Smallest disentangling state spaces for general entangled bipartite quantum states

Hussain Anwar,[1] Sania Jevtic,[2, 3] Oliver Rudolph,[4] and Shashank Virmani[5, *]

[1]*Department of Physics, Imperial College London, Prince Consort Road, London SW7 2AZ, UK.*
[2]*Department of Mathematics, Huxley Building, Imperial College, London SW7 2AZ, UK.*
[3]*Institut für Theoretische Physik, Leibniz Universität Hannover, Appelstraße 2, 30167 Hannover, Germany.*
[4]*Leonardo da Vinci Gymnasium, Im Spitzerfeld 25, 69151 Neckargemünd,*
Germany. & Hector-Seminar, Waldhoferstr. 100, 69123 Heidelberg, Germany.
[5]*Department of Mathematical Sciences, Brunel University, Uxbridge, Middlesex UB8 3PH, UK.*

Entangled quantum states can be given a separable decomposition if we relax the restriction that the local operators be quantum states. Motivated by the construction of classical simulations and local hidden variable models, we construct 'smallest' local sets of operators that achieve this. In other words, given an arbitrary bipartite quantum state we construct convex sets of local operators that allow for a separable decomposition, but that cannot be made smaller while continuing to do so. We then consider two further variants of the problem where the local state spaces are required to contain the local quantum states, and obtain solutions for a variety of cases including a region of pure states around the maximally entangled state. The methods involve calculating certain forms of cross norm. Two of the variants of the problem have a strong relationship to theorems on ensemble decompositions of positive operators, and our results thereby give those theorems an added interpretation. The results generalise those obtained in our previous work on this topic [New J. Phys. **17**, 093047 (2015)].

I.   BACKGROUND AND OVERVIEW

In recent years a variety of work has examined the nature of quantum entanglement by considering quantum systems as examples of more general structures. In the study of generalised probabilistic theories [1], for example, various axioms are postulated (such as no instant signalling) for reasonable physical theories, without necessarily demanding that the theory has an underlying Hilbert space or operator space structure. Other works have considered theories with an operator structure, but where the operators are allowed for instance to have negative eigenvalues [2], so long as the measurements are restricted in a way that leads to valid probabilities. The work of [3] was one of the first to consider the consequences of this line of thinking for entangled quantum states. A feature emerging from these previous constructions is the idea that if the available observables we have are restricted, then, in principle, our perception of the entanglement in a given system may change. In other articles this idea motivated us to consider whether this can help design new local hidden variable (LHV) models [4] and classical simulation methods for entangled quantum states measured in restricted ways [5–7].

One of the observations underlying the present work is that by allowing local state spaces to contain non-quantum operators, an entangled quantum state may in fact admit a separable decomposition that can be considered to be *unentangled* if the measurements available are restricted. This motivates the question: if we want to use such generalised separable decompositions for the construction of LHV models or for efficient classical simulation algorithms, what such decompositions are useful? This question motivated a problem that we considered in a previous work [8]: what are the *smallest* sets of operators such that a given entangled quantum state admits a separable decomposition? In this work we will generalise the results of [8] to a much wider family of states.

Structure—This paper is structured as follows. In section II we explain and motivate the four problems (Problems 1,2,2*,3) that we consider in this work. In section III we present some norms that we use to characterise how small a given set of states is. In section IV we present a summary of our main results, which constitute three Theorems. Theorem 1 characterises all solutions to Problem 1 for all bipartite states for a subset of the norms that we consider, and it is proven in section V. Theorem 2 solves Problem 2 for all bipartite states, and is proven in section VI. Theorem 3 presents solutions to problems 2* and 3 for certain families of bipartite state (including the pure states in a region around the maximally entangled state), and is proven in section VII. Those not interested in proofs may skip sections V,VI, and VII.

---

*Electronic address: shashank.virmani@brunel.ac.uk

II. GENERALISED SEPARABILITY AND PROBLEMS CONSIDERED

In this work we will be concerned with a bipartite setting in which there are two systems A and B with arbitrary finite local Hilbert space dimensions. Their joint Hilbert space is $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ and their quantum state is described by a density operator ρ_{AB} with $\rho_{AB} \geq 0$ and $\mathrm{tr}(\rho_{AB}) = 1$. According to the standard definition of quantum entanglement, a state ρ_{AB} is said to be *separable* if

$$\rho_{AB} = \sum_j p_j \rho_j^A \otimes \rho_j^B, \tag{1}$$

where the local quantum operators ρ_j^A and ρ_j^B are positive semidefinite with $\mathrm{tr}(\rho_j^A) = \mathrm{tr}(\rho_j^B) = 1$, and p_j forms a probability distribution, otherwise it is said to be *entangled* [9]. In this work we will be interested in generalising these definitions, and so we will instead use the expressions *quantum*-separable or *quantum*-entangled. The reason for this is that if one relaxes the requirement that ρ_j^A and ρ_j^B be positive operators, then one can indeed construct 'separable' decompositions for any bipartite state ρ_{AB}. An illustrative example can be constructed from the operator-Schmidt decomposition of any given quantum state ρ_{AB}:

$$\rho_{AB} = \sum_{i=1}^{D} \lambda_i X_i \otimes Y_i, \tag{2}$$

where X_i and Y_i are local orthonormal operator bases (i.e. $\mathrm{tr}(X_i^\dagger X_j) = \mathrm{tr}(Y_i^\dagger Y_j) = \delta_{ij}$), the Schmidt coefficients $\lambda_i > 0$ are real positive numbers, and the operator-Schmidt rank D satisfies $D \leq \dim(\mathcal{H}_A)^2, \dim(\mathcal{H}_B)^2$. Defining $\lambda := \sum_i \lambda_i$, and $\widetilde{X}_i := \sqrt{\lambda}X_i, \widetilde{Y}_i := \sqrt{\lambda}Y_i$, we can rewrite this decomposition in the separable form:

$$\rho_{AB} = \sum_{i=1}^{D} \frac{\lambda_i}{\lambda} \widetilde{X}_i \otimes \widetilde{Y}_i. \tag{3}$$

As the coefficients λ_i/λ form a probability distribution, Eq. (3) is a separable decomposition in terms of typically non-positive operators $\widetilde{X}_i, \widetilde{Y}_j$. Although Eq. (3) may seem like a trivial rewriting of the Schmidt decomposition, we shall show later that this decomposition has an interesting property:

> If we consider the local operator sets formed from the convex hulls of the $\widetilde{X}_i$ and $\widetilde{Y}_i$ for each subsystem separately, then the resulting sets cannot be made *smaller* while continuing to admit a separable decomposition for ρ_{AB}.

Hence we say that the decomposition Eq. (3) is an example of a 'smallest' separable decomposition for the state ρ_{AB}. Finding such decompositions for various definitions of 'smallest' will be the central aim of this article.

The motivation for looking for such smallest separable decompositions arises from their utility in the construction of LHV models or efficient classical simulation algorithms (see, for example, [7, 8]). The basic principle that underpins such constructions is the idea that a non-positive operator (a *non-physical* quantum state—i.e. cannot be *prepared* by a physical process) can in fact be considered a valid description of quantum system when the available measurements are restricted. Variations of this idea have a long history, going back at least as far as the Wigner function [10]. For convenience we briefly review why this is.

Consider an experiment in which a quantum state ρ undergoes a physical process $\mathcal{P}$ consisting of a transformation $L : \rho \mapsto L(\rho)$, followed by a measurement yielding an outcome corresponding to a measurement operator $M \geq 0$. The whole process, denoted by $\mathcal{P}(L, M)$, occurs with probability $\mathrm{tr}(ML(\rho))/\mathrm{tr}(L(\rho))$. In such situations, it would be instructive to decompose ρ into a probabilistic ensemble of positive density matrices ρ_i, each prepared with classical probability p_i, such that $\rho = \sum_i p_i \rho_i$. Due to linearity, one can analyse the experiment in terms of the process $\mathcal{P}(L, M)$ acting individually on each of the ρ_i operators in the decomposition. This approach usually assumes that each ρ_i is a positive operator, however, it works even when the operators ρ_i that appear in the decomposition are *not* positive, provided that the processes $\mathcal{P}(L, M)$ being considered are restricted so that all the ρ_i satisfy $\mathrm{tr}(ML(\rho_i)) \geq 0$ and $\mathrm{tr}(L(\rho_i)) > 0$. These properties ensure that the classical manipulation of the probabilities p_i follows the usual Bayesian rules, and hence there are no obstacles to considering the system as a mixture of the otherwise non-physical operators ρ_i [25].

We refer to the properties $\mathrm{tr}(ML(\rho_i)) \geq 0$ and $\mathrm{tr}(L(\rho_i)) > 0$ as notions of *generalised* positivity with respect to $\mathcal{P}(L, M)$. In this paper the choice of L and M will be either arbitrary or clear from the context, and so for brevity we will usually just use the term *generalised positivity* without explicitly mentioning the transformations L or measurements M.

LHV construction and motivating examples for Bell states— One of the motivations for looking for such generalised separable decompositions is that they can provide LHV models for measurements for which the local operators are positive. For convenience we now explain this standard argument (as used, for example in the use of Wigner functions to construct LHV models [11]). Suppose that

$$\rho_{AB} = \sum_i q_i A_i \otimes B_i,$$

is a separable decomposition of a quantum state ρ_{AB}, where A_i and B_i are generalised-positive for local positive-operator valued measures (POVMs) $\mathcal{M} = \{M_i | M_i \geq 0, \sum_i M_i = \mathbb{1}\}$ and $\mathcal{N} = \{N_i | N_i \geq 0, \sum_i N_i = \mathbb{1}\}$. The q_i must be positive but need not be normalised. The probability of obtaining outcomes $M_k \in \mathcal{M}$ and $N_l \in \mathcal{N}$ is

$$
\begin{aligned}
\text{tr}(\rho_{AB} M_k \otimes N_l) &= \sum_i q_i \text{tr}(A_i M_k)\text{tr}(B_i N_l), \\
&= \sum_i q_i a(k|i)b(l|i),
\end{aligned}
\tag{4}
$$

where $a(k|i) := \text{tr}(A_i M_k) \geq 0$ and $b(l|i) := \text{tr}(B_i N_l) \geq 0$. The completeness of the POVMs implies that $\sum_k a(k|i) = \text{tr}(A_i)$ and $\sum_l b(l|i) = \text{tr}(B_i)$. From this we see that if operators A_i or B_i are traceless, then the quantities $a(k|i)$ or $b(l|i)$ will be all zero, and will not contribute to the sum in Eq. (4). Hence we may ignore those terms and rewrite Eq. (4) as:

$$
\text{tr}(\rho_{AB} M_k \otimes N_l) = \sum_{i:\text{tr}\{A_i\},\text{tr}\{B_i\}>0} q_i \text{tr}(A_i)\text{tr}(B_i)\frac{a(k|i)}{\text{tr}(A_i)}\frac{b(l|i)}{\text{tr}(B_i)} = \sum_{i:\text{tr}\{A_i\},\text{tr}\{B_i\}>0} p_i r(k|i)s(l|i)
\tag{5}
$$

where $p_i := q_i \text{tr}(A_i)\text{tr}(B_i), r(k|i) := a(k|i)/\text{tr}(A_i), s(l|i) := b(l|i)/\text{tr}(B_i)$. From the fact that $\text{tr}(\rho_{AB}) = 1$ it follows that p_i is a normalised probability distribution, and $r(k|i)$ and $s(l|i)$ are normalised conditional probability distributions. Hence we see that Eq. (5) supplies a LHV model for the probability distribution $\text{tr}(\rho_{AB} M_k \otimes N_l)$. Note that as demonstrated by the Werner states [9] in the case of projective quantum measurements, a lack of generalised separability for a given class of measurements does not necessarily imply that a state is non-local w.r.t. those measurements (although for a small enough number of measurements and measurement outcomes the two are equivalent [5]).

In order for a separable decomposition to be useful for constructing LHV models in this way, we would like the local operators appearing in the decomposition to be generalised positive for as large a set of measurements as possible. In this respect there are some separable decompositions that are more useful than others. We will illustrate this point, as well as the theorems that we present later, in reference to three example separable decompositions (all of which are actually examples of 3) of the density matrix corresponding to the canonical Bell state $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$:

1. It is well known from the stabiliser formalism that the Bell state can be written a in terms of the separable decomposition:

$$
|\phi^+\rangle\langle\phi^+| = \frac{1}{4}I \otimes I + \frac{1}{4}\sigma_x \otimes \sigma_x + \frac{1}{4}\sigma_y \otimes \sigma_y^T + \frac{1}{4}\sigma_z \otimes \sigma_z,
\tag{6}
$$

   where $\{\sigma_x, \sigma_y, \sigma_z\}$ are the conventional Pauli operators. While the operators $\sigma_x, \sigma_y, \sigma_z$ can be generalised positive for a large class of measurement operators (e.g. measurement operators with Bloch vectors in the positive orthant), this decomposition is useless for constructing LHV models, as the traceless operators $\sigma_x, \sigma_y, \sigma_z$ cannot be generalised positive for a *complete* POVM other than the trivial identity measurement.

2. Consider now a second separable decomposition:

$$
|\phi^+\rangle\langle\phi^+| = \frac{1}{4}(\sqrt{2}|0\rangle\langle0|) \otimes (\sqrt{2}|0\rangle\langle0|) + \frac{1}{4}(\sqrt{2}|1\rangle\langle1|) \otimes (\sqrt{2}|1\rangle\langle1|) + \frac{1}{4}\sigma_x \otimes \sigma_x + \frac{1}{4}\sigma_y \otimes \sigma_y^T.
\tag{7}
$$

   This decomposition better than the first, but still only yields a LHV model only for trivial local σ_z observable measurements. By following the procedure given above, the LHV terms for the σ_z measurement are

$$
p_i = \left(\frac{1}{2}, \frac{1}{2}, 0, 0\right),
$$
$$
r(k = +1|i) = s(k = +1|i) = (1, 0, 0, 0),
$$
$$
r(k = -1|i) = s(k = -1|i) = (0, 1, 0, 0).
$$

3. Consider now another decomposition of the same Bell state that occurs in the study of discrete Wigner functions ([8, 12]):

$$|\phi^+\rangle\langle\phi^+| = \frac{1}{4}\sum_{i=1}^{4} W_i \otimes W_i^T, \tag{8}$$

where the $\{W_i\}$ form a subset of the *phase-point operators* [12], defined as

$$W_1 = \frac{1}{2}(\mathbb{1} + \sigma_x + \sigma_y + \sigma_z),$$

$$W_2 = \frac{1}{2}(\mathbb{1} + \sigma_x - \sigma_y - \sigma_z),$$

$$W_3 = \frac{1}{2}(\mathbb{1} - \sigma_x + \sigma_y - \sigma_z),$$

$$W_4 = \frac{1}{2}(\mathbb{1} - \sigma_z - \sigma_y + \sigma_z).$$

The decomposition in Eq. (8) is much better than the first two - it yields a LHV model for a much larger set $\mathcal{S}$ of local measurement operators. For system A, this set turns out to be the measurement operators with Bloch vectors in $\mathcal{Q} \cap \mathcal{T}$, where $\mathcal{Q}$ is the set of single qubit operators (i.e. the Bloch sphere) and $\mathcal{T}$ is the tetrahedron whose four vertices are the W_i (see Ref. [8]). For system B, the set of local measurement operators is the transpose of those for A. The set $\mathcal{S}$ includes all the Pauli observables. In addition, it includes local non-stabilizer POVMs such as $\{c\,|m\rangle\langle m|, \mathbb{1} - c\,|m\rangle\langle m|\}$ for small enough $c > 0$ where $|m\rangle$ is the (magic [13]) state with Bloch vector $(1, 1, 1)/\sqrt{3}$.

These examples demonstrate the value of finding a separable decomposition with local operators that are generalised positive for the largest possible classes of local measurements or processes. How may we identify decompositions with such properties? In order to address this question we will follow the strategy described in our previous work [8], which was built upon a key property implicitly shared by all notions of generalised positivity: because operators satisfying a given notion of generalised positivity form a convex set, if a set of local operators $\mathcal{V}_A$ is positive for a given notion of generalised positivity, then so is the convex hull of any subset $\mathcal{W}_A \subset \mathcal{V}_A$. This implies that if two separable decompositions exist such that the local convex *state spaces* of one decomposition are contained within those of the other decomposition, then the smaller decomposition can only be more useful. By minimising the local state spaces, for example, we can only enlarge the corresponding dual space of local measurements, thereby achieving better LHV models. This is the motivation behind the question that was posed in [8] and that we continue to explore in this work: what are the *smallest* sets of local operators such that a given entangled quantum state ρ_{AB} is separable with respect to them?

Problem variants— To make this question precise we must define the term 'smallest' more precisely. In a previous article [8] we considered three variants of the problem motivated by different definitions of the word 'smallest', and presented solutions for some states. In this article we generalise those arguments, obtaining solutions for all bipartite states for two of the problems, and for certain families of bipartite state for the other one. As a technical tool to facilitate the discussion we will also add another variant (which we call Problem 2*)—while this variant was in fact considered in [8], it was not explicitly defined. Hence there are four problems that we consider, which are:

- **Problem 1:** For a given bipartite quantum state ρ_{AB} consider all pairs of convex sets $(\mathcal{V}_A, \mathcal{V}_B)$ (which we call 'state spaces') for which ρ_{AB} is separable, and work out the minimum (in all cases considered in this paper the minimum exists) value of $\|\mathcal{V}_A\|\|\mathcal{V}_B\|$, where we define the size $\|\mathcal{V}\|$ of a state space $\mathcal{V}$ relative to some choice of norm $\|\bullet\|$ as $\|\mathcal{V}\| := \sup\{\|X\|\,|X \in \mathcal{V}\}$. In [8] this was shown to be equivalent to computing a form of cross-norm [14, 15] built from $\|\bullet\|$ (readers that are not familiar with cross norms may see Eq. (11) for a definition of the versions that we consider here). Depending upon the choice of norm, this problem may not have strong physical meaning by itself. However, we use it as a mathematical tool to solve other variants of the problem.

- **Problem 2:** For a given quantum state ρ_{AB}, consider all pairs of state spaces $(\mathcal{V}_A, \mathcal{V}_B)$ for which ρ_{AB} is separable, and identify pairs $(\mathcal{V}_A, \mathcal{V}_B)$ such that no strict subsets $(\mathcal{W}_A \subset \mathcal{V}_A, \mathcal{W}_B \subseteq \mathcal{V}_B)$ or $(\mathcal{W}_A \subseteq \mathcal{V}_A, \mathcal{W}_B \subset \mathcal{V}_B)$ exist for which ρ_{AB} is $(\mathcal{W}_A, \mathcal{W}_B)$-separable. The motivation for this problem is the fact that generalised positivity of a set implies generalised positivity of any subset, so it is in our advantage to look for the smallest separable decompositions. As shown in [8], all the separable decompositions given above for $|\phi^+\rangle\langle\phi^+|$ are solutions to Problem 2 (the fact that two of them are not of use for constructing LHV models is a motivation for the next two problems).

- **Problem 2*:** The same as Problem 2, but with the additional constraint that we identify smallest state spaces under the restriction that they must contain the *local quantum states* and they only contain operators with unit trace. For the purposes of this article the motivation for Problem 2* is primarily technical, because it turns out that the conic hull [26] of solutions to Problem 2* automatically give a solution our next problem (see [8]). However, in applications of the separable decompositions to classically efficient simulations of quantum systems (not just LHV models), the restriction to unit trace operators can be important for making sure that sampling occurs correctly [7].

- **Problem 3:** The same as Problem 2, but with the additional constraint that the state spaces $(\mathcal{V}_A, \mathcal{V}_B)$ must be convex *cones* of positive trace operators that contain the quantum states (in this case the weights appearing in the separable decomposition need only be positive—whether they form a normalised probability distribution is irrelevant as we are now dealing with cones). The motivation for this problem is explained in [8] and footnote [17].

In the next few sections we present a solutions to these problems for more general states than were considered in [8]. The key technique is the use of certain norms to characterise the local operators and the entangled state. Loosely speaking, the strategy of our arguments follows the reasoning used in [8]: by evaluating a type of cross-norm (defined later) on an entangled state we obtain an achievable lower bound on the size of the local operator spaces appearing in the decomposition. We are able to generalise [8] to much larger families of states by considering a broader family of norms than were considered there.

III.   LINEARLY TRANSFORMED 2-NORMS AND CROSS-NORMS

In this article we will use a particular class of norms on a single state space. These norms are given by a 2-norm of the output of a fixed invertible linear transformation Λ on the operator from the state space, and we denote them in the following way:

$$\|X\|_\Lambda := \|\Lambda(X)\|_2. \tag{9}$$

Here Λ is an invertible linear transformation, and $\|\Lambda(X)\|_2$ refers to the 2-norm of $\Lambda(X)$ (the 2-norm of an operator Y is defined as $\sqrt{\operatorname{tr}(Y^\dagger Y)}$). As described above this norm can induce a norm on sets $\mathcal{V}$ of operators via the definition

$$\|\mathcal{V}\|_\Lambda := \sup\{\|X\|_\Lambda | X \in \mathcal{V}\}.$$

There are two reasons why these norms are useful to us: firstly the triangle inequality is *strict* for them, i.e.

$$\|X + Y\|_\Lambda < \|X\|_\Lambda + \|Y\|_\Lambda, \tag{10}$$

unless $X = cY$ for some positive number c [18], and secondly we can calculate the resulting cross norms (see e.g. [14, 15, 19] for use in entanglement theory, or Eq. (11) for the ones we use here) explicitly in a number of cases. We note that strictness of the triangle inequality holds for a variety of other norms (such as the so-called Schatten p-norms with $1 < p < \infty$), so it is possible that our arguments could generalise as long as the resulting cross norms can be calculated.

We will use these single state space norms to construct types of cross norm on more than one state space. Given two invertible linear transformations Λ, Γ we will define the Λ, Γ cross-norm as:

$$\|\rho_{AB}\|_{\Lambda,\Gamma} := \inf\{\sum_i \|A_i\|_\Lambda \|B_i\|_\Gamma \ : \ \rho_{AB} = \sum_i A_i \otimes B_i\} \tag{11}$$

The arguments presented in [8] show that $\|\rho_{AB}\|_{\Lambda,\Gamma}$ is equivalent to the infimum value of $\|\mathcal{V}_\mathcal{A}\|_\Lambda \|\mathcal{V}_\mathcal{B}\|_\Gamma$ over all pairs of sets $\mathcal{V}_\mathcal{A}, \mathcal{V}_\mathcal{B}$ such that ρ_{AB} is $(\mathcal{V}_\mathcal{A}, \mathcal{V}_\mathcal{B})$-separable. We will use these cross norms to help us apply solutions presented in [8] for the maximally entangled state to more general settings. It will be important to our arguments that we can explicitly calculate optimal decompositions (11) in a variety of cases, thereby adding to solutions that have been computed in previous works [14, 15].

IV. SUMMARY OF RESULTS

In this section we summarize the main results of the paper. We describe the linear transformations Λ, Γ by expanding in the orthonormal basis that describes the operator-Schmidt decomposition of the state ρ_{AB} of interest. So for instance we write:

$$\Lambda(X_i) = \sum_{i,j} R_{ij} X_j, \tag{12}$$

where $R = \{R_{ij}\}$ is a $D \times D$ matrix. With this notation our main results are as follows.

Theorem 1. *(Solutions to Problem 1) Consider a bipartite quantum state ρ_{AB} with operator-Schmidt decomposition $\rho_{AB} = \sum_{i=1}^{D} s_i X_i \otimes Y_i$ with $s_i > 0$. Define $S = \text{diag}(s_i)$ and make arbitrary choices of the following: an invertible $D \times D$ positive diagonal matrix R, positive constants $c_k, k = 1, \ldots, N$, probability distribution $p_k, k = 1, \ldots, N$, and $D \times N$ isometry U. Then*

$$\rho_{AB} = \sum_{k=1}^{N} p_k A^k \otimes B^k, \tag{13}$$

with

$$
\begin{aligned}
A^k &= \sqrt{\frac{1}{p_k c_k}} \sum_{i=1}^{D} (\sqrt{S} R^{-1} U)_{ik} X_i, \\
B^k &= \sqrt{\frac{c_k}{p_k}} \sum_{i=1}^{D} (\sqrt{S} R U)^{*}_{ik} Y_i,
\end{aligned} \tag{14}
$$

gives a separable decomposition for ρ_{AB} that solves Problem 1, i.e. it achieves the minimal R, R^{-1} cross-norm. Moreover, all R, R^{-1} cross-norm achieving separable decompositions can be written in this form. By choosing $R = \sqrt{S}$, $U_{ij} = \delta_{ij}$, $c_k = p_k$, we find that the operator-Schmidt decomposition itself is a solution to Problem 1.

In the examples given for the Bell state in the introduction all the decompositions satisfy this theorem and are hence solutions to Problem 1. More generally, the operator-Schmidt decomposition in Eq. (3) satisfies this theorem.

Theorem 2. *(Solutions to Problem 2) Consider a bipartite quantum state ρ_{AB} with operator-Schmidt decomposition $\rho_{AB} = \sum_{i=1}^{D} s_i X_i \otimes Y_i$ with $s_i > 0$. Define $S = \text{diag}(s_i)$ and make arbitrary choices of the following: an invertible $D \times D$ positive diagonal matrix R, a positive constants c and $D \times D$ unitary matrix U. Then*

$$\rho_{AB} = \sum_{k=1}^{N} p_k A^k \otimes B^k, \tag{15}$$

with

$$
\begin{aligned}
A^k &= \sqrt{\frac{1}{p_k c}} \sum_{i=1}^{D} (\sqrt{S} R^{-1} U)_{ik} X_i, \\
B^k &= \sqrt{\frac{c}{p_k}} \sum_{i=1}^{D} (\sqrt{S} R U)^{*}_{ik} Y_i, \\
p_k &= \sum_{i=1}^{D} |U_{ik}|^2 \frac{s_i}{\sum_j s_j},
\end{aligned} \tag{16}
$$

gives a separable decomposition for ρ_{AB} that solves Problem 2, i.e. it achieves the minimal R, R^{-1} cross-norm and is a smallest one, in the sense that the sets:

$$
\begin{aligned}
\mathcal{V}_A &:= \text{conv}\{A^k | k = 1, \ldots, D\}, \\
\mathcal{V}_B &:= \text{conv}\{B^k | k = 1, \ldots, D\},
\end{aligned}
$$

cannot be made smaller while continuing to admit a separable decomposition for ρ_{AB}. By choosing $R = \sqrt{\frac{1}{c}} \mathbb{1}$ and $U_{ij} = \delta_{ij}$ we find that the operator-Schmidt decomposition itself is a solution to Problem 2.

Note the close parallels between this result and theorems on ensemble decompositions of density matrices (see [20–22] and Theorem 2.6 in [23]).

In the examples given for the Bell state in the introduction all decompositions are satisfy this theorem and are hence solutions to Problem 2. More generally, the operator-Schmidt decomposition in Eq. (3) satisfies this theorem and thus provides a solution to Problem 2.

Theorem 3. *(Solutions to Problem 2* and Problem 3 for some states.) Consider a bipartite quantum state that can be expressed as $\rho_{AB} = \mathcal{E} \otimes \mathcal{F}(\Psi)$, where Ψ is the canonical maximally entangled state and $\mathcal{E}, \mathcal{F}$ are invertible linear maps. Then provided $\mathcal{E}, \mathcal{F}$ satisfy the following:*

(A) a basis of operators W_i can be found such that $\operatorname{tr}(W_i^\dagger W_j) = d\delta_{ij}$ and $\operatorname{tr}(\mathcal{E}(W_i)) = 1$ and $\operatorname{tr}(\mathcal{F}(W_i^T)) = 1$,

(B) $||\mathcal{E}^{-1}(\sigma)||_2, ||\mathcal{F}^{-1}(\sigma)||_2 < \sqrt{d}$ for any local quantum states σ,

then ρ_{AB} is separable w.r.t both the unit trace convex hulls

$$\operatorname{conv}\{\mathcal{E}(W_i) \bigcup \mathcal{Q}_A\} \ , \ \operatorname{conv}\{\mathcal{F}(W_i^T) \bigcup \mathcal{Q}_B\} \tag{17}$$

where $\mathcal{Q}_A, \mathcal{Q}_B$ are the local quantum states for particles A, B respectively, and the positive trace conic hulls:

$$\operatorname{coni}\{\mathcal{E}(W_i) \bigcup \mathcal{Q}_A\} \ , \ \operatorname{coni}\{\mathcal{F}(W_i^T) \bigcup \mathcal{Q}_B\} \tag{18}$$

Moreover, these convex (resp. conic) sets contain no convex (resp. conic) strict subsets that both contain the local quantum states and maintain separability of ρ_{AB}. Hence they give a solution to Problem 2* (resp. Problem 3) for any ρ_{AB} satisfying the two conditions. Condition (A) holds for all bipartite pure states, and all bipartite states that can be obtained from Ψ by local CPTP maps. Condition (B) holds for all bipartite states satisfying $\min_k s_k > 1/d^2$, where s_k are the operator-Schmidt coefficients, which in turn holds for states in a finite region around the maximally entangled state.

In the examples given for the Bell state in the introduction, taking the convex/conic hull of the local quantum states with the local operators of the last decomposition gives an example of this solution.

V. PROOF OF THEOREM 1: CHARACTERISATION OF SOLUTIONS TO PROBLEM 1 FOR SOME CROSS NORMS.

In this section we will characterise solutions to Problem 1 for a particular subset of the cross norms defined by picking $\Gamma = \Lambda^{-1}$. We choose these norms because they enable us to treat the problem in a very similar way to the problem of finding decompositions of positive operators, and hence draw upon solutions to that problem (see [20–22], or the discussion around Theorem 2.6 in [23]). In later sections we will show that among the solutions we obtain to Problem 1 there are many that also solve Problem 2.

We begin by deriving a simple lower bound on the cross norms defined by picking $\Gamma = \Lambda^{-1}$. Consider an operator written out in an operator-Schmidt decomposition:

$$\rho_{AB} = \sum_{i=1}^{D} s_i X_i \otimes Y_i \equiv \sum_i s_i \delta_{ij} X_i \otimes Y_j. \tag{19}$$

Now consider any generalised separable decomposition of ρ_{AB}:

$$\rho_{AB} = \sum_{k=1}^{N} p_k A^k \otimes B^k. \tag{20}$$

We will expand A^k and B^k in terms of the operators X_i, Y_j in the following way: $A^k = \sum a_i^k X_i$ and $B^k = \sum (b_i^k)^* Y_i$, where a_i^k and b_i^k are expansion coefficients (the complex conjugate on the b_i^k is for later convenience). We will perform all calculations in this representation. Let us use the symbols $\boldsymbol{a}^k$ and $\boldsymbol{b}^k$ to represent vectors of these expansion coefficients a_i^k and b_i^k, respectively. It is useful to note that the the 2-norms of the local operators are given by

$\|A^k\|_2 = \|\boldsymbol{a}^k\|_2$ and $\|B^k\|_2 = \|\boldsymbol{b}^k\|_2$. For the two expressions for ρ_{AB} to be equal, we require the following three equivalent equations to hold:

$$s_i \delta_{ij} = \sum_{k=1}^{N} p_k \alpha_i^k (\beta_j^k)^*,$$

$$\Leftrightarrow \quad \sqrt{s_i} \sqrt{s_j} \delta_{ij} = \sum_{k=1}^{N} p_k \alpha_i^k (\beta_j^k)^*,$$

$$\Leftrightarrow \quad \delta_{ij} = \sum_{k=1}^{N} p_k \frac{\alpha_i^k}{\sqrt{s_i}} \left(\frac{\beta_j^k}{\sqrt{s_i}} \right)^*. \tag{21}$$

If we sum the first of these equations over $i = j$ we get:

$$\sum_i s_i = \sum_k p_k \langle \boldsymbol{b}^k, \boldsymbol{a}^k \rangle \leq \sum_k p_k \|\boldsymbol{a}^k\| \|\boldsymbol{b}^k\|, \tag{22}$$

where the inequality follows from Cauchy-Schwarz inequality. This equation gives us a lower bound for the usual cross 2-norm, which is in fact tight [14, 15]. However, the route taken to this equation can be used to give a lower bound for the larger class of norms involving linear transformations $\Lambda = \Gamma^{-1}$ on the two subsystems. To see this consider representing Λ by a linear transformation R acting on the space of vectors $\boldsymbol{a}^k$. Then we note that:

$$\langle \boldsymbol{b}^k, \boldsymbol{a}^k \rangle = \langle R^{-1} \boldsymbol{b}^k, R \boldsymbol{a}^k \rangle.$$

Given R we may rewrite Eq. (22) as:

$$\sum_i s_i = \sum_k p_k \langle R^{-1} \boldsymbol{b}^k, R \boldsymbol{a}^k \rangle,$$

$$\leq \sum_k p_k \|R \boldsymbol{a}^k\| \|R^{-1} \boldsymbol{b}^k\|. \tag{23}$$

This is the lower bound that we set out to derive. We may consider $\|R\boldsymbol{a}^k\|, \|R^{-1}\boldsymbol{b}^k\|$ as defining a cross norm calculated by first applying invertible linear transformations R to the first system, and R^{-1} to the second system, and then computing the 2-norm. In other words we are considering a cross norm of the type discussed above, but with Λ set by R and Γ set by R^{-1}.

Although the lower bound (23) will not be tight in general, we will now show that it is true if we restrict R that to be diagonal and positive. In other words we will demonstrate that for diagonal positive R the reverse inequality holds too:

$$\sum_k p_k \|R \boldsymbol{a}^k\| \|R^{-1} \boldsymbol{b}^k\| \leq \sum_i s_i. \tag{24}$$

The Cauchy-Schwarz inequality implies that the only way that the previous two equations could possibly be compatible is if we have $\|R\boldsymbol{a}^k\|_2 \|R^{-1}\boldsymbol{b}^k\|_2 = \langle R^{-1}\boldsymbol{b}^k, R\boldsymbol{a}^k \rangle$ for all k, and hence $R\boldsymbol{a}^k$ and $R^{-1}\boldsymbol{b}^k$ must be proportional:

$$R^{-1} \boldsymbol{b}^k = c_k R \boldsymbol{a}^k,$$

$$\Rightarrow \quad \boldsymbol{b}^k = c_k R^2 \boldsymbol{a}^k, \tag{25}$$

for some positive numbers c_k that may depend upon k. As we now show, it is this condition that links our problem to previously known results on the decomposition of positive operators. Let us insert Eq. (25) into one of Eqs. (21) to eliminate the b_j^k:

$$s_i \delta_{ij} = \sum_k p_k c_k a_i^k \sum_g \left(R^2 \right)_{jg} (a_g^k)^*. \tag{26}$$

Multiplying both sides by R^{-2} gives:

$$\sum_j s_i \delta_{ij} \left(R^{-2} \right)_{hj} = \sum_k p_k c_k a_i^k (a_h^k)^*,$$

$$\Rightarrow \quad \sum_j s_i \delta_{ij} \left(R^{-2} \right)_{jh} = \sum_k p_k c_k a_i^k (a_h^k)^*. \tag{27}$$

We are looking for p_k, c_k (which are positive) and a_i^k that satisfy these equations. Hence we see from Eq. (27) that we are looking for a positive decomposition of the matrix on the left-hand-side, and may thereby draw upon previous results on that problem ([20–23]). It would of course not be possible to do this if the operator on the left-hand-side were not positive, and this is why we picked R deliberately to be diagonal and positive so that the LHS is positive (although there is a slightly more general class of R that would satisfy this, it turns out that the more general argument still leads to the same decompositions). If we define a $D \times D$ matrix S as the diagonal matrix with matrix elements $s_i \delta_{ij}$, and another $D \times N$ matrix Z with matrix elements $Z_{ik} = \sqrt{p_k c_k} a_i^k$, then Eq. (27) can be rewritten more concisely as:

$$SR^{-2} = ZZ^\dagger. \tag{28}$$

The separable decomposition is represented by the matrix Z and the coefficients p_k, c_k. We would hence like to find values for these variables that solve this equation. Notice that Eq. (28) is simply the statement that Z is a square root of the positive operator SR^{-2}. Hence:

$$Z = \sqrt{S}R^{-1}U, \tag{29}$$

for an isometry U.

Let us summarise where we are. Given a positive invertible diagonal matrix R, a probability distribution p_k, and a set of positive numbers c_k, a separable decomposition satisfying the assertion (24) can exist only if Z is given by (29). Writing these equations out explicitly, we find that any separable decomposition achieving (24) must *necessarily* be of the form:

$$\rho_{AB} = \sum_{k=1}^N p_k A^k \otimes B^k,$$

$$A^k = \sum a_i^k X_i = \sqrt{\frac{1}{p_k c_k}} \sum_i (\sqrt{S}R^{-1}U)_{ik} X_i,$$

$$B^k = \sum (b_i^k)^* Y_i = \sqrt{\frac{c_k}{p_k}} \sum_i (\sqrt{S}RU)_{ik}^* Y_i. \tag{30}$$

To show that this is also sufficient, we must verify that this is indeed a separable decomposition (the fact that it satisfies (24) is by construction). A short calculation verifies that this is true for arbitrary $p_k, c_k > 0$ [24]. Hence, we see that Eq. (30) is indeed a separable decomposition for ρ_{AB}, and moreover it characterises all separable decompositions that achieve the cross-norm defined by R, R^{-1}. The c_k simply serves as a 'scale' factor that cancels out when forming $A^k \otimes B^k$—so we can use it to make the norm of any operator on Bob's side small, but only at the expense of making the corresponding operator on Alice's side large. Note that (as follows from the connections in [8]) Eqs. (30) characterise all cross-norm achieving decompositions, not just separable ones (indeed nothing in our above arguments requires $\sum_k p_k = 1$).

VI. PROOF OF THEOREM 2: DECOMPOSITIONS WITH A MINIMAL NUMBER OF OPERATORS HAVING THE SAME NORM

The previous section characterises all separable decompositions achieving the minimal value of the cross-norm defined through R and R^{-1}. We will now show that some of the decomposition of Theorem 1 also yield solutions to Problem 2. We will do this by picking decompositions with two properties: (a) for which there is the minimal number of operators ($N = D$) on each side, (b) for which the local expansion vectors on Alice's side all have the same norm, and for which the local expansion vectors on Bob's side all have the same norm. In such cases, the only operators from within the local state spaces that have a high enough norm to be part of a separable decomposition are hence the extremal points (the fact that the triangle inequality is strict means that *only* the the extremal points have a high enough norm). Moreover, as we need all D of them from each local state space, such state spaces have the property that we cannot make them smaller while preserving separability.

To complete the argument we hence need to show how to pick decompositions with the two properties (a),(b). Property (a) can be guaranteed by restricting the U appearing in the decompositions to be a square unitary matrix. This means that we only have decompositions with the minimal number D of local operators, hence satisfying (a). We must now work out which of these decompositions satisfy (b), i.e. with all operators having the same norm. The

squares of the norms of the vectors $\boldsymbol{a}^k, \boldsymbol{b}^k$ in the decomposition are given by:

$$\|R\boldsymbol{a}^k\|^2 = \frac{1}{p_k c_k} \sum_i |U_{ik}|^2 s_i,$$

$$\|R^{-1}\boldsymbol{b}^k\|^2 = \frac{c_k}{p_k} \sum_i |U_{ik}|^2 s_i. \tag{31}$$

If we wish $\|R\boldsymbol{a}^k\|^2$, say, to be same for all k then we require it to be a constant, which we denote w:

$$w = \frac{1}{p_k c_k} \sum_i |U_{ik}|^2 s_i, \tag{32}$$

i.e. we require that:

$$w p_k c_k = \sum_i |U_{ik}|^2 s_i. \tag{33}$$

Note that the $|U_{ik}|^2$ form the entries of a doubly stochastic matrix (note that it is square as $N = D$, the minimal number allowed). The value of w is fixed by summing both sides over k, i:

$$w = \frac{\sum_i s_i}{\sum_k p_k c_k}. \tag{34}$$

Hence Eq. (33) becomes:

$$\frac{p_k c_k}{\sum_l p_l c_l} = \sum_i |U_{ik}|^2 \frac{s_i}{\sum_j s_j}. \tag{35}$$

We may go through the same argument for the operators on Bob's side. The result is obtained by replacing c_k with $1/c_k$:

$$\frac{p_k/c_k}{\sum_l p_l/c_l} = \sum_i |U_{ik}|^2 \frac{s_i}{\sum_j s_j}. \tag{36}$$

These two equations tell us that the probability distributions defined by:

$$\frac{p_k c_k}{\sum_l p_l c_l} \quad ; \quad \frac{p_k/c_k}{\sum_l p_l/c_l}, \tag{37}$$

must be majorized by the probability distribution defined by:

$$\frac{s_i}{\sum_j s_j}, \tag{38}$$

with the corresponding doubly stochastic matrix (taking the latter distribution to the former distributions) given by the absolute square elements of a unitary:

$$|U_{ik}|^2. \tag{39}$$

The only way that this can happen is if the two probability distributions (37) are identical, which requires that for all k, k':

$$\frac{p_k c_k}{p_k/c_k} = \frac{p_{k'} c_{k'}}{p_{k'}/c_{k'}},$$

$$\Rightarrow c_k = c_{k'}. \tag{40}$$

So we must set $c_k = c$ where c is some positive constant, and hence in Eq. (37) the c's cancel and the two probability distributions become identical. Hence, we find that a separable decomposition:

$$\rho_{AB} = \sum_{k=1}^D p_k A^k \otimes B^k,$$

$$A^k = \sqrt{\frac{1}{p_k c}} \sum_{i=1}^D (\sqrt{S} R^{-1} U)_{ik} X_i,$$

$$B^k = \sqrt{\frac{c}{p_k}} \sum_{i=1}^D (\sqrt{S} R U)_{ik}^* Y_i, \tag{41}$$

is a smallest one if the p_k are chosen so that:

$$p_k = \sum_{i=1}^{D} |U_{ik}|^2 \frac{s_i}{\sum_j s_j}. \tag{42}$$

The above argument can also be used to construct solutions to Problem 2 under the restriction that the local state spaces are Hermitian. As a quantum state ρ_{AB} is Hermitian, the X_i and Y_i in the operator-Schmidt decomposition can be taken to be Hermitian. If we want to restrict the separable decomposition to involve only Hermitian operators, then we can do so by further restricting the unitary U to in fact be a real orthogonal matrix O.

Note the close parallels between this result and theorems on ensemble decompositions of density matrices (see [20–22] and Theorem 2.6 in [23]). By choosing $R = \sqrt{\frac{1}{c}}\mathbb{1}$ and $U_{ij} = \delta_{ij}$ we find that the operator-Schmidt decomposition itself is a solution to Problem 2, as we claimed in the introduction.

VII. PROOF OF THEOREM 3: OBTAINING SOLUTIONS TO PROBLEM 3 FROM THE MAXIMALLY ENTANGLED CASE

The solutions to Problems 1 and 2 in the previous sections do not provide solutions to Problem 2* or Problem 3. The motivation for Problem 3 is that its solution maximises the set of quantum effects for which the separable decomposition gives a LHV model: taking the conic hull of a state space does not reduce the dual set of quantum measurements for which the state space is positive; requiring the inclusion of the quantum states in the the conic hull means that the dual only contains quantum *effects* (that is, positive semidefinite operators); and requiring positive trace means that all such quantum effects can be completed to give POVMs (in the sense that for positive trace operators, positivity w.r.t. quantum POVM element M automatically implies positivity w.r.t. $I - M$ too).

In order to solve Problem 3, we must identify conic state spaces containing the local quantum states that cannot be made smaller while maintaining separability of ρ_{AB}. To do this we will first identify, for a family of bipartite quantum states that includes all pure states, local state spaces consisting of unit trace operators that solve Problem 2. We will then argue that, provided that the quantum state ρ_{AB} is correlated enough, adding the unit trace local quantum operators to these state spaces also solves Problem 2*. The conic hull of these state solutions automatically gives solutions to Problem 3, as explained in [8].

So our main aim is to hence obtain solutions to Problem 2*. One might attempt a solution to Problem 2* by taking the solutions to Problem 2 described in Theorem 2, and looking for U, p_k, c that make the operators in Eq. (41) have unit trace. Although [8] shows that was possible to do this for the maximally entangled state, we have not been able to carry out that approach for more general ρ_{AB}. However, we will see that it is possible to transform the solutions for the maximally entangled state into solutions to both Problem 2* (and hence Problem 3) for more general quantum states ρ_{AB} that possess two technical properties that we describe below.

Let us first recap the solution to Problem 2 in the case of the maximally entangled state on particles with $d = \dim(\mathcal{H}_A) = \dim(\mathcal{H}_B)$:

$$|\Psi\rangle = \frac{1}{\sqrt{d}} \sum_{i}^{d} |ii\rangle. \tag{43}$$

In [8] it was shown that *any* basis of orthogonal local operators $\{C_i | i = 1...d^2\}$ normalised to $\text{tr}(C_i C_j^\dagger) = d\delta_{ij}$ gives a smallest separable decomposition of $\Psi := |\Psi\rangle\langle\Psi|$ via:

$$\Psi = \frac{1}{d^2} \sum_{i} C_i \otimes C_i^T. \tag{44}$$

In order to see how this can be of use for solving Problem 2* for more general ρ_{AB}, consider writing ρ_{AB} in terms of local invertible transformations on Ψ, i.e.

$$\rho_{AB} = \mathcal{E} \otimes \mathcal{F}(\Psi), \tag{45}$$

where $\mathcal{E}, \mathcal{F}$ are linear invertible maps. As long as ρ_{AB} has the same operator-Schmidt rank as Ψ, i.e. d^2, then this is possible. Note that we do not require $\mathcal{E}, \mathcal{F}$ to be completely positive or trace preserving maps—they only have to be linear and invertible. Although there are many possible choices for $\mathcal{E}, \mathcal{F}$, we will explicitly consider the following

ones that are constructed from the Schmidt form of ρ_{AB}:

$$\mathcal{E}(\sigma) = \sum_j \sqrt{s_j} X_j \mathrm{tr}(C_j^\dagger \sigma),$$

$$\mathcal{F}(\sigma) = \sum_j \sqrt{s_j} Y_j \mathrm{tr}((C_j^T)^\dagger \sigma). \tag{46}$$

Let us now consider an alternative separable decomposition for Ψ in terms of another set of basis operators (where $\mathrm{tr}(W_i^\dagger W_j) = d\delta_{ij}$), i.e. now decompose Ψ as:

$$\Psi = \frac{1}{d^2} \sum_i W_i \otimes W_i^T. \tag{47}$$

For the state ρ_{AB} and for any suitable choice of $\mathcal{E}, \mathcal{F}$ we can get a separable decomposition for ρ_{AB} by acting upon this decomposition of Ψ with $\mathcal{E}, \mathcal{F}$ to give:

$$\rho_{AB} = \frac{1}{d^2} \sum_{i=1}^{d^2} \mathcal{E}(W_i) \otimes \mathcal{F}(W_i^T). \tag{48}$$

It is not difficult to see that this gives a solution to Problem 2: the local state spaces given by the convex hulls of the $\{\mathcal{E}(W_i)\}$ and the $\{\mathcal{F}(W_i^T)\}$ are solutions to Problem 2 for ρ_{AB} because if they weren't, we could identify smaller subset that give a separable decomposition for ρ_{AB}, and then apply the inverse maps to get smaller state spaces than the convex hulls of the W_i, W_i^T for which Ψ is separable. As this is not possible, this means that $\{\mathcal{E}(W_i)\}$ and $\{\mathcal{F}(W_i^T)\}$ give solutions to Problem 2 for ρ_{AB} of full operator-Schmidt rank. Note that this method of solution is different to the one developed earlier sections—there is no restriction on $\mathcal{E}, \mathcal{F}$ other than invertibility, and so far we have not considered any norms on ρ_{AB} anyway.

We will now see how this method can be extended to solve Problem 2* provided that we make two further assumptions about ρ_{AB}:

(A) There exist W_i, W_i^T satisfying $\mathrm{tr}(W_i^\dagger W_j) = d\delta_{ij}$ such that the $\{\mathcal{E}(W_i)\}$ and $\{\mathcal{F}(W_i^T)\}$ are of unit trace. In the appendix we show that an infinite number of such Ws can be found if ρ_{AB} is such that $\mathrm{tr}(X_i) = \mathrm{tr}(Y_i)$, and this is in turn true for instance if ρ_{AB} is a bipartite pure state. Furthermore, in the case that $\mathcal{E}, \mathcal{F}$ are CPTP maps then this can always be achieved by picking Ws with unit trace.

(B) There exist $\mathcal{E}, \mathcal{F}$ such that $\mathcal{E}^{-1}, \mathcal{F}^{-1}$ don't increase 2-norms too much, by which we mean that for all local quantum states σ on either Alice or Bob's side, it holds that $||\mathcal{E}^{-1}(\sigma)||_2, ||\mathcal{F}^{-1}(\sigma)||_2 < \sqrt{d}$. In the appendix we show that this is guaranteed provided that $\min_k s_k \geq 1/d^2$, which is in turn true provided for states in a region around the maximally entangled state (because for the maximally entangled state $\min_k s_k = 1/d$).

Let us now see why if ρ_{AB} satisfies these two conditions we are able to solve Problem 3. Define a cross norm as follows:

$$\|\rho_{AB}\|_{\mathcal{E}^{-1}, \mathcal{F}^{-1}} := \inf\{\sum_i ||\mathcal{E}^{-1}(A_i)||_2 ||\mathcal{F}^{-1}(B_i)||_2 \; : \; \rho_{AB} = \sum_i A_i \otimes B_i\}. \tag{49}$$

Then from the fact that $\rho_{AB} = \mathcal{E} \otimes \mathcal{F}(\Psi)$ we are able to calculate the value of this norm on ρ_{AB}. Indeed we see that:

$$\|\rho_{AB}\|_{\mathcal{E}^{-1}, \mathcal{F}^{-1}} = \|\Psi\|_{I,I} = d, \tag{50}$$

because if a product decomposition of ρ_{AB} existed giving a lower value of $\|\rho_{AB}\|_{\mathcal{E}^{-1}, \mathcal{F}^{-1}}$, then by applying the inverse transformations we would end up with a decomposition for Ψ with a lower value than d for $\|\Psi\|_{I,I}$, which is not possible (note that $\|\Psi\|_{I,I}$ is more conventionally denoted $\|\Psi\|_\gamma$ [14]).

Now let us consider the unit trace convex hulls of the local quantum states with either $\{\mathcal{E}(W_i)\}$ or $\{\mathcal{F}(W_i^T)\}$. We will argue that these state spaces are solutions to Problem 2* by invoking properties (A),(B). First we note that the operators $\{\mathcal{E}(W_i)\}$ and $\{\mathcal{F}(W_i^T)\}$ have the property that $||\mathcal{E}^{-1}(\mathcal{E}(W_i))||_2 = ||\mathcal{F}^{-1}(\mathcal{F}(W_i^T))||_2 = \sqrt{d}$. Hence, if we consider the unit trace convex hulls made from $\{\mathcal{E}(W_i)\}$ and $\{\mathcal{F}(W_i^T)\}$ with the local quantum states, then a separable decomposition for ρ_{AB} in terms of these unit trace convex sets can only involve the operators $\{\mathcal{E}(W_i)\}$ and $\{\mathcal{F}(W_i^T)\}$. This is because by applying the strictness of the triangle inequality (10), all other operators in these sets have a norm $||\mathcal{E}^{-1}(\bullet)||_2$ or $||\mathcal{E}^{-1}(\bullet)||_2$ that is too small - only the $\{\mathcal{E}(W_i)\}$ and $\{\mathcal{F}(W_i^T)\}$ can allow us to attain $\|\rho_{AB}\|_{\mathcal{E}^{-1}, \mathcal{F}^{-1}} = d$. Moreover, all of the operators $\{\mathcal{E}(W_i)\}$ and $\{\mathcal{F}(W_i^T)\}$ must appear in the separable decomposition, due to the operator-Schmidt rank of ρ_{AB}. This means that the convex hulls of $\{\mathcal{E}(W_i)\}$ and $\{\mathcal{F}(W_i^T)\}$ with the local quantum states are solutions to Problem 2* for ρ_{AB}, and hence the conic hulls of these sets give a solution to Problem 3.

VIII. CONNECTIONS WITH DECOMPOSITIONS OF LINEAR OPERATORS

There is a standard association between tensor products and linear operators that has appeared in multiple ways in the literature (equivalence of the singular value decompositions and the Schmidt-decomposition, Choi-Jamiołkowski isomorphism, etc). All our results can hence be expressed in terms of linear operators rather than tensor products, and indeed the results on Problems 1 and 2 have a close relationship to previous results on decompositions of positive operators. In this section we clarify these connections.

Following the association made in [15], Eq. (21) can be viewed as a decomposition of a linear operator on the left hand side in terms of a ket-bra of vectors $\boldsymbol{a}^k$ and $\boldsymbol{b}^k$ appearing on the right hand side. With this viewpoint, Problem 2 can hence be reexpressed in the following way. Given a linear operator G acting on some finite dimensional complex vector space, find convex decompositions in terms of rank-1 operators:

$$G = \sum_i p_i \left| A_i \right\rangle \left\langle B_i \right|, \tag{51}$$

that are 'smallest' ones, in the sense that the convex hulls $\mathrm{conv}\{|A_i\rangle\}$ and $\mathrm{conv}\{|B_i\rangle\}$ cannot be made smaller while continuing to admit a rank-1 convex decomposition for G. Problem 1 on the other hand asks for convex decompositions into rank-1 operators where the sums of the products of various norms of the $\{|A_i\rangle\}$ and $\{|B_i\rangle\}$ are minimised. The proofs of Theorems 1 and 2 show that some such solutions can be identified by reducing the problem to one of finding decompositions of positive operators, and moreover that the Schmidt/Singular value decomposition and simple transformations of it lead to these solutions. Problem 3 appears to have no similar natural linear algebraic interpretation.

IX. CONCLUSIONS

We have considered the construction of separable decompositions for entangled quantum states that are obtained by relaxing the requirement that the local operators in the decomposition be positive unit trace quantum states. The motivation for this problem is the construction of LHV models and classically efficient simulations for bipartite entangled quantum states, or the multipartite quantum states that are built from them.

In this context it is of interest to find 'smallest' decompositions, where the sets of local operators cannot be made smaller while continuing to allow a separable description. We consider four variants of this problem. The first (Problem 1) uses norms to quantify how 'small' a set of operators is. The second (Problem 2) uses set inclusion (so that one set is smaller than another if it is contained within it). The third (Problem 2*) and fourth (Problem 3) variants we add restrictions (requiring the sets to be unit trace, contain the quantum states, or be cones) that are important when constructing LHV models.

For the first problem we present all solutions for bipartite states for some norms. For the second problem we present some solutions for all bipartite quantum states. For the remaining two problems we obtain solutions for some bipartite states, including pure states in a region around the maximally entangled state.

Our results generalise those of [8], and have strong relationships to the study of generalised probabilistic theories with operator spaces [2], the study of discrete phase space distributions [12], cross norm entanglement measures [14, 15], and decompositions of linear operators [20–23]. In the manner of [7] we believe that they may find applications in the study and simulation of entangled many-body quantum states.

Acknowledgements

X.   APPENDIX: CONDITIONS (A) AND (B)

A: We show that condition (A) can be met if the operators in the operator-Schmidt decomposition of ρ_{AB} satisfy $\text{tr}(X_i) = \text{tr}(Y_i)$. If we take $\mathcal{E}, \mathcal{F}$ to be as in Eq. (46), then we are looking for Ws that satisfy:

$$
\begin{aligned}
1 = \text{tr}(\mathcal{E}(W_k)) &= \sum_j \sqrt{s_j}\,\text{tr}(X_j)\text{tr}(C_j^\dagger W_k), \\
&= \sum_j \sqrt{s_j}\,dT_{kj}\text{tr}(X_j), \\
1 = \text{tr}(\mathcal{F}(W_k)) &= \sum_j \sqrt{s_j}\,\text{tr}(Y_j)\text{tr}((C^T)_j^\dagger W_k^T), \\
&= \sum_j \sqrt{s_j}\,dT_{kj}\text{tr}(Y_j),
\end{aligned}
$$

where T_{kj} form the elements of a unitary defined by $dT_{kj} = \text{tr}(C_j^\dagger W_k) = \text{tr}((C^T)_j^\dagger W_k^T)$. Working directly with T rather than the Ws, we find that we must solve:

$$
\begin{aligned}
\frac{1}{d} &= \sum_j \sqrt{s_j}\,T_{kj}\text{tr}(X_j), \\
\frac{1}{d} &= \sum_j \sqrt{s_j}\,T_{kj}\text{tr}(Y_j).
\end{aligned}
$$

Denoting the vector with coefficients $\sqrt{s_j}\text{tr}(X_j)$ by e, the vector with coefficients $\sqrt{s_j}\text{tr}(Y_j)$ by f (both of which are real by Hermiticity of X_i, Y_i), and the unit vector with constant coefficients $1/d$ by g, we see that we must find T such that:

$$ Te = Tf = g. \tag{52} $$

Hence e, f must be identical unit vectors if a solution is to exist, and moreover an infinite number of T can be found in such cases. So we simply need to work out when e, f are identical unit vectors. From the fact that:

$$ \text{tr}(\rho_{AB}) = 1 = \sum_i s_i \text{tr}(X_i)\text{tr}(Y_i) = e.f \le \|e\|_2 \|f\|_2, \tag{53} $$

and using Cauchy-Schwarz we see that e, f are identical unit vectors iff $\text{tr}(X_i) = \text{tr}(Y_i)$.

B: Let us represent a quantum state σ by $\sigma = \sum_k x_k X_k$, so that the square 2-norm of σ is given by $\|\sigma\|_2^2 = \sum_k x_k^2 \le 1$. $\mathcal{E}^{-1}(\sigma)$ is given by:

$$ \mathcal{E}^{-1}(\sigma) = \sum_k \frac{1}{d\sqrt{s_k}} C_k \text{tr}(X_k^\dagger \sigma). \tag{54} $$

A similar equation holds for $\mathcal{F}$. The square of the 2-norm of this expression is given by:

$$ \|\mathcal{E}^{-1}(\sigma)\|_2^2 = \sum_k \frac{|x_k|^2}{ds_k} \le \frac{1}{d\min_k s_k}, \tag{55} $$

so if we have:

$$ 1/(d\min_k s_k) < d \Rightarrow \min_k s_k > 1/d^2, \tag{56} $$

then condition (B) holds. In the case of the maximally entangled state $\min_k s_k = 1/d$, so by continuity there is a region of states around the maximally entangled state that satisfy this.

References

[1] J. Barrett, Phys. Rev. A **75**, 032304 (2007); L. Hardy, quant-ph/0101012; A.S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, North Holland (1982); S. Popescu and D. Rohrlich, Found. Phys. **24**, 379, (1994); P. Janotta and H. Hinrichsen, J. Phys. A **47**, 323001 (2014).

[2] A. Acín, R. Augusiak, D. Cavalcanti, C. Hadley, J.K. Korbicz, M. Lewenstein, Ll. Masanes, M. Piani, Phys. Rev. Lett. **104**, 140404 (2010); J. Geller and M. Piani, J. Phys. A: Math. Theor. **47**, 424030 (2014).

[3] H. Barnum, E. Knill, G. Ortiz, and L. Viola, Phys. Rev. A **68**, 032308 (2003).

[4] J.S. Bell, Physics **1**, 195 (1964); N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014); R. Augusiak, M. Demianowicz, and A. Acin, J. Phys. A **47**, 424002 (2014).

[5] N. Ratanje and S. Virmani, Phys. Rev. A **83**, 032309 (2011).

[6] N. Ratanje and S. Virmani, arXiv:1201.0613.

[7] H. Anwar, S. Jevtic, O. Rudolph, and S. Virmani, arXiv:1412.3780.

[8] H. Anwar, S. Jevtic, O. Rudolph, and S. Virmani, New J. Phys. **17**, 093047 (2015).

[9] R.F. Werner, Phys. Rev. A **40**, 4277 (1989).

[10] E.P. Wigner, Phys. Rev. **40**, 749 (1932).

[11] J.S. Bell, *Speakable and unspeakable in Quantum Mechanics*, 2nd edition, Cambridge University Press (2004).

[12] W.K. Wootters, Ann. Phys. **176**, 1 (1987); K.S. Gibbons, M.J. Hoffman, and W.K. Wootters, Phys. Rev. A. **70**, 062101 (2004); E.F. Galvão, Phys. Rev. A. **71** 042302, (2005); D. Gross, J. Math. Phys. **47**, 122107 (2006).

[13] S. Bravyi and A. Kitaev, Phys. Rev. A **71**, 022316 (2005).

[14] O. Rudolph, J. Phys. A: Math. Gen. **33**, 3951 (2000); O. Rudolph, J. Math. Phys. **42**, 5306 (2001).

[15] O. Rudolph, Lett. Math. Phys. **70**, 57 (2004); O. Rudolph, Quant. Inf. Proc. **4**, 219 (2005).

[16] S. Boyd and L. Vandenberghe, '*Convex Optimization*', Cambridge University Press (2004).

[17] Note that if a set of operators, say $\{A_i\}$, is generalised-positive for a given process, then so is the *cone* of operators $\{r_i A_i\}$ generated by multiplying the operators by arbitrary positive numbers r_i. This means that if our motivation is simply to construct LHV models for a particular class of measurements without considering any transformation L (i.e. the process is $\mathcal{P}(I, M)$) then a separable decomposition w.r.t. a given pair of convex sets does not imply a LHV model for any more measurements than separability w.r.t. the cones generated by those sets, as described in section I. In Problem 3 the additional condition of positive trace is added in order to ensure that the operators are generalised positive for complete POVMs, and the condition that the cones contain the quantum states is added to ensure that we only consider state spaces that are generalised positive for quantum effects.

[18] For convenience we reproduce the following standard argument for showing that the triangle inequality is strict for the norms that we consider:

$$
\begin{aligned}
\|X + Y\|_\Lambda^2 &= \|X\|_\Lambda^2 + \|Y\|_\Lambda^2 + \text{tr}(\Lambda(X)^\dagger \Lambda(Y)) + \text{tr}(\Lambda(X)\Lambda(Y)^\dagger), \\
&= \|X\|_\Lambda^2 + \|Y\|_\Lambda^2 + 2\text{Re}\{\text{tr}(\Lambda(X)^\dagger \Lambda(Y))\}, \\
&\leq \|X\|_\Lambda^2 + \|Y\|_\Lambda^2 + 2\|X\|_\Lambda \|Y\|_\Lambda, \\
&\leq (\|X\|_\Lambda + \|Y\|_\Lambda)^2, \\
\Rightarrow \|X + Y\|_\Lambda &\leq \|X\|_\Lambda + \|Y\|_\Lambda.
\end{aligned}
\tag{57}
$$

The third line follows from the Cauchy-Schwarz inequality, which is strict unless $\Lambda(X) = c\Lambda(Y)$ for some positive number c. As Λ is taken to be invertible this is equivalent to saying that the inequalities are strict unless $X = cY$.

[19] N. Johnston and D.W. Kribs, Houston J. Math. **41**,831 (2015).

[20] K. Kornelson and D. Larson, Contemp. Math. **345**, 203 (2004).

[21] D. Han, K. Kornelson, D. Larson and E. Weber, *Frames for Undergraduates*, Amer. Math. Soc. (2008).

[22] L.P. Hughston, R. Jozsa, and W.K. Wootters, Phys. Lett. A **183**, 14 (1993).

[23] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press (2000).

[24] Verification:

$$
\begin{aligned}
\sum_{k=1}^N p_k A^k \otimes B^k &= \sum_{k=1}^N \sum_{i,j=1}^D (\sqrt{S}R^{-1}U)_{ik} X_i \otimes (\sqrt{S}RU)_{jk}^* Y_j, \\
&= \sum_{i,j=1}^D \left(\sum_{k=1}^N (\sqrt{S}R^{-1}U)_{ik} (\sqrt{S}RU)_{kj}^\dagger \right) X_i \otimes Y_j, \\
&= \sum S_{ij} X_i \otimes Y_j = \sum s_i X_i \otimes Y_i = \rho_{AB}.
\end{aligned}
\tag{58}
$$

[25] The existence of non-positive operators satisfying these conditions is possible when restricted (L, M) are considered. In situations where the full set of quantum operations and measurements are considered, then the ρ_i are necessarily positive.

[26] A conic hull of a set H is the set of all $ah_1 + bh_2$ where $a, b > 0$ and $h_1, h_2 \in H$, see [16] for details.