{R}} + \mathbf{i} \int p_{\mathrm{Im}} d\mu^{\mathbb{R}} = \int p d\mu.$$

This completes the proof of Theorem 8 (i). We now derive its part (ii).

Fix an integer $k \in \mathbb{N}$. By Theorem 36 (ii), there exists $\widehat{L} : \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}] \to \mathbb{R}$ such that $\widehat{L}(f) = L^{\mathbb{R}}(f)$ for all $f \in \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}]_k$ and $\widehat{L} = \sum_{\ell=1}^{K} \lambda_\ell L_{a^\ell}$ for some $K \in \mathbb{N}$, $\lambda_\ell > 0$ and $a^\ell \in \mathscr{D}^{\mathbb{R}}(S_{\mathrm{Re}})$. Define the complex linear map $\widetilde{L} : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathbb{C}$ by $\widetilde{L}(p) := \widehat{L}(p_{\mathrm{Re}}) + \mathbf{i}\,\widehat{L}(p_{\mathrm{Im}})$ for any $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$. Then, in view of (89), we have $\widetilde{L}(p) = L(p)$ for any $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]_k$. For each $\ell \in [K]$ let w^ℓ be the complex vector such that $(w^\ell_{\mathrm{Re}}, w^\ell_{\mathrm{Im}}) = a^\ell$. Then each w^ℓ belongs to $\mathscr{D}(S)$ and we have

$$\widetilde{L}(p) = \widehat{L}(p_{\mathrm{Re}}) + \mathbf{i}\,\widehat{L}(p_{\mathrm{Im}}) = \sum_{\ell} \lambda_{\ell}(p_{\mathrm{Re}}(a^{\ell}) + \mathbf{i}\,p_{\mathrm{Im}}(a^{\ell})) = \sum_{\ell} \lambda_{\ell} p(w^{\ell}),$$

which shows $\widetilde{L} = \sum_{\ell} \lambda_{\ell} L_{w^{\ell}}$, and thus concludes the proof of Theorem 8 (ii).

A.3. Deriving Theorem 9 from its real analog

In this section we prove the implication (ii) $\implies$ (i) in Theorem 9 from its real analog in [10], which we restate below for convenience.

Theorem 37. *[10] Let $S \subseteq \mathbb{R}[\mathbf{x}]$ be a set of polynomials such that the quadratic module $\mathcal{M}^{\mathbb{R}}(S)$ is Archimedean. Let $\mathcal{L} : \mathbb{R}[\mathbf{x}] \to \mathcal{S}^m$ be a matrix-valued linear functional that is positive on $\mathcal{M}^{\mathbb{R}}(S)$, i.e., $\mathcal{L}(gf^2) \succeq 0$ for all $g \in S \cup \{1\}$ and $f \in \mathbb{R}[\mathbf{x}]$. Then there exists a matrix-valued measure μ that is supported on $\mathscr{D}^{\mathbb{R}}(S)$ and takes values in the cone $\mathcal{S}^m_+$ of $m \times m$ positive semidefinite matrices such that $\mathcal{L}(f) = \int f d\mu$ for all $f \in \mathbb{R}[\mathbf{x}]$.*

We now indicate how to derive the implication (ii) $\implies$ (i) in Theorem 9 from Theorem 37. For this let $S \subseteq \mathbb{C}[\mathbf{x},\overline{\mathbf{x}}]^h$ such that $\mathcal{M}(S)$ is Archimedean and let $\mathcal{L} : \mathbb{C}[\mathbf{x},\overline{\mathbf{x}}] \to \mathcal{H}^m$ which is Hermitian and satisfies $\mathcal{L}(gp\overline{p}) \succeq 0$ for all $g \in S$ and $p \in \mathbb{C}[\mathbf{x},\overline{\mathbf{x}}]$. Then the set $S_{\mathrm{Re}} \subseteq \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}]$ from (84) has a Archimedean quadratic module $\mathcal{M}^{\mathbb{R}}(S_{\mathrm{Re}})$ by Lemma 32. Consider the linear map $\mathcal{L}^{\mathbb{R}} : \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}] \to \mathcal{S}^{2m}$ defined via (90). Then, by Lemma 34, $\mathcal{L}^{\mathbb{R}}$ is positive on $\mathcal{M}^{\mathbb{R}}(S_{\mathrm{Re}})$. Hence we can apply Theorem 37 and conclude that $\mathcal{L}^{\mathbb{R}}$ has a representing measure $\mu^{\mathbb{R}}$, which is supported on $\mathscr{D}^{\mathbb{R}}(S_{\mathrm{Re}})$ and takes values in the cone $\mathcal{S}^{2m}_+$. We will now construct a (complex) measure μ, which represents $\mathcal{L}$ and is supported on the set $\mathscr{D}(S)$, using the following two claims.

Claim 38. *The map $\mathcal{L}^{\mathbb{R}}$ takes values in the set $\mathcal{W} := \left\{ \begin{bmatrix} A & B^T \\ B & C \end{bmatrix} \in \mathcal{S}^{2m} : A = C,\ B^T = -B \right\}$.*

Proof. Since $\mathcal{L}^{\mathbb{R}}$ takes values in $\mathcal{S}^{2m}$ it has the following block-form

$$\mathcal{L}^{\mathbb{R}} = \begin{bmatrix} \mathcal{L}^{\mathbb{R}}_{11} & (\mathcal{L}^{\mathbb{R}}_{21})^T \\ \mathcal{L}^{\mathbb{R}}_{21} & \mathcal{L}^{\mathbb{R}}_{22} \end{bmatrix},$$

where $\mathcal{L}^{\mathbb{R}}_{11}$ and $\mathcal{L}^{\mathbb{R}}_{22}$ take values in $\mathcal{S}^m$, $\mathcal{L}^{\mathbb{R}}_{21}$ takes values in $\mathbb{R}^{m \times m}$ and, by construction,

$$\mathcal{L}^{\mathbb{R}}_{11}(f) = \mathcal{L}^{\mathbb{R}}_{22}(f) = \mathrm{Re}\big(\mathcal{L}\big(f\big(\tfrac{\mathbf{x}+\overline{\mathbf{x}}}{2}, \tfrac{\mathbf{x}-\overline{\mathbf{x}}}{2\mathbf{i}}\big)\big)\big), \quad \mathcal{L}^{\mathbb{R}}_{21}(f) = \mathrm{Im}\big(\mathcal{L}\big(f\big(\tfrac{\mathbf{x}+\overline{\mathbf{x}}}{2}, \tfrac{\mathbf{x}-\overline{\mathbf{x}}}{2\mathbf{i}}\big)\big)\big)$$

(91)

and $(\mathcal{L}^{\mathbb{R}}_{21}(f))^T = -\mathrm{Im}(\mathcal{L}(f(\tfrac{\mathbf{x}+\overline{\mathbf{x}}}{2}, \tfrac{\mathbf{x}-\overline{\mathbf{x}}}{2\mathbf{i}})))$ for any $f \in \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}]$. Hence $(\mathcal{L}^{\mathbb{R}}_{21})^T = -\mathcal{L}^{\mathbb{R}}_{21}$ and thus $\mathcal{L}^{\mathbb{R}}$ takes values in $\mathcal{W}$ as desired. $\quad\square$

Claim 39. *Without loss of generality we may assume the measure $\mu^{\mathbb{R}}$ takes values in $\mathcal{W} \cap \mathcal{S}_+^{2m}$.*

Proof. We can write the measure $\mu^{\mathbb{R}}$ in block-form as

$$\mu^{\mathbb{R}} = \begin{bmatrix} \mu_{11}^{\mathbb{R}} & (\mu_{21}^{\mathbb{R}})^T \\ \mu_{21}^{\mathbb{R}} & \mu_{22}^{\mathbb{R}} \end{bmatrix},$$

where each block is a measure taking its values in $\mathbb{R}^{m \times m}$. Then we can define the following new matrix-valued measure

$$\mu' := \frac{1}{2} \begin{bmatrix} \mu_{11}^{\mathbb{R}} + \mu_{22}^{\mathbb{R}} & -(\mu_{21}^{\mathbb{R}} - (\mu_{21}^{\mathbb{R}})^T) \\ (\mu_{21}^{\mathbb{R}} - (\mu_{21}^{\mathbb{R}})^T) & \mu_{11}^{\mathbb{R}} + \mu_{22}^{\mathbb{R}} \end{bmatrix} =: \begin{bmatrix} \mu_{11}' & -(\mu_{21}')^T \\ \mu_{21}' & \mu_{11}' \end{bmatrix}.$$

First, by construction, μ' takes its values in the set $\mathcal{W}$. Second, μ' takes its values in $\mathcal{S}_+^{2m}$. Indeed, by Theorem 37, $\mu^{\mathbb{R}}$ takes values in $\mathcal{S}_+^{2m}$ and we have

$$\mu' = \frac{1}{2} \begin{bmatrix} 0 & -I_m \\ I_m & 0 \end{bmatrix} \mu^{\mathbb{R}} \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix} + \frac{1}{2} \mu^{\mathbb{R}}.$$

Finally, μ' also represents $\mathcal{L}^{\mathbb{R}}$. Indeed, for any $f \in \mathbb{R}[\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}}]$ we have $\mathcal{L}_{11}^{\mathbb{R}}(f) = \mathcal{L}_{22}^{\mathbb{R}}(f)$ and $-\mathcal{L}_{21}^{\mathbb{R}}(f) = (\mathcal{L}_{21}^{\mathbb{R}}(f))^T$ by Claim 38. This implies

$$\mathcal{L}_{11}^{\mathbb{R}}(f) = \frac{1}{2}(\mathcal{L}_{11}^{\mathbb{R}}(f) + \mathcal{L}_{22}^{\mathbb{R}}(f)) = \frac{1}{2} \int f(d\mu_{11}^{\mathbb{R}} + d\mu_{22}^{\mathbb{R}}) = \int f d\mu_{11}',$$

$$\mathcal{L}_{21}^{\mathbb{R}}(f) = \frac{1}{2}(\mathcal{L}_{21}^{\mathbb{R}}(f) - (\mathcal{L}_{21}^{\mathbb{R}}(f))^T) = \frac{1}{2} \int f(d\mu_{21}^{\mathbb{R}} - d(\mu_{21}^{\mathbb{R}})^T) = \int f d\mu_{21}',$$

and thus $\mathcal{L}^{\mathbb{R}}(f) = \int f d\mu'$. Therefore we may replace the measure $\mu^{\mathbb{R}}$ by μ', which shows the claim. $\square$

We now define the complex measure μ by setting

$$d\mu := d\mu_{11}^{\mathbb{R}} \circ \phi + \mathbf{i} \, d\mu_{21}^{\mathbb{R}} \circ \phi, \tag{92}$$

where ϕ is the complex/real bijection in Eq. (79). So μ takes values in $\mathbb{C}^{m \times m}$. As shown above in Claim 39, $\mu^{\mathbb{R}}$ takes values in the set $\mathcal{W} \cap \mathcal{S}_+^{2m}$. Hence, in view of Eq. (80), we can conclude that μ takes its values in $\mathcal{H}_+^m$. In addition, as $\mu^{\mathbb{R}}$ is supported by $\mathscr{D}^{\mathbb{R}}(S_{\mathrm{Re}})$, it follows that μ is supported by $\mathscr{D}(S)$. Finally, we verify that μ represents $\mathcal{L}$. Indeed, for any $p \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$, using (91) we obtain

$$\mathcal{L}(p) = \mathcal{L}\big(p_{\mathrm{Re}}\big(\frac{\mathbf{x} + \overline{\mathbf{x}}}{2}, \frac{\mathbf{x} - \overline{\mathbf{x}}}{2i}\big)\big) + \mathbf{i} \, \mathcal{L}\big(p_{\mathrm{Im}}\big(\frac{\mathbf{x} + \overline{\mathbf{x}}}{2}, \frac{\mathbf{x} - \overline{\mathbf{x}}}{2i}\big)\big)$$

$$= (\mathcal{L}_{11}^{\mathbb{R}}(p_{\mathrm{Re}}(\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}})) + \mathbf{i} \, \mathcal{L}_{21}^{\mathbb{R}}(p_{\mathrm{Re}}(\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}})))$$

$$+ \mathbf{i} \left(\mathcal{L}_{11}^{\mathbb{R}}(p_{\mathrm{Im}}(\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}})) + \mathbf{i} \, \mathcal{L}_{21}^{\mathbb{R}}(p_{\mathrm{Im}}(\mathbf{x}_{\mathrm{Re}}, \mathbf{x}_{\mathrm{Im}})) \right)$$

$$= \int p_{\mathrm{Re}} d\mu_{11}^{\mathbb{R}} + \mathbf{i} \int p_{\mathrm{Re}} d\mu_{21}^{\mathbb{R}} + \mathbf{i} \int p_{\mathrm{Im}} d\mu_{11}^{\mathbb{R}} - \int p_{\mathrm{Im}} d\mu_{21}^{\mathbb{R}}$$

$$= \int (p_{\mathrm{Re}} + \mathbf{i} \, p_{\mathrm{Im}}) d\mu_{11}^{\mathbb{R}} + \mathbf{i} \int (p_{\mathrm{Re}} + \mathbf{i} \, p_{\mathrm{Im}}) d\mu_{21}^{\mathbb{R}}$$

$$= \int (p_{\mathrm{Re}} + \mathbf{i} \, p_{\mathrm{Im}})(d\mu_{11}^{\mathbb{R}} + \mathbf{i} \, d\mu_{21}^{\mathbb{R}})$$

$$= \int p d\mu.$$

This concludes the proof of the implication (ii) $\implies$ (i) in Theorem 9.

References

[1] B. Podolsky, A. Einstein, N. Rosem, Can quantum-mechanical description of physical reality be considered complete?, Phys. Rev. 47 (10) (1935) 777–780.
[2] A. Berman, N. Shaked-Monderer, Completely Positive Matrices, World Scientific Publishing Company Pte Limited, 2003.
[3] J. Bezanson, A. Edelman, S. Karpinski, V.B. Shah, Julia: a fresh approach to numerical computing, SIAM Rev. 59 (1) (2017) 65–98.
[4] I. Bomze, W. Schachinger, R. Ullrich, From seven to eleven: completely positive matrices with high cp-rank, Linear Algebra Appl. 459 (2014) 208–221.
[5] S. Burer, On the copositive representation of binary and continuous nonconvex quadratic programs, Math. Program. 120 (2009) 479–495.
[6] C.M. Caves, C.A. Fuchs, R. Schack, Unknown quantum states: the quantum de Finetti representation, J. Math. Phys. 43 (2002) 4537.
[7] L. Chen, D.Z. Dokovic, Qubit-qudit states with positive partial transpose, Phys. Rev. A 86 (6) (2012).
[8] L. Chen, D.Z. Dokovic, Dimensions, lengths, and separability in finite-dimensional quantum systems, J. Math. Phys. 54 (2) (2013) 022201.
[9] M.-D. Choi, Positive linear maps, in: Operator Algebras and Applications, in: Proc. Sympos. Pure Math., vol. 38(2), Amer. Math. Soc., 1982, pp. 583–590.
[10] J. Cimpric, A. Zalar, Moment problems for operator polynomials, J. Math. Anal. Appl. 401 (2013) 307–316.
[11] E. de Klerk, D.V. Pasechnik, Approximation of the stability number of a graph via copositive programming, SIAM J. Optim. 12 (4) (2002) 875–892.
[12] G. de las Cuevas, T. Drescher, T. Netzer, Separability for mixed states with operator Schmidt rank two, Quantum 3 (2019) 203.
[13] G. de las Cuevas, T. Netzer, Mixed states in one spatial dimension: decompositions and correspondence with nonnegative matrices, J. Math. Phys. 61 (4) (2020) 041901.
[14] D.P. Divincenzo, B.M. Terhal, A.V. Thapliyal, Optimal decompositions of barely separable states, J. Mod. Opt. 47 (2–3) (2000) 377–385.
[15] Research Computing Documentation, Mosek — research computing documentation, 2017. [Online; accessed 5-September-2021].
[16] A.C. Doherty, P.A. Parrilo, F.M. Spedalieri, Distinguishing separable and entangled states, Phys. Rev. Lett. 88 (2002) 187904.
[17] M. Dressler, J. Nie, Z. Yang, Separability of hermitian tensors and PSD decompositions, arXiv:2011. 08132, 2020, J. Linear Multilinear Algebra (2021), https://doi.org/10.1080/03081087.2021.1965078, in press.
[18] I. Dunning, J. Huchette, M. Lubin, Jump: a modeling language for mathematical optimization, SIAM Rev. 59 (2) (2017) 295–320.
[19] K. Fang, H. Fawzi, The sum-of-squares hierarchy on the sphere and application in quantum information theory, in: Mathematical Programming, 2020.

[20] H. Fawzi, The set of separable states has no finite semidefinite representation except in dimension 2×3, Commun. Math. Phys. 386 (2021) 1319–1335.
[21] H. Fawzi, J. Gouveia, P.A. Parrilo, R. Robinson, R.R. Thomas, Positive semidefinite rank, Math. Program., Ser. B 153 (1) (2015) 133–177.
[22] H. Fawzi, P.A. Parrilo, Self-scaled bounds for atomic cone ranks: applications to nonnegative rank and cp-rank, Math. Program. 158 (1) (2016) 417–465.
[23] K. Gaterman, P. Parrilo, Symmetry groups, semidefinite programs and sums of squares, J. Pure Appl. Algebra 192 (2004) 95–128.
[24] S. Gharibian, Strong NP-hardness of the quantum separability problem, Quantum Inf. Comput. 10 (3) (2010) 343–360.
[25] N. Gillis, Nonnegative Matrix Factorization, SIAM, 2020.
[26] S. Gribling, D. de Laat, M. Laurent, Lower bounds on matrix factorization ranks via noncommutative polynomial optimization, Found. Comput. Math. 19 (5) (2019) 1013–1070.
[27] L. Gurvits, Classical deterministic complexity of Edmonds' problem and quantum entanglement, in: Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing, 2003, pp. 10–19.
[28] D.W. Hadwin, Dilations and Hahn decompositions for linear maps, Can. J. Math. 33 (4) (1981) 826–839.
[29] A.W. Harrow, A. Natarajan, X. Wu, An improved semidefinite programming hierarchy for testing entanglement, Commun. Math. Phys. 352 (3) (2017) 881–904.
[30] J. Håstad, Tensor rank is NP-complete, J. Algorithms 11 (4) (1990) 644–654.
[31] M. Horodecki, P. Horodecki, R. Horodecki, Separability of mixed states: necessary and sufficient conditions, Phys. Lett. A 223 (1–2) (1996) 1–8.
[32] P. Horodecki, Separability criterion and inseparable mixed states with positive partial transposition, Phys. Lett. A 232 (5) (1997) 333–339.
[33] D.P. Kimsey, An operator-valued generalization of Tchakaloff's theorem, J. Funct. Anal. 266 (2014) 1170–1184.
[34] T.G. Kolda, B.W. Bader, Tensor decompositions and applications, SIAM Rev. 51 (3) (2009) 455–500.
[35] J.B. Lasserre, Global optimization with polynomials and the problem of moments, SIAM J. Optim. 11 (2001) 796–817.
[36] J.B. Lasserre, Moments, Positive Polynomials and Their Applications, Imperial College Press, 2009.
[37] M. Laurent, Sums of squares, moment matrices and optimization over polynomials, in: Emerging Applications of Algebraic Geometry, in: IMA Volumes in Mathematics and its Applications, vol. 149, Springer, New York, 2009, pp. 157–270.
[38] Y. Li, G. Ni, Separability discrimination and decomposition of m-partite quantum mixed states, Phys. Rev. A 102 (1) (2020).
[39] M. Nakata, A numerical evaluation of highly accurate multiple-precision arithmetic version of semidefinite programming solver: SDPA-GMP, -QD and -DD, in: 2010 IEEE International Symposium on Computer-Aided Control System Design, 2010, pp. 29–34.
[40] M. Navascués, M. Owari, M.B. Plenio, Power of symmetric extensions for entanglement detection, Phys. Rev. A 80 (2009) 052306.
[41] J. Nie, The $\mathcal{A}$-truncated K-moment problem, Found. Comput. Math. 14 (2014) 1243–1276.
[42] J. Nie, Symmetric tensor nuclear norms, SIAM J. Appl. Algebra Geom. 1 (1) (2017) 599–625.
[43] J. Nie, X. Zhang, Positive maps and separable matrices, SIAM J. Optim. 26 (2) (2016) 1236–1256.
[44] M.N. Nielsen, I.L. Chuang, Quantum Computation and Quantum Information, Cambridge University Press, 2000.
[45] V. Paulsen, Completely Bounded Maps and Operator Algebras, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2003.
[46] A. Peres, Separability criterion for density matrices, Phys. Rev. Lett. 77 (1996) 1413–1415.
[47] M. Putinar, Positive polynomials on compact semi-algebraic sets, Indiana Univ. Math. J. 42 (1993) 969–984.
[48] Y. Shitov, A short proof that NMF is NP-hard, arXiv:1605.04000, 2016.
[49] Y. Shitov, A universality theorem for nonnegative matrix factorizations, arXiv:1606.09068, 2016.
[50] Y. Shitov, The complexity of positive semidefinite matrix factorization, SIAM J. Optim. 27 (3) (2017) 1898–1909.
[51] M. Dutour Sikirić, A. Schürmann, F. Vallentin, A simplex algorithm for rational CP-factorization, Math. Program. 187 (2021) 25–45.
[52] J. Sponsel, M. Dür, Factorization and cutting planes for completely positive matrices by copositive projection, Math. Program. 143 (1–2) (2014) 211–229.

[53] G. Tang, P. Shah, Guaranteed tensor decomposition: a moment approach, in: Proceedings of the 32nd International Conference on Machine Learning, ICML-15, 2015, pp. 1491–1500.

[54] V. Tchakaloff, Formules de cubature mécaniques à coéfficients non négatifs, Bull. Sci. Math. 81 (1957) 123–134.

[55] A. Uhlmann, Entropy and optimal decompositions of states relative to a maximal commutative subalgebra, Open Syst. Inf. Dyn. 5 (1998) 209–228.

[56] S.A. Vavasis, On the complexity of nonnegative matrix factorization, SIAM J. Optim. 20 (3) (2009) 1364–1377.

[57] J. Wang, V. Magron, J.B. Lasserre, N.H. Anh Mai, CS-TSSOS: correlative and term sparsity for large-scale polynomial optimization, arXiv:2005.02828, 2020.

[58] J. Watrous, The Theory of Quantum Information, Cambridge University Press, 2018.

[59] S.L. Woronowicz, Positive maps of low dimensional matrix algebras, Rep. Math. Phys. 10 (2) (1976) 165–183.

[60] M. Yamashita, K. Fujisawa, M. Kojima, Implementation and evaluation of SDPA 6.0 (semidefinite programming algorithm 6.0), Optim. Methods Softw. (2003) 491–505.

[61] M. Yamashita, K. Fujisawa, K. Nakata, M. Nakata, M. Fukuda, K. Kobayashi, K. Goto, A high-performance software package for semidefinite programs: SDPA7, Technical report, 2010.