

# Lower bounds for testing complete positivity and quantum separability<sup>★</sup>

Costin Bădescu<sup>1</sup> and Ryan O'Donnell<sup>1</sup>

Computer Science Department, Carnegie Mellon University, Pittsburgh PA 15213, USA  
{cbadescu,odonnell}@cs.cmu.edu

**Abstract.** In this work, we study the *separability* problem in quantum property testing, where one is given  $n$  copies of an unknown mixed quantum state  $\varrho$  on  $\mathbb{C}^d \otimes \mathbb{C}^d$ , and one wants to test whether  $\varrho$  is separable or  $\epsilon$ -far from all separable states in trace distance. We prove that  $n = \Omega(d^2/\epsilon^2)$  copies are necessary to test separability, assuming  $\epsilon$  is not too small, viz.  $\epsilon = \Omega(1/\sqrt{d})$ . We also study completely positive distributions on the grid  $[d] \times [d]$  as a classical analog of separable states and prove that  $\Omega(d/\epsilon^2)$  samples from an unknown distribution  $p$  are necessary to decide whether  $p$  is completely positive or  $\epsilon$ -far from all completely positive distributions in total variation distance.

**Keywords:** quantum entanglement · complete positivity · property testing

## 1 Introduction

A quantum state  $\varrho$  on  $\mathbb{C}^d \otimes \mathbb{C}^d$  is said to be *separable* if it can be written as a convex combination of product states, meaning states of the form  $\rho_1 \otimes \rho_2$  where  $\rho_1$  and  $\rho_2$  are quantum states on  $\mathbb{C}^d$ . Separable quantum states are precisely those states which do not exhibit any form of quantum entanglement. These are the only states that can be prepared by separated parties who can only share classical information. Understanding the general structure and properties of the set of separable states in higher dimensions is a difficult problem and is the subject of ongoing research. For instance, deciding whether a given  $d^2 \times d^2$  matrix represents a separable state on  $\mathbb{C}^d \otimes \mathbb{C}^d$  – also known as the *separability problem* in the quantum literature – is NP-hard [Gur04]. In this work, we study the following property testing version of the separability problem:

Provided unrestricted measurement access to  $n$  copies of an unknown quantum state  $\varrho$  on  $\mathbb{C}^d \otimes \mathbb{C}^d$ , decide with high probability if  $\varrho$  is separable or  $\epsilon$ -far from all separable states in trace distance.

The ultimate goal is to determine the number of copies of  $\varrho$  that is necessary and sufficient to solve the problem, up to constant factors, as a function of  $d$  and  $\epsilon$ .

By estimating  $\varrho$  using recent algorithms for quantum state tomography [HHJ<sup>+</sup>16, OW16] and checking if the estimate is sufficiently close to a separable state, this problem can be solved using  $O(d^4/\epsilon^2)$  copies of  $\varrho$ . In this paper, we prove a lower bound, showing that  $\Omega(d^2/\epsilon^2)$  copies of  $\varrho$  are necessary, provided  $\epsilon = \Omega(1/\sqrt{d})$ .

---

<sup>★</sup> Supported by NSF grant CCF-1717606. This material is based upon work supported by the National Science Foundation under grant numbers listed above. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation (NSF).

Analogies between quantum states and classical probability distributions have proven to be a helpful source of inspiration throughout quantum theory. Unfortunately, entanglement is understood to be a purely quantum phenomenon; every finitely-supported discrete distribution can be expressed as a convex combination of product point distributions, so there are no “entangled” distributions. But motivated by the characterization of separable quantum states using symmetric extensions and the quantum de Finetti theorem [DPS04], we study mixtures of i.i.d. bivariate distributions which arise in the classical de Finetti theorem. [DPS04] uses the quantum de Finetti theorem to show that a quantum state  $\rho$  on  $\mathbb{C}^d \otimes \mathbb{C}^d$  is separable (i.e. a mixture of product states) if and only if  $\rho$  has a symmetric extension to  $\mathbb{C}^d \otimes (\mathbb{C}^d)^{\otimes k}$  for any positive integer  $k$ . Somewhat analogously, the classical de Finetti theorem states that a sequence of real random variables is a mixture of i.i.d. sequences of random variables if and only if it is exchangeable [Dia77].

We call distributions which are mixtures of i.i.d. bivariate distributions *completely positive* due to their connection with completely positive matrices. We show that, given sample access to an unknown distribution  $p$  over  $[d] \times [d]$ ,  $\Omega(d/\epsilon^2)$  samples are necessary to decide with high probability if  $p$  is completely positive or  $\epsilon$ -far from all completely positive distributions in total variation distance. Our proof also yields a generalization of Paninski’s lower bound for testing if a distribution is uniform [Pan08].

## 1.1 Previous work

The property testing version of the separability problem, as defined above, appears in [MdW16], where a lower bound of  $\Omega(d^2)$  is proven for constant  $\epsilon$ . As in [MdW16], our proof also reduces the problem of testing if a state is separable to the problem of testing if a state is the maximally mixed state. However, we do not make use of the entanglement of formation measure, as [MdW16] does, and instead rely on results about the convex structure of the set of separable states. This approach yields a more direct proof that certain random states are w.h.p. far from separable, which allows us to take advantage of a lower bound from [OW15] (see [Theorem 1](#)).

There is an extensive literature on the subject of entanglement detection (see e.g. [GT09, HHHH09]), establishing different criteria for detecting or verifying entanglement. However, it is not obvious how these results can be applied in the property testing setting. In particular, few of these criteria are specifically concerned with states that are far from separable in trace distance and many only apply to certain restricted classes of quantum states.

Our proof of the lower bound for testing if a distribution is completely positive is inspired by and generalizes Paninski’s lower bound for testing if a distribution is uniform [Pan08]. In our lower bound, we use distributions over  $[d] \times [d]$  that are more structured than the distributions on  $[d]$  in Paninski’s work. Nevertheless, our overall proof strategy is similar to Paninski’s.

## 1.2 Outline

In [Section 2](#) we cover background material on completely positive distributions, quantum states and separability, and the property testing framework that our results are concerned with. In [Section 3](#), we prove that testing if a distribution  $p$  on  $[d] \times [d]$  is completely positive or  $\epsilon$ -far from all completely positive distributions in total variation distance requires  $\Omega(d/\epsilon^2)$  samples from  $p$ . Finally, in [Section 4](#), we show that testing if a quantum state  $\rho$  on  $\mathbb{C}^d \otimes \mathbb{C}^d$  is separable or  $\epsilon$ -far from all separable states in trace distance requires  $\Omega(d^2/\epsilon^2)$  copies of  $\rho$  when  $\epsilon = \Omega(1/\sqrt{d})$ .

## 2 Preliminaries

This section covers the mathematical background and notation used in the rest of the paper.

### 2.1 Completely positive distributions

There is a well-developed theory of completely positive and copositive matrices (see e.g. [GM12, Chapter 7]). In this section, we review some known material.

Let  $d$  be a positive integer. We consider distributions over the grid  $[d]^2 = [d] \times [d] = \{(1, 1), (1, 2), \dots, (d, d)\}$  which we represent as matrices  $A \in \mathbb{R}^{d \times d}$  with  $A_{ij}$  being the probability of sampling  $(i, j)$ .

*Example 1.* If  $p \in \mathbb{R}^d$  is a distribution on  $[d] = \{1, \dots, d\}$  represented as a column vector, then  $pp^\top$  is the natural i.i.d. product probability distribution on  $[d] \times [d]$  derived from  $p$  with  $p_i p_j$  being the probability of sampling  $(i, j)$ .

**Definition 1.** A matrix  $A \in \mathbb{R}^{d \times d}$  is completely positive (CP) if there exist vectors  $v_1, \dots, v_k \in \mathbb{R}_{\geq 0}^d$  with nonnegative entries such that  $A$  can be expressed as a convex combination of their projections  $v_1 v_1^\top, \dots, v_k v_k^\top$ , viz.  $A = \sum_{i=1}^k c_i v_i v_i^\top$  for some nonnegative real numbers  $c_1, \dots, c_k \in \mathbb{R}$  with  $c_1 + \dots + c_k = 1$ .

A distribution on  $[d]^2$  represented as a matrix  $A$  is completely positive if  $A$  is a CP matrix.

*Remark 1.* For a CP distribution  $A$ , the associated vectors  $v_i$  may be taken to be probability distributions, since one can replace  $v_i$  by  $v_i / \|v_i\|_1$  and  $c_i$  by  $c_i \|v_i\|_1^2$ . Thus, CP distributions are mixtures of i.i.d. distributions.

It follows immediately from Definition 1 that a CP matrix  $A$  satisfies three basic properties:

- (i)  $A$  is symmetric ( $A^\top = A$ ),
- (ii)  $A_{ij} \geq 0$  for all  $i, j \in [d]$ , and
- (iii)  $A$  is positive semidefinite (PSD), denoted  $A \geq 0$ .

A matrix satisfying these three properties is called *doubly nonnegative*. However, if  $d \geq 5$ , then there exist doubly nonnegative matrices which are not completely positive [MM62].

*Example 2.* Let  $J$  denote the  $d \times d$  matrix with  $J_{ij} = 1$  for all  $i, j \in [d]$  and let  $\text{Unif}_{d^2} = J/d^2$  denote the uniform distribution on  $[d]^2$ . Since  $\text{Unif}_{d^2} = (\frac{1}{d}, \dots, \frac{1}{d})(\frac{1}{d}, \dots, \frac{1}{d})^\top$ , the uniform distribution on  $[d]^2$  is completely positive.

Let  $\text{CP}_d$  denote the set of completely positive  $d \times d$  matrices and let  $\text{CPD}_d$  denote its subset of completely positive distributions on  $[d]^2$ . It is well known that  $\text{CP}_d$  is a cone and that its dual cone consists of *copositive* matrices, i.e. matrices  $M$  such that  $x^\top M x \geq 0$  for all nonnegative vectors  $x \in \mathbb{R}_{\geq 0}^d$ . Thus, by cone duality, if  $B \notin \text{CP}_d$  is a non-CP matrix, then there exists a copositive matrix  $W$  such that  $\text{tr}(AW) \geq 0$  for all  $A \in \text{CP}_d$  and  $\text{tr}(BW) < 0$ . This result yields witnesses certifying nonmembership in  $\text{CPD}_d$ . However, its usefulness is limited by the fact that it provides no quantitative information about how far a nonmember  $A$  is from the set  $\text{CPD}_d$ .

In what follows, we interpret distributions on  $[d]^2$  as weighted directed graphs with self-loops and obtain a sufficient condition for a distribution to be  $\epsilon$ -far in total variation distance from  $\text{CPD}_d$  in terms of the maximum value of a cut in the corresponding graph.

We interpret a distribution  $A$  on  $[d]^2$  as a weighted directed graph  $G$  with vertices  $V(G) = [d]$  and edges

$$E(G) = \{(i, j) \in [d]^2 \mid A_{ij} > 0\}.$$

A cut  $x \in \{\pm 1\}^d$  in a  $G$  is a bipartition of the vertices  $V(G) = E_1 \cup E_2$  with  $E_1 = \{i \in [d] \mid x_i < 0\}$  and  $E_2 = \{i \in [d] \mid x_i > 0\}$ . The total weight of edges cut by this bipartition is

$$\sum_{(i,j) \in [d]^2} \frac{1 - x_i x_j}{2} A_{ij} = \mathbf{E}_{(i,j) \sim A} \frac{1 - x_i x_j}{2} = \frac{1}{2} - \frac{1}{2} \mathbf{E}_{(i,j) \sim A} x_i x_j = \frac{1}{2} - \frac{1}{2} x^\top A x.$$

In particular, if  $A = pp^\top$  with  $p \in \mathbb{R}^d$ , then

$$x^\top A x = x^\top p p^\top x = (x^\top p)^2 \geq 0.$$

By [Remark 1](#), a CP distribution is a convex combination of matrices of the form  $pp^\top$ . Thus, it holds that

**Proposition 1.** *If  $A$  is a CP distribution, then the total weight of a cut in the graph represented by  $A$  is at most  $\frac{1}{2}$ .*

This fact allows us to prove the following result which gives a sufficient condition for a distribution to be  $\epsilon$ -far from all CP distributions in  $\ell^1$  distance:

**Proposition 2.** *Let  $A$  be a distribution on  $[d]^2$ . If there exists a cut  $x \in \{\pm 1\}^d$  with  $x^\top A x \leq -\epsilon$ , then  $\|B - A\|_1 \geq \epsilon$  for all  $B \in \text{CPD}_d$ .*

*Proof.* Let  $B \in \text{CPD}_d$  be arbitrary. By Hölder's inequality, for all  $U \in \mathbb{R}^{d \times d}$  with  $\|U\|_\infty = 1$ ,

$$\|B - A\|_1 \geq \text{tr}(U^\top (B - A)) = \text{tr}(U^\top B) - \text{tr}(U^\top A).$$

Let  $U = xx^\top$ . Since  $x^\top B x \geq 0$  and  $\text{tr}(U^\top A) = x^\top A x \leq -\epsilon$ ,

$$\|B - A\|_1 \geq x^\top B x - x^\top A x \geq \epsilon.$$

## 2.2 Quantum states and separability

This section serves as a brief introduction to quantum states and separability. For a more comprehensive introduction, see e.g. [\[Wat18\]](#).

We work over  $\mathbb{C}$  and use bra-ket notation to denote vectors in  $\mathbb{C}^d$ , viz. for all vectors  $x, y \in \mathbb{C}^d$  and matrices  $A \in \mathbb{C}^{d \times d}$ ,  $|x\rangle = x$ ,  $\langle x| = x^\dagger = \bar{x}^\top$ ,  $\langle x \otimes y| = \langle x| \otimes \langle y|$ ,  $|x \otimes y\rangle = |x\rangle \otimes |y\rangle$ ,  $\langle x|y\rangle = x^\dagger y$ ,  $|x\rangle\langle y| = xy^\dagger$ , and  $\langle x|A|y\rangle = x^\dagger A y$ .

**Definition 2.** A quantum state  $\rho$  on  $\mathbb{C}^d$  is a positive semidefinite matrix  $\rho \in \mathbb{C}^{d \times d}$  with  $\text{tr}(\rho) = 1$ . A measurement is a set  $\{E_1, \dots, E_k\}$  of positive semidefinite matrices on  $\mathbb{C}^d$  with  $E_1 + \dots + E_k = \mathbb{1}$ , where  $\mathbb{1}$  denotes the identity matrix.

Let  $\rho$  and  $\{E_1, \dots, E_k\}$  be as in the definition above and let  $p_i = \text{tr}(\rho E_i)$  for  $i = 1, \dots, k$ . Since  $\rho$  and the  $E_i$  are PSD,  $p_i \geq 0$  for all  $i = 1, \dots, k$ , and

$$p_1 + \dots + p_k = \text{tr}(\rho E_1) + \dots + \text{tr}(\rho E_k) = \text{tr}(\rho(E_1 + \dots + E_k)) = \text{tr}(\rho) = 1.$$

Hence,  $(p_1, \dots, p_k)$  is a distribution on  $[k]$ . Applying the measurement  $\{E_1, \dots, E_k\}$  to the quantum state  $\rho$  yields outcome  $i \in [k]$  with probability  $p_i = \text{tr}(\rho E_i)$ .

*Example 3.*  $\frac{1}{d}$  is a quantum state on  $\mathbb{C}^d$  called the *maximally mixed state*; it is analogous to the uniform distribution on  $[d]$ .

**Definition 3.** A state of the form  $\rho = |x\rangle\langle x|$  for some  $x \in \mathbb{C}^d$  is called a *pure state*.

Given quantum states  $\rho$  and  $\sigma$  on  $\mathbb{C}^d$ , the tensor product  $\rho \otimes \sigma$  is a quantum state on  $\mathbb{C}^d \otimes \mathbb{C}^d$ . If  $\rho$  and  $\sigma$  represent the individual states of two isolated particles, then  $\rho \otimes \sigma$  is the state of the physical system comprising both particles. Thus, the system composed of  $n$  identical copies of the state  $\rho$  is represented as the state  $\rho^{\otimes n}$  on  $(\mathbb{C}^d)^{\otimes n}$ .

**Definition 4.** A quantum state  $\varrho$  on  $\mathbb{C}^d \otimes \mathbb{C}^d$  is separable if  $\varrho$  can be expressed as a convex combination of product states, viz.  $\varrho = \sum_{i=1}^k c_i \rho_i \otimes \sigma_i$ , where  $\rho_i$  and  $\sigma_i$  are states on  $\mathbb{C}^d$  for  $i = 1, \dots, k$  and  $c_1, \dots, c_k \in \mathbb{R}_{\geq 0}$  satisfy  $c_1 + \dots + c_k = 1$ . Thus, the physical system represented by  $\varrho$  may be regarded as being in the state  $\rho_i \otimes \sigma_i$  with probability  $c_i$ .

A state that is not separable is called *entangled*.

*Example 4.* Since  $\frac{1}{d^2} = \frac{1}{d} \otimes \frac{1}{d}$ , the maximally mixed state is separable.

**Definition 5.** Let  $\text{Sep}$  denote the set of separable states on  $\mathbb{C}^d \otimes \mathbb{C}^d$  and let  $\text{Sep}_{\pm}$  denote its cylindrical symmetrization (cf. [AS17, p. 81]), viz.  $\text{Sep}_{\pm} = \text{conv}(\text{Sep} \cup (-\text{Sep}))$ , where  $\text{conv}(E)$  denotes the convex hull of the set  $E$ .

Similar to the duality between completely positive and copositive matrices, the set  $\text{Sep}$  generates a cone of separable operators whose dual is the cone of *block-positive* operators (see e.g. [AS17]). A block-positive operator acts as an entanglement witness certifying that a given quantum state is not separable. Thus, [Proposition 4](#) below is comparable to [Proposition 2](#) in that it describes witnesses certifying that a quantum state is not just entangled but actually  $\epsilon$ -far from all separable states in trace distance.

### 2.3 The property testing framework

In the property testing model, we have a set  $\mathcal{O}$  of objects and also a distance function  $\text{dist} : \mathcal{O} \times \mathcal{O} \rightarrow \mathbb{R}$ . A *property*  $\mathcal{P}$  is a subset of  $\mathcal{O}$  and the distance between an object  $x \in \mathcal{O}$  and the property  $\mathcal{P}$  is defined by  $\text{dist}(x, \mathcal{P}) = \inf_{y \in \mathcal{P}} \text{dist}(x, y)$ . An algorithm  $\mathcal{T}$  is said to test  $\mathcal{P}$  if, given some type of access to  $x \in \mathcal{O}$  (e.g. independent samples or identical copies),  $\mathcal{T}$  accepts  $x$  w.h.p. when  $x \in \mathcal{P}$  and  $\mathcal{T}$  rejects  $x$  w.h.p. when  $\text{dist}(x, \mathcal{P}) \geq \epsilon$ .

In [Section 3](#),  $\mathcal{O}$  is the set of distributions on  $[d] \times [d]$ ,  $\text{dist}$  is the total variation distance, and  $\mathcal{P} = \text{CPD}_d \subseteq \mathcal{O}$  is the set of CP distributions. Given samples  $\mathbf{x}_1, \dots, \mathbf{x}_n$  from a distribution  $p$  on  $[d]^2$ , a testing algorithm  $\mathcal{T}$  for  $\text{CPD}_d$  satisfies

$$\begin{aligned} p \in \text{CPD}_d &\implies \mathbf{P}[\mathcal{T}(\mathbf{x}_1, \dots, \mathbf{x}_n) \text{ accepts}] \geq \frac{2}{3}, \\ p \epsilon\text{-far from } \text{CPD}_d &\implies \mathbf{P}[\mathcal{T}(\mathbf{x}_1, \dots, \mathbf{x}_n) \text{ accepts}] \leq \frac{1}{3}. \end{aligned}$$

In [Section 4](#),  $\mathcal{O}$  is the set of quantum states on  $\mathbb{C}^d \otimes \mathbb{C}^d$ ,  $\text{dist}(\varrho, \sigma) = \frac{1}{2} \|\varrho - \sigma\|_1$  is the trace distance between quantum states, and  $\mathcal{P} = \text{Sep}$  is the set of separable states on  $\mathbb{C}^d \otimes \mathbb{C}^d$ . Given measurement access to  $n$  copies  $\varrho^{\otimes n}$  of a state  $\varrho \in \mathbb{C}^d \otimes \mathbb{C}^d$ , a testing algorithm for  $\text{Sep}$  is a two-outcome measurement  $\{E_0, E_1\}$  on  $(\mathbb{C}^d)^{\otimes n}$  satisfying:

$$\varrho \in \text{Sep} \implies \text{tr}(E_1 \varrho^{\otimes n}) = 1] \geq \frac{2}{3}, \quad \varrho \epsilon\text{-far from } \text{Sep} \implies \text{tr}(E_1 \varrho^{\otimes n}) = 1] \leq \frac{1}{3}.$$

### 3 Testing complete positivity

Let  $d$  be a positive integer. If  $d$  is odd, we can reduce to the case of  $d - 1$  by using distributions that don't involve outcome  $d \in [d]$ , and the asymptotics of  $\Omega(d/\epsilon^2)$  remain unchanged. Hence we may assume, without loss of generality, that  $d$  is even.

We begin by defining a family of distributions on  $[d]^2$  which are  $\epsilon$ -far from  $\text{CPD}_d$ . Let  $S \subseteq [d]$  be a subset of size  $|S| = \frac{d}{2}$ . Thus,  $|S^c| = \frac{d}{2}$  and

$$|S \times S^c \cup S^c \times S| = |S \times S^c| + |S^c \times S| = \frac{d^2}{2}.$$

Let  $\phi_S : [d]^2 \rightarrow \mathbb{R}$  be the function defined by

$$\phi_S(x) = \begin{cases} 1 + \epsilon, & x \in S \times S^c \cup S^c \times S \\ 1 - \epsilon, & \text{otherwise.} \end{cases}$$

Hence,

$$\text{avg}_{x \in [d]^2} \phi_S(x) = \frac{1}{d^2} \left( \frac{d^2}{2} (1 + \epsilon) + \frac{d^2}{2} (1 - \epsilon) \right) = 1.$$

So we may think of  $\phi_S$  as a density function with respect to the uniform distribution on  $[d]^2$ .

Let  $x \in \{\pm 1\}^d$  be defined as follows: for all  $i \in [d]$ , if  $i \in S$ , then  $x_i = 1$ , otherwise  $x_i = -1$ . Let  $A^S$  be the matrix defined by  $A_{ij}^S = \phi_S((i, j))/d^2$ . Thus,  $A^S$  is a symmetric distribution on  $[d]^2$  and  $x$  is a cut. The total weight of this cut is  $\frac{d^2}{2} \cdot \frac{1+\epsilon}{d^2} = \frac{1}{2} + \frac{\epsilon}{2}$ . Therefore, for every subset  $S \subseteq [d]$ , the distribution  $A^S$  is *not* completely positive. Moreover,  $x^\top A^S x = -\epsilon$ , so, by [Proposition 2](#),  $\|A^S - B\|_1 \geq \epsilon$  for every CP distribution  $B$ . In other words, for every subset  $S \subseteq [d]$  with  $|S| = \frac{d}{2}$ ,  $A^S$  is a distribution on  $[d]^2$  which is  $\epsilon$ -far in  $\ell^1$  distance from every CP distribution on  $[d]^2$ .

Fix  $\Omega = [d]^2$  and let  $\phi : \Omega^n \rightarrow \mathbb{R}$  denote the function defined by

$$\phi(x) = \text{avg}_{\substack{S \subseteq [d] \\ |S|=d/2}} \phi_S(x_1) \cdots \phi_S(x_n).$$

Let  $\mathcal{D}_n$  denote the distribution on  $\Omega^n$  defined by the density  $\phi$  and let  $d_{\chi^2}(-, -)$  denote the  $\chi^2$ -distance between probability distributions, i.e. for distributions  $\mathcal{P}$  and  $\mathcal{Q}$  on  $\Omega$ ,

$$d_{\chi^2}(\mathcal{P}, \mathcal{Q}) = \mathbf{E}_{x \sim \mathcal{Q}} \left[ \left( \frac{\mathcal{P}(x)}{\mathcal{Q}(x)} - 1 \right)^2 \right].$$

The following proposition will be shown to imply our lower bound:

**Proposition 3.** *If  $d_{\chi^2}(\mathcal{D}_n, \text{Unif}_{d^2}^{\otimes n}) \geq \frac{1}{3}$ , then  $n = \Omega(d/\epsilon^2)$ .*

*Proof.* Let  $\mathcal{H}$  denote the uniform distribution over subsets  $S \subseteq [d]$  with  $|S| = d/2$ . Thus,

$$\begin{aligned}
d_{\chi^2}(\mathcal{D}_n, \text{Unif}_{d^2}^{\otimes n}) &= \left( \sum_{x \in \Omega^n} \frac{\mathcal{D}_n(x)^2}{\text{Unif}_{d^2}^{\otimes n}(x)} \right) - 1 \\
&= \left( \sum_{x \in \Omega^n} \frac{\phi(x)^2}{d^{2n}} \right) - 1 \\
&= \mathbf{E}_{\mathbf{x} \sim \text{Unif}_{d^2}^{\otimes n}} \phi(\mathbf{x})^2 - 1 \\
&= \mathbf{E}_{\mathbf{x} \sim \text{Unif}_{d^2}^{\otimes n}} \left[ \left( \mathbf{E}_{\mathbf{S} \sim \mathcal{H}} \phi_{\mathbf{S}}(\mathbf{x}_1) \cdots \phi_{\mathbf{S}}(\mathbf{x}_n) \right)^2 \right] - 1 \\
&= \mathbf{E}_{\mathbf{x} \sim \text{Unif}_{d^2}^{\otimes n}} \left[ \mathbf{E}_{\mathbf{S}, \mathbf{S}' \sim \mathcal{H}} \phi_{\mathbf{S}}(\mathbf{x}_1) \cdots \phi_{\mathbf{S}}(\mathbf{x}_n) \phi_{\mathbf{S}'}(\mathbf{x}_1) \cdots \phi_{\mathbf{S}'}(\mathbf{x}_n) \right] - 1 \\
&= \mathbf{E}_{\mathbf{S}, \mathbf{S}' \sim \mathcal{H}} \mathbf{E}_{\mathbf{x} \sim \text{Unif}_{d^2}^{\otimes n}} \phi_{\mathbf{S}}(\mathbf{x}_1) \cdots \phi_{\mathbf{S}}(\mathbf{x}_n) \phi_{\mathbf{S}'}(\mathbf{x}_1) \cdots \phi_{\mathbf{S}'}(\mathbf{x}_n) - 1 \\
&= \mathbf{E}_{\mathbf{S}, \mathbf{S}' \sim \mathcal{H}} \left[ \left( \mathbf{E}_{\mathbf{x} \sim \text{Unif}_{d^2}} \phi_{\mathbf{S}}(\mathbf{x}) \phi_{\mathbf{S}'}(\mathbf{x}) \right)^n \right] - 1.
\end{aligned}$$

For a subset  $E \subseteq [d]$ , let  $\chi_E$  be the  $\pm 1$ -valued indicator function defined by  $\chi_E(x) = 1$  if  $x \in E$  and  $\chi_E(x) = -1$  otherwise. Note that  $\phi_E(x) = 1 - \chi_E(x_1)\chi_E(x_2)\epsilon$  for all  $x \in \Omega$ . Hence,

$$\phi_{\mathbf{S}}(\mathbf{x})\phi_{\mathbf{S}'}(\mathbf{x}) = 1 - (\chi_{\mathbf{S}}(\mathbf{x}_1)\chi_{\mathbf{S}}(\mathbf{x}_2) + \chi_{\mathbf{S}'}(\mathbf{x}_1)\chi_{\mathbf{S}'}(\mathbf{x}_2))\epsilon + \chi_{\mathbf{S}}(\mathbf{x}_1)\chi_{\mathbf{S}}(\mathbf{x}_2)\chi_{\mathbf{S}'}(\mathbf{x}_1)\chi_{\mathbf{S}'}(\mathbf{x}_2)\epsilon^2.$$

For a fixed outcome of  $\mathbf{S}$  and  $\mathbf{x}$  uniformly random,  $\chi_{\mathbf{S}}(\mathbf{x}_1)$  and  $\chi_{\mathbf{S}}(\mathbf{x}_2)$  are independent uniform  $\pm 1$ -valued bits. So, in expectation, the terms involving just  $\epsilon$  in the expression above drop out. Moreover,  $\chi_{\mathbf{S}}(\mathbf{x}_1)\chi_{\mathbf{S}'}(\mathbf{x}_1)$  and  $\chi_{\mathbf{S}}(\mathbf{x}_2)\chi_{\mathbf{S}'}(\mathbf{x}_2)$  are independent. Hence,

$$\mathbf{E}_{\mathbf{x} \sim \text{Unif}_{d^2}} \phi_{\mathbf{S}}(\mathbf{x})\phi_{\mathbf{S}'}(\mathbf{x}) = 1 - \epsilon^2 \cdot \left( \mathbf{E}_{\mathbf{x} \sim \text{Unif}_{d^2}} \chi_{\mathbf{S}}(\mathbf{x}_1)\chi_{\mathbf{S}'}(\mathbf{x}_1) \right)^2$$

Let  $\mathbf{r} = |\mathbf{S} \cap \mathbf{S}'|$ , where  $\mathbf{S}, \mathbf{S}' \sim \mathcal{H}$ , and let  $\delta$  denote the mean of  $\chi_{\mathbf{S}}(\mathbf{x}_1)\chi_{\mathbf{S}'}(\mathbf{x}_1)$  appearing above. It is easy to check that  $\delta = 4\mathbf{r}/d - 1$ . Thus,

$$d_{\chi^2}(\mathcal{D}_n, \text{Unif}_{d^2}^{\otimes n}) \leq \mathbf{E}_{\mathbf{S}, \mathbf{S}' \sim \mathcal{H}} \left[ (1 + \epsilon^2 \delta^2)^n \right] - 1 \leq \mathbf{E}_{\mathbf{S}, \mathbf{S}' \sim \mathcal{H}} [\exp(\epsilon^2 \delta^2)^n] - 1 = \mathbf{E}_{\mathbf{S}, \mathbf{S}' \sim \mathcal{H}} [\exp(n\epsilon^2 \delta^2)] - 1.$$

Since  $\exp(n\epsilon^2 \delta^2) - 1 \geq 0$ ,

$$\mathbf{E}_{\mathbf{S}, \mathbf{S}' \sim \mathcal{H}} [\exp(n\epsilon^2 \delta^2)] - 1 = \int_0^\infty \mathbf{P}_{\mathbf{S}, \mathbf{S}' \sim \mathcal{H}} [\exp(n\epsilon^2 \delta^2) - 1 \geq t] dt.$$

Since  $\exp(n\epsilon^2 \delta^2) - 1 \geq t$  is equivalent to

$$\mathbf{r} \geq \frac{d}{4} + \frac{d}{4} \cdot \left( \frac{\log(1+t)}{n\epsilon^2} \right)^{\frac{1}{2}}$$

it follows that

$$d_{\chi^2}(\mathcal{D}_n, \text{Unif}_{d^2}^{\otimes n}) \leq \int_0^\infty \mathbf{P}_{\mathbf{S}, \mathbf{S}' \sim \mathcal{H}} \left[ \mathbf{r} \geq \frac{d}{4} + \frac{d}{4} \sqrt{f(t)} \right] dt,$$

where  $f(t) = \log(1+t)/n\epsilon^2$ .

Since  $\mathbf{r} = |\mathbf{S} \cap \mathbf{S}'|$  is invariant under permutations of  $[d]$ , it follows that  $\mathbf{r}$  is distributed according to the hypergeometric distribution with  $d/2$  draws from a set of  $d$  elements with  $d/2$  successes. If  $\mathbf{X}$  is a random variable distributed according to the hypergeometric distribution with  $m$  draws from a set of  $N$  elements with  $k$  successes, then (see e.g. [Ska13])

$$\mathbf{P} \left[ \frac{\mathbf{X}}{m} \geq \frac{k}{N} + s \right] \leq \exp(-2s^2 m).$$

Hence,

$$\mathbf{P}_{\mathbf{S}, \mathbf{S}' \sim \mathcal{H}} \left[ \mathbf{r} \cdot \frac{2}{d} \geq \frac{1}{2} + t \right] = \mathbf{P}_{\mathbf{S}, \mathbf{S}' \sim \mathcal{H}} \left[ \mathbf{r} \geq \frac{d}{4} + \frac{dt}{2} \right] \leq \exp(-dt^2),$$

whence,

$$\mathbf{P}_{\mathbf{S}, \mathbf{S}' \sim \mathcal{H}} \left[ \mathbf{r} \geq \frac{d}{4} (\sqrt{f(t)} + 1) \right] \leq \exp(-df(t)/4).$$

Therefore,

$$\begin{aligned} d_{\chi^2}(\mathcal{D}_n, \text{Unif}_{d^2}^{\otimes n}) &\leq \int_0^\infty \exp(-df(t)/4) dt \\ &= \int_0^\infty \exp\left(-\frac{d}{4n\epsilon^2} \log(1+t)\right) dt \\ &= \int_0^\infty \left(\frac{1}{1+t}\right)^c dt = \frac{1}{c-1}, \end{aligned}$$

where  $c = d/4n\epsilon^2$ . Since  $d_{\chi^2}(\mathcal{D}_n, \text{Unif}_{d^2}^{\otimes n}) \geq 1/3$ , it follows that  $c \leq 4$ , so  $n \geq d/16\epsilon^2$ . Therefore,  $n = \Omega(d/\epsilon^2)$ , as needed.

Let  $d_{\text{TV}}(-, -)$  denote the total variation distance between probability distributions. Let  $p \in \text{CPD}_d$  and let  $q$  be a distribution  $\epsilon$ -far from  $\text{CPD}_d$ .

A testing algorithm  $f : ([d]^2)^n \rightarrow \{0, 1\}$  for complete positivity determines a probability event  $E \subseteq ([d]^2)^n$  satisfying  $p^{\otimes n}(E) \geq 2/3$  and  $q^{\otimes n}(E) \leq 1/3$ . Hence,  $\text{Unif}_{d^2}^{\otimes n}(E) \geq 2/3$  and, since  $\mathcal{D}_n$  is supported on distributions  $\epsilon$ -far from  $\text{CPD}_d$ ,  $\mathcal{D}_n(E) \leq 1/3$ . Therefore,  $d_{\text{TV}}(\mathcal{D}_n, \text{Unif}_{d^2}^{\otimes n}) \geq 1/3$  and the following corollary establishes the lower bound:

**Corollary 1.** *If  $d_{\text{TV}}(\mathcal{D}_n, \text{Unif}_{d^2}^{\otimes n}) \geq 1/3$ , then  $n = \Omega(d/\epsilon^2)$ .*

*Proof.* For all distributions  $\mu$  and  $\nu$ ,  $2d_{\text{TV}}(\mu, \nu)^2 \leq d_{\chi^2}(\mu, \nu)$ . Hence,

$$(d/4n\epsilon^2 - 1)^{-1} \geq d_{\chi^2}(\mathcal{D}_n, \text{Unif}_{d^2}^{\otimes n}) \geq 2d_{\text{TV}}(\mathcal{D}_n, \text{Unif}_{d^2}^{\otimes n})^2 \geq \frac{2}{9},$$

where the first inequality is obtained in the proof of [Proposition 3](#). Therefore,  $n = \Omega(d/\epsilon^2)$ .



## 4 Testing separability

Let  $d$  be a positive integer. As in the previous section, we may assume, without loss of generality, that  $d$  is even.

Let  $\mathcal{H} = \mathbb{C}^d \otimes \mathbb{C}^d$ , let  $\mathcal{U}(\mathcal{H})$  denote the set of unitary operators on  $\mathcal{H}$ , and recall that  $\text{Sep}$  denotes the set of separable states on  $\mathcal{H}$ . For all operators  $T$  on  $\mathcal{H}$ , let  $\|T\|_p$  denote the Schatten  $p$ -norm of  $T$ , viz.  $\|T\|_p = (\text{tr}(|T|^p))^{\frac{1}{p}}$ , where  $|T| = \sqrt{T^\dagger T}$  is the absolute value of the operator  $T$ . Let  $d_{\text{tr}}(\varrho, \sigma) = \frac{1}{2}\|\varrho - \sigma\|_1$  denote the trace distance between quantum states  $\varrho$  and  $\sigma$ .

We begin by defining a family of quantum states which are with high probability  $O(\epsilon)$ -far from  $\text{Sep}$ . For  $0 \leq \epsilon \leq 1/2$ , let  $D_\epsilon$  be the diagonal matrix on  $\mathcal{H}$  defined by

$$D_\epsilon = \text{diag}\left(\frac{1+2\epsilon}{d^2}, \dots, \frac{1+2\epsilon}{d^2}, \frac{1-2\epsilon}{d^2}, \dots, \frac{1-2\epsilon}{d^2}\right),$$

where  $\text{tr}(D_\epsilon) = 1$ , and let  $\mathcal{D}$  denote the family of all quantum states on  $\mathcal{H}$  with the same spectrum as  $D_\epsilon$ , viz.  $\mathcal{D} = \{UD_\epsilon U^\dagger \mid U \in \mathcal{U}(\mathcal{H})\}$ .

Our lower bound will rely on the following theorem from [OW15]:

**Theorem 1.**  $\Omega(d^2/\epsilon^2)$  copies are necessary to test whether a quantum state  $\varrho$  on  $\mathcal{H}$  is the maximally mixed state or  $\varrho \in \mathcal{D}$ .

If  $U$  is a random unitary on  $\mathcal{H}$  distributed according to the Haar measure, then  $\varrho = UD_\epsilon U^\dagger$  is a random element of  $\mathcal{D}$ . This induced probability measure is invariant under conjugation by a fixed unitary: for all  $V \in \mathcal{U}(\mathcal{H})$ ,  $V\varrho V^\dagger$  has the same distribution as  $\varrho$ . We want to show that:

**Lemma 1.** *There is a universal constant  $C_0$  such that for all  $C_0/\sqrt{d} \leq \epsilon \leq 1/2$ , the following holds when  $\varrho = UD_\epsilon U^\dagger$  is a uniformly random state in  $\mathcal{D}$ :  $\mathbf{P}[\forall \sigma \in \text{Sep}, \|\varrho - \sigma\|_1 \geq 2\epsilon] \geq \frac{2}{3}$ .*

As  $\epsilon$  tends to zero, the elements of  $\mathcal{D}$  get closer to the maximally mixed state and eventually become separable, by the Gurvits–Barnum theorem [GB02]. Indeed, if  $\epsilon \leq 1/(2\sqrt{d^2-1})$ , then  $\mathcal{D} \subseteq \text{Sep}$ . Hence, some assumption on  $\epsilon$  is necessary for Lemma 1 to hold.

Lemma 1 and Theorem 1 easily imply the desired lower bound:

**Theorem 2.** *Let  $\varrho$  be a quantum state on  $\mathbb{C}^d \otimes \mathbb{C}^d$  and let  $\epsilon = \Omega(1/\sqrt{d})$ . Testing if  $\varrho$  is separable or  $\epsilon$ -far from  $\text{Sep}$  in trace distance requires  $\Omega(d^2/\epsilon^2)$  copies of  $\varrho$ .*

*Proof.* Let  $\{E_0, E_1\}$  be a measurement corresponding to a separability testing algorithm using  $n$  copies of  $\varrho$ . To apply the lower bound in Theorem 1, we use  $\{E_0, E_1\}$  to define an algorithm that decides w.h.p. if a state  $\varrho$  is equal to the maximally mixed state  $\frac{1}{d^2}$  or  $\varrho \in \mathcal{D}$ .

Let  $\varrho^{\otimes n}$  be given with either  $\varrho \in \mathcal{D}$  or  $\varrho = \frac{1}{d^2}$ . Note that, for all  $\varrho \in \mathcal{D}$ ,  $d_{\text{tr}}(\varrho, \frac{1}{d^2}) \geq \epsilon$  holds. Let  $U$  be a random unitary. If  $\varrho$  is the maximally mixed state, then  $V\varrho V^\dagger = \varrho$  for all  $V \in \mathcal{U}(\mathcal{H})$ , so  $(U\varrho U^\dagger)^{\otimes n} = \varrho^{\otimes n}$ . Otherwise,  $U\varrho U^\dagger$  is a random state in  $\mathcal{D}$ .

Applying the separability test  $\{E_0, E_1\}$  to  $U\varrho U^\dagger$ , we have that:

- (i) if  $U\varrho U^\dagger = \varrho = \frac{1}{d^2}$ , then  $U\varrho U^\dagger$  is separable, so  $\text{tr}((U\varrho U^\dagger)^{\otimes n} E_1) = \text{tr}(\varrho^{\otimes n} E_1) \geq \frac{2}{3}$ .
- (ii) if  $\varrho \in \mathcal{D}$ , then the probability of error is

$$\mathbf{E} \text{tr}((U\varrho U^\dagger)^{\otimes n} E_1) \leq \mathbf{P}[U\varrho U^\dagger \text{ is } \epsilon\text{-close to Sep}] + \mathbf{P}[\text{test fails} \mid U\varrho U^\dagger \text{ is } \epsilon\text{-far from Sep}]$$

which is at most  $\frac{1}{3} + \frac{1}{3} \cdot \frac{2}{3} = \frac{5}{9}$ , using Lemma 1.

Thus, using the separability test, we can distinguish w.h.p. between  $\varrho = \frac{\mathbb{1}}{d^2}$  and  $\varrho \in \mathcal{D}$  using  $n$  copies of  $\varrho$ . Therefore, by [Theorem 1](#),  $n = \Omega(d^2/\epsilon^2)$ .

It remains to show that [Lemma 1](#) holds. Its proof relies on two main facts: first, that Sep is approximated by a polytope with  $\exp(O(d))$  vertices which are separable pure states; and, second, that a random element of  $\mathcal{D}$  is  $\epsilon$ -far from a fixed pure state except with probability  $\exp(-O(d))$ .

The first fact follows from the next lemma which is a rephrasing of [\[AS17, Lemma 9.4\]](#):

**Lemma 2.** *There exists a constant  $C > 0$  such that, for every dimension  $d$ , there is a family  $\mathcal{N}$  of pure product states on  $\mathcal{H}$  (i.e. states of the form  $|x \otimes y\rangle\langle x \otimes y|$  with  $x, y \in \mathbb{C}^d$ ) with  $|\mathcal{N}| \leq C^d$  satisfying*

$$\text{conv}(\mathcal{N} \cup -\mathcal{N}) \subseteq \text{Sep}_{\pm} \subseteq 2 \text{conv}(\mathcal{N} \cup -\mathcal{N}).$$

Now, we wish to upper bound the probability that a random element of  $\mathcal{D}$  is  $\epsilon$ -far from a fixed pure state. The following result provides a sufficient condition for a state  $\sigma$  on  $\mathcal{H}$  to be  $\epsilon$ -far from a state  $\varrho \in \mathcal{D}$ :

**Proposition 4.** *Let  $\varrho \in \mathcal{D}$  be arbitrary and let  $W = \frac{\mathbb{1}}{d^2} - \varrho$ . For all quantum states  $\sigma$  on  $\mathcal{H}$ , if  $\text{tr}(\sigma W) \geq -\epsilon \|W\|_{\infty}$ , then  $\|\varrho - \sigma\|_1 \geq \epsilon$ .*

*Proof.* Note that

$$\text{tr}(\varrho W) = \frac{1}{d^2} - \text{tr}(\varrho^2) = \frac{1}{d^2} - \frac{1 + 4\epsilon^2}{d^2} = -\frac{4\epsilon^2}{d^2}, \quad \text{and} \quad \|W\|_{\infty} = \left\| \frac{\mathbb{1}}{d^2} - D_{\epsilon} \right\|_{\infty} = \frac{2\epsilon}{d^2}.$$

By Hölder's inequality for matrices,  $\text{tr}((\sigma - \varrho)W) \leq \|\sigma - \varrho\|_1 \cdot \|W\|_{\infty}$ . Hence,

$$\|\sigma - \varrho\|_1 \geq \frac{\text{tr}(\sigma W) - \text{tr}(\varrho W)}{\|W\|_{\infty}} = 2\epsilon + \frac{\text{tr}(\sigma W)}{\|W\|_{\infty}}.$$

When  $\sigma = |x\rangle\langle x|$  with  $x \in \mathcal{H}$  and  $\varrho = UD_{\epsilon}U^{\dagger}$ , we have

$$\text{tr}(|x\rangle\langle x|W) = \langle x|W|x\rangle = \langle x| \left( \frac{\mathbb{1}}{d^2} - UD_{\epsilon}U^{\dagger} \right) |x\rangle = \langle x|U \left( \frac{\mathbb{1}}{d^2} - D_{\epsilon} \right) U^{\dagger} |x\rangle = \|W\|_{\infty} \cdot \langle x|UZU^{\dagger}|x\rangle, \quad (1)$$

where  $Z = \text{diag}(-1, \dots, -1, 1, \dots, 1)$  is just  $\mathbb{1}/d^2 - D_{\epsilon}$  divided by  $\|W\|_{\infty}$ . Hence,  $\|\varrho - |x\rangle\langle x|\|_1 \geq \epsilon$  holds if  $\langle x|UZU^{\dagger}|x\rangle \geq -\epsilon$ .

Since we are interested in the case when  $\varrho = UD_{\epsilon}U^{\dagger}$  is random, it suffices to show that  $\langle x|UZU^{\dagger}|x\rangle$  concentrates in the interval  $[-\epsilon, \epsilon]$ . This fact follows easily from the next lemma:

**Lemma 3.** *Let  $k$  be a positive even integer. If  $\mathbf{u} \in \mathbb{C}^k$  is a uniformly random unit vector, then, for sufficiently large  $k$ ,*

$$\mathbf{P} \left[ |\langle \mathbf{u}|Z|\mathbf{u} \rangle| \geq \frac{1}{2}ck^{-1/4} \right] \leq 4\exp(-\sqrt{k}c^2/8),$$

where  $Z = \text{diag}(1, \dots, 1, -1, \dots, -1)$  is a  $k \times k$  diagonal matrix with  $\text{tr}(Z) = 0$  and  $c$  may be any positive constant.

*Proof.* Let  $\mathbf{u} = (\mathbf{a}_1 + i\mathbf{b}_1, \dots, \mathbf{a}_k + i\mathbf{b}_k) \in \mathbb{C}^k$  be a uniformly random unit vector with  $\mathbf{a}_1, \dots, \mathbf{a}_k, \mathbf{b}_1, \dots, \mathbf{b}_k \in \mathbb{R}$  and let  $\mathbf{v} \in \mathbb{R}^{2k}$  be defined by

$$\mathbf{v} = (\mathbf{a}_1, \dots, \mathbf{a}_{\frac{k}{2}}, \mathbf{b}_1, \dots, \mathbf{b}_{\frac{k}{2}}, \mathbf{a}_{\frac{k}{2}+1}, \dots, \mathbf{a}_k, \mathbf{b}_{\frac{k}{2}+1}, \dots, \mathbf{b}_k).$$

Let  $D$  be the  $2k \times 2k$  diagonal matrix  $D = \text{diag}(1, \dots, 1, -1, \dots, -1)$  with  $\text{tr}(D) = 0$ . Thus,  $\mathbf{v}$  is a uniformly random real unit vector such that  $\langle \mathbf{v} | D | \mathbf{v} \rangle = \langle \mathbf{u} | Z | \mathbf{u} \rangle$ .

Let  $\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{y}_1, \dots, \mathbf{y}_k \in \mathbb{R}$  be  $2k$  standard Gaussian random variables. Let  $\mathbf{X} = \mathbf{x}_1^2 + \dots + \mathbf{x}_k^2$  and  $\mathbf{Y} = \mathbf{y}_1^2 + \dots + \mathbf{y}_k^2$ . By the rotational symmetry of multivariate Gaussian random variables,  $\mathbf{v}$  has the same distribution as

$$\frac{(\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{y}_1, \dots, \mathbf{y}_k)}{\sqrt{\mathbf{X} + \mathbf{Y}}}.$$

Hence,  $\langle \mathbf{v} | D | \mathbf{v} \rangle$  and  $\frac{\mathbf{X} - \mathbf{Y}}{\mathbf{X} + \mathbf{Y}}$  have the same distribution. Since  $\mathbf{X}$  and  $\mathbf{Y}$  are independent  $\chi^2$  random variables with  $k$  degrees of freedom each, it holds that (see e.g. [Wai19, Example 2.11])

$$\mathbf{P} \left[ \left| \frac{\mathbf{X}}{k} - 1 \right| \geq t \right] \leq 2 \exp(-kt^2/8),$$

for all  $t \in (0, 1)$  and similarly for  $\mathbf{Y}$ . Hence, for  $t = ck^{-1/4}$ , we have  $\mathbf{P}[|\mathbf{X} - k| \geq ck^{3/4}] \leq 2 \exp(-\sqrt{k}c^2/8)$ .

If  $|\mathbf{X} - k| < ck^{3/4}$  and  $|\mathbf{Y} - k| < ck^{3/4}$ , then, for  $k$  sufficiently large,

$$|\langle \mathbf{v} | D | \mathbf{v} \rangle| = \frac{|\mathbf{X} - \mathbf{Y}|}{\mathbf{X} + \mathbf{Y}} \leq \frac{2ck^{3/4}}{2k - 2ck^{3/4}} = \frac{c}{k^{1/4} - 1} < \frac{1}{2}ck^{-1/4}.$$

Hence,  $\mathbf{P}[|\langle \mathbf{v} | D | \mathbf{v} \rangle| < \frac{1}{2}ck^{-1/4}] \geq 1 - 4 \exp(-\sqrt{k}c^2/8)$ .

If  $\mathbf{U}$  is a random unitary distributed according to the Haar measure on  $\mathbf{U}(\mathcal{H})$  and  $x \in \mathcal{H}$  is a fixed unit vector, then  $\mathbf{u} = \mathbf{U}|x\rangle$  is a uniformly random unit vector in  $\mathcal{H}$ . Hence, we can apply [Lemma 3](#) to  $|\langle \mathbf{u} | Z | \mathbf{u} \rangle|$  to get

$$\mathbf{P}[|\langle x | \mathbf{U} Z \mathbf{U}^\dagger | x \rangle| \geq \epsilon] \leq 4 \exp(-dc^2/8), \quad (2)$$

where  $c$  is an arbitrary positive constant and  $\epsilon \geq \frac{1}{2}cd^{-1/2}$ .

We now have all the elements needed to prove [Lemma 1](#):

*Proof (Proof of [Lemma 1](#)).* Let  $\mathbf{g} = \mathbf{U} D_\epsilon \mathbf{U}^\dagger$  be a uniformly random element of  $\mathcal{D}$  and let  $\mathbf{W} = \frac{1}{d^2} - \mathbf{g}$ . Thus, assuming  $\epsilon \geq cd^{-1/2}$ ,

$$\begin{aligned} & \mathbf{P}[\forall \sigma \in \text{Sep}, d_{\text{TV}}(\mathbf{g}, \sigma) \geq \epsilon] \\ &= \mathbf{P}[\forall \sigma \in \text{Sep}, \|\mathbf{g} - \sigma\|_1 \geq 2\epsilon] \\ &\geq \mathbf{P}[\forall \sigma \in \text{Sep}, \text{tr}(\sigma \mathbf{W}) \geq -2\epsilon\|\mathbf{W}\|_\infty] && \text{(by [Proposition 4](#))} \\ &\geq \mathbf{P}[\forall \sigma \in 2 \text{conv}(\mathcal{N} \cup -\mathcal{N}), \text{tr}(\sigma \mathbf{W}) \geq -2\epsilon\|\mathbf{W}\|_\infty] && \text{(by [Lemma 2](#))} \\ &= \mathbf{P}[\forall |x\rangle\langle x| \in \mathcal{N} \cup -\mathcal{N}, 2 \text{tr}(|x\rangle\langle x| \mathbf{W}) \geq -2\epsilon\|\mathbf{W}\|_\infty] && \text{(by convexity)} \\ &= \mathbf{P}[\forall |x\rangle\langle x| \in \mathcal{N}, |\langle x | \mathbf{U} Z \mathbf{U}^\dagger | x \rangle| \leq \epsilon] && \text{(by [Equation \(1\)](#))} \end{aligned}$$

By a union bound, this probability is at least

$$1 - \sum_{|x\rangle\langle x| \in \mathcal{N}} \mathbf{P}\left[|\langle x|UZU^\dagger|x\rangle| > \epsilon\right] \geq 1 - |\mathcal{N}| \cdot 4\exp(-dc^2/8) = 1 - 4\exp(d(\log C - c^2/8)),$$

where the first inequality was by Equation (2). Hence, if  $c = \sqrt{8(\log C + 1)}$ , then

$$\mathbf{P}[\forall \sigma \in \text{Sep}, d_{\text{TV}}(\boldsymbol{\rho}, \sigma) \geq \epsilon] \geq 1 - 4\exp(d(\log C - c^2/8)) = 1 - 4\exp(-d) \geq \frac{2}{3},$$

for  $d \geq \log 12$ .

## References

- AS17. Guillaume Aubrun and Stanisław Szarek. *Alice and Bob Meet Banach*. American Mathematical Society, 2017. 5, 10
- Dia77. Persi Diaconis. Finite forms of de Finetti's theorem on exchangeability. *Synthese*, 36(2):271–281, October 1977. 2
- DPS04. Andrew C. Doherty, Pablo A. Parrilo, and Federico M. Spedalieri. Complete family of separability criteria. *Physical Review A*, 69(2), Feb 2004. 2
- GB02. Leonid Gurvits and Howard Barnum. Largest separable balls around the maximally mixed bipartite quantum state. *Physical Review A*, 66(6), Dec 2002. 9
- GM12. Bernd Gärtner and Jiří Matoušek. *Approximation Algorithms and Semidefinite Programming*. Springer Berlin Heidelberg, 2012. 3
- GT09. Otfried Gühne and Géza Tóth. Entanglement detection. *Physics Reports*, 474(1-6):1–75, Apr 2009. 2
- Gur04. Leonid Gurvits. Classical complexity and quantum entanglement. *Journal of Computer and System Sciences*, 69(3):448–484, Nov 2004. 1
- HHHH09. Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Reviews of Modern Physics*, 81(2):865–942, Jun 2009. 2
- HHJ<sup>+</sup>16. Jeongwan Haah, Aram W. Harrow, Zhengfeng Ji, Xiaodi Wu, and Nengkun Yu. Sample-optimal tomography of quantum states. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing - STOC 2016*. ACM Press, 2016. 1
- MdW16. Ashley Montanaro and Ronald de Wolf. A survey of quantum property testing. *Theory of Computing*, 1(1):1–81, 2016. 2
- MM62. John E. Maxfield and Henryk Minc. On the matrix equation  $X'X = A$ . *Proceedings of the Edinburgh Mathematical Society*, 13(02):125, 1962. 3
- OW15. Ryan O'Donnell and John Wright. Quantum spectrum testing. In *Proceedings of the 47th Annual ACM SIGACT Symposium on Theory of Computing - STOC 2015*. ACM Press, 2015. 2, 9
- OW16. Ryan O'Donnell and John Wright. Efficient quantum tomography. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing - STOC 2016*. ACM Press, 2016. 1
- Pan08. Liam Paninski. A coincidence-based test for uniformity given very sparsely sampled discrete data. *IEEE Transactions on Information Theory*, 54(10):4750–4755, Oct 2008. 2
- Ska13. Matthew Skala. Hypergeometric tail inequalities: ending the insanity. 2013. Technical report, arXiv:1311.5939. 8
- Wai19. Martin J. Wainwright. *High-Dimensional Statistics*. Cambridge University Press, 2019. 11
- Wat18. John Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018. 4