Dimension Independent Disentanglers from Unentanglement and Applications

Fernando Granha Jeronimo
Institute for Advanced Studies, Princeton, NJ, USA
Simons Institute, Berkeley, CA, USA

Pei Wu
Weizmann Institute of Science, Rehovot, Israel

Abstract

Quantum entanglement, a distinctive form of quantum correlation, has become a key enabling ingredient in diverse applications in quantum computation, complexity, cryptography, etc. However, the presence of unwanted adversarial entanglement also poses challenges and even prevents the correct behaviour of many protocols and applications.

In this paper, we explore methods to "break" the quantum correlations. Specifically, we construct a *dimension-independent* k-partite disentangler (like) channel from bipartite unentangled input. In particular, we show: For every $d, \ell \geq k \in \mathbb{N}^+$, there is an efficient channel $\Lambda\colon \mathbb{C}^{d\ell} \otimes \mathbb{C}^{d\ell} \to \mathbb{C}^{dk}$ such that for every bipartite separable density operator $\rho_1 \otimes \rho_2$, the output $\Lambda(\rho_1 \otimes \rho_2)$ is close to a k-partite separable state. Concretely, for some distribution μ on states from $\mathbb{C}^d$,

$$\left\| \Lambda(\rho_1 \otimes \rho_2) - \int |\psi\rangle\langle\psi|^{\otimes k} d\mu(\psi) \right\|_1 \leq \tilde{O}\left(\left(\frac{k^3}{\ell} \right)^{1/4} \right).$$

Moreover, $\Lambda(|\psi\rangle\langle\psi|^{\otimes \ell} \otimes |\psi\rangle\langle\psi|^{\otimes \ell}) = |\psi\rangle\langle\psi|^{\otimes k}$. Without the bipartite unentanglement assumption, the above bound is conjectured to be impossible and would imply QMA(2) = QMA.

Leveraging multipartite unentanglement ensured by our disentanglers, we achieve the following: (i) a new proof that QMA(2) admits arbitrary gap amplification; (ii) a variant of the swap test and product test with improved soundness, addressing a major limitation of their original versions. More importantly, we demonstrate that unentangled quantum proofs of almost general real amplitudes capture NEXP, thereby greatly relaxing the non-negative amplitudes assumption in the recent work of $\mathrm{QMA}^+(2) = \mathrm{NEXP}$ [Jeronimo and Wu, STOC 2023]. Specifically, our findings show that to capture NEXP, it suffices to have unentangled proofs of the form $|\psi\rangle = \sqrt{a}|\psi_+\rangle + \sqrt{1-a}|\psi_-\rangle$ where $|\psi_+\rangle$ has non-negative amplitudes, $|\psi_-\rangle$ only has negative amplitudes and $|a - (1-a)| \geq 1/\mathrm{poly}(n)$ with $a \in [0, 1]$. Additionally, we present a protocol achieving an almost largest possible completeness-soundness gap before obtaining $\mathrm{QMA}^{\mathbb{R}}(k) = \mathrm{NEXP}$, namely, a $1/\mathrm{poly}(n)$ additive improvement to the gap results in this equality.

2012 ACM Subject Classification Theory of computation → Interactive proof systems; Theory of computation → Quantum information theory

Keywords and phrases QMA(2), disentangler, quantum proofs

Digital Object Identifier 10.4230/LIPIcs.CCC.2024.26

Related Version *Full Version*: https://arxiv.org/abs/2402.15282

Funding This material is based on work supported by the National Science Foundation under Grant No. CCF-1900460. Part of the work is done when P.W. was at IAS and the Simons Institute.

1 Introduction

Quantum entanglement is a fundamental form of quantum correlation that can be stronger than any classical correlation [13, 5, 11, 21]. It plays a crucial role in a myriad of areas such as quantum computing, quantum information, quantum complexity, quantum cryptography,

condensed matter physics, etc [19, 25, 32]. Hence, comprehending both the capabilities and constraints of quantum entanglement stands as a crucial research endeavor. However, entanglement can also pose challenges in numerous applications, such as quantum key distribution and quantum proof systems [28, 23, 26, 15]. This raises the natural question of designing quantum channels that convert quantum states into unentangled states. For the purpose of applications, such channel, also called *disentangler*, $\Phi : \mathcal{H} \to \mathcal{K} \otimes \mathcal{K}$ can be defined to satisfy two conditions: (i) for any $|\psi\rangle \in \mathcal{K}$, there is preimage $|\phi\rangle$, such that $\Phi(\phi) = \psi \otimes \psi$; and (ii) for any density operator $\phi \in \mathcal{H}$, $\Phi(\phi)$ is close to *separable*.

The *quantum de Finetti* type theorems [10, 24, 28] provide examples of disentanglers. A quantum de Finetti theorem quantifies the closeness of a *permutation-invariant* ℓ-partite quantum state, to k-partite separable states when all but k subsystems are traced out. A standard quantum de Finetti theorem reads

▶ **Theorem 1** (Quantum de Finetti [24]). *For every $d, \ell \geq k \in \mathbb{N}^+$, the channel $\Lambda\colon (\mathbb{C}^d)^{\otimes \ell} \to (\mathbb{C}^d)^{\otimes k}$ defined as $\Lambda(\rho) = \mathrm{Tr}_{\ell-k}(1/\ell! \sum_{\pi \in \mathrm{Sym}_\ell} \pi\rho\pi^\dagger)$ satisfies*

$$\left\| \Lambda(\rho) - \int |\psi\rangle\langle\psi|^{\otimes k} d\mu \right\|_1 \leq \frac{2kd^2}{\ell}.$$

Note that the error bound scales at least[1] as d/ℓ, and in this version of the quantum de Finetti theorem, the parameters are known to be essentially tight. Consequently, if each subsystem is composed of n qubits, then obtaining a non-trivial error bound requires at least $\ell \geq d = 2^n$ subsystems, making this channel impractical for many applications. This is conjectured to be essentially the best you can achieve. In particular, it is conjectured that for any disentangler, the input dimension will be exponential in the output dimension [1] to achieve that the output is always ε close in trace distance to some separable states for any constant $\varepsilon < 1$.

Dimension Independent Disentangler from Unentanglement

While the original disentangler conjecture remains widely open, in this work, we show that there is an explicit, efficient (BQP), and *dimension independent* quantum disentangler for k-partite (output) system starting from a bipartite unentangled system. More precisely, we prove

▶ **Theorem 2** (Disentangler from unentanglement). *Let $d, \ell \geq k \in \mathbb{N}^+$. There is an efficient channel $\Lambda\colon (\mathbb{C}^d)^{\otimes \ell} \otimes (\mathbb{C}^d)^{\otimes \ell} \to (\mathbb{C}^d)^{\otimes k}$ such that for any density operators $\rho_1, \rho_2 \in \mathbb{C}^{d\ell}$ there is a distribution μ on pure states $|\psi\rangle \in \mathbb{C}^d$ satisfying*

$$\left\| \Lambda(\rho_1 \otimes \rho_2) - \int |\psi\rangle\langle\psi|^{\otimes k} d\mu \right\|_1 \leq \tilde{O}\left(\left(\frac{k^3}{\ell} \right)^{1/4} \right).$$

Furthermore, product states of the form $\rho_1 = \rho_2 = |\psi\rangle\langle\psi|^{\otimes \ell}$ are mapped to $|\psi\rangle\langle\psi|^{\otimes k}$.

In contrast to the de Finetti disentangler, our disentangler from unentanglement features error parameters that are independent of the input dimension entirely! Subsequently, we discuss applications of Theorem 2 in testing product states and the gap amplification in

[1] If instead of making the state permutation invariant, we project it onto the symmetric subspace, which is a perfectly valid and efficient operation in the quantum setting, then the dependence on d in Theorem 1 improves from d^2 to d.

quantum proof systems, culminating in a near-optimal gap amplification for the $\mathrm{QMA}^{+}(k)$ class: Any improvement on this gap amplification would imply $\mathrm{QMA}^{\mathbb{R}}(k) = \mathrm{NEXP}$.[2] We also anticipate that our tool will find further applications beyond those discussed in this paper.

1.1 Super Product Test

The *product test* was designed to test if a state $|\phi\rangle$ is close to k-partite product state, i.e., $|\phi\rangle \approx |\phi_1\rangle \otimes \cdots \otimes |\phi_k\rangle$, given two copies of $|\phi\rangle$. This test involves applying a sequence of swap tests to each of the k subsystems of the two copies $|\phi\rangle$. Clearly, if $|\phi\rangle$ is indeed a k-partite product state, all the swap tests accept with certainty. On the other hand, if $|\phi\rangle$ is entangled across the k subsystems, some swap test will reject with a probability that depends on the amount of entanglement. It can be argued that the product test is optimal for ensuring perfect completeness, i.e., accepting product states with certainty [17].

Despite its utility and elegance, the product test has two limitations. Firstly, it only provides a guarantee concerning its input $|\phi\rangle \otimes |\phi\rangle$ which are destroyed after the test, yielding a single classical bit as output. Very often in applications, one also needs some extra certified input states $|\phi\rangle$ to manipulate in subsequent computations after the test. Secondly, and probably more irritatingly, the product test always accepts with some constant probability (say $\geq 1/2$) no matter how far $|\phi\rangle$ is from being k-partite product, i.e., it has poor soundness. These limitations can be resolved if you have more than 2 copies of $|\phi\rangle$ [22, 29]. For instance, given ℓ copies of $|\phi\rangle$, then one can adapt the product test to sequentially apply projections on to symmetric subspace on the first, second, and subsequent subsystems of all the copies of ϕ. Intuitively, this should give us a stronger test whose analysis was left as an open problem in [17]. Recently, She and Yuen [29] analyzed this higher order version of product test achieving improved soundness. We restate this higher order product test as relying on some ℓ unentangled equal copies of $|\psi\rangle$ to deduce a k-partite product structure of the input state. One can require something even stronger on the input to achieve what we call *super product test*.

▶ **Lemma 3.** *The super product test on input* $|\psi\rangle \otimes (|\phi_1\rangle \dots |\phi_k\rangle)^{\otimes \ell}$ *accepts with probability*

$$\frac{\ell}{(\ell+1)} \cdot |\langle \psi \,|\, \phi_1\rangle \dots |\phi_k\rangle|^2 + \frac{1}{(\ell+1)}\,.$$

This super product test focuses on determining whether a target state $|\psi\rangle$ is a product state or not. In addition to the target state, there are ℓ copies of an already k-partite product state that come to help. This test is very natural and simple, except it seems to ask too much of its inputs: To compare, the high-order product test requires some copies of a state whereas Lemma 3 requires some copies of an already k-partite product state of the form $|\phi_1\rangle \otimes \cdots \otimes |\phi_k\rangle$. We claim the super product test is not really asking for too much because our disentangler channel effectively "amplifies" the number of unentangled systems. In particular, we can rely on just two unentangled proofs to enforce a state close to $(|\phi_1\rangle \otimes \cdots \otimes |\phi_k\rangle)^{\otimes \ell}$ by Theorem 2. For simplicity, consider k unentangled pairs of untangled proofs where the i^{th} pair applied Theorem 2 yields $|\phi_i\rangle^{\otimes \ell}$. Then run the super product test on a target state

[2] We don't want to distract the readers by the issue about quantum states over real or complex numbers. In many cases, quantum computation over reals captures that over complex numbers. However, to the best of the authors' knowledge, this is unclear in the context of QMA(2). We have to use $\mathrm{QMA}^{\mathbb{R}}(k)$ to denote the proof systems where the proofs are guaranteed to have real amplitudes.

$|\psi\rangle$ and the ℓ copies of already product states from our disentangler. Furthermore, note that it is very cheap to instead enforce a state close to $(|\phi_1\rangle \otimes \cdots \otimes |\phi_k\rangle)^{\otimes 2\ell}$, allowing us to reserve the extra ℓ copies of $|\phi_1\rangle \otimes \cdots \otimes |\phi_k\rangle$ as once the super product test passes, they can be used in any other computations as a very good proxy of $|\psi\rangle$. With this combination, we achieve arbitrarily good soundness without requiring more than $2k$ unentangled states[3] while obtaining a guarantee about the output, rather than having just a single classical bit of output.

1.2 A Gap Amplification for QMA$^+$(2) up to Criticality

Next, we turn to the unentangled quantum proofs, the so-called QMA(2) class [23] and its variants. First, we provide some background on this subject.

The complexity of QMA(2) was shown to be closely related to a variety of quantum and classical computational problems, e.g., determining if a mixed state is entangled given its classical description, as well as, various forms of classical polynomial/tensor optimization (see [17] for a more comprehensive list). Despite considerable interest and effort (e.g., [12, 1, 6, 4, 7, 14, 30, 27, 9, 8, 18]), we still only know the trivial complexity bounds QMA $\subseteq$ QMA(2) $\subseteq$ NEXP.

Even the fact that QMA(2) admits strong gap amplification is non-trivial and remained open for about 10 years before the seminar work of Harrow and Montanaro [17]. With Theorem 2, it is easy to give a new proof of this fact.

A variant of QMA(2), denoted QMA$^+$(2), with proofs of nonnegative amplitudes was introduced by Jeronimo and Wu in [20]. The goal of this variant was to capture many properties of QMA(2) while having more structure in order to obtain a greater understanding. Indeed, they showed that QMA$^+$(2) = NEXP by designing a QMA$^+$(2) protocol for a NEXP-complete problem with a constant gap. On the other end of their result is the observation that QMA$^+$(2) $\subseteq$ QMA(2) provided that the completeness-soundness gap of QMA$^+$(2) is a sufficiently large constant. This makes QMA$^+$(2) an intriguing class to study since either (i) showing that QMA$^+$(2) = QMA(2), via possibly a gap amplification approach for QMA$^+$(2), would characterize the complexity of QMA(2), or (ii) showing QMA$^+$(2) $\neq$ QMA(2) would give a better upper bound QMA(2) $\subsetneq$ NEXP.

By virtue of the unentanglement assumption of QMA$^+$(2) and the product test [17], QMA$^+$(2) admits some non-trivial gap amplification. For example, a gap of $1/\mathrm{poly}(n)$ can be amplified to a constant gap in which the completeness becomes $1 - \exp(-\mathrm{poly}(n))$ and the soundness becomes some constant strictly less than 1. Recently, Bassirian, Fefferman and Marwaha [3], building on [20], curiously showed that QMA$^+$(1) = NEXP also with a constant gap.[4] Since in the large constant gap regime of QMA$^+$(1), we have QMA$^+$(1) = QMA $\subseteq$ PP, their result rules out the strong gap amplification for QMA$^+$(1) unless NEXP $\subseteq$ PP. Moreover, it also suggests that strategies aimed at amplifying the gap for QMA$^+$(2) must rely on the unentanglement assumption. This is precisely where the tools like the product test or our disentangler become essential.

With our disentangler, we make progresses towards understanding of QMA$^+$(2) versus QMA(2). In particular, our progresses can be summarized as two aspects with two motivating questions.

[3] Naturally, the $2k$ unentangled states need to get larger in dimension to achieve better soundness.
[4] It is not clear that their gap can be made as large as the one for QMA$^+$(2) = NEXP.

Motivating question 1. How crucial is the nonnegative amplitudes assumption to obtain $\mathrm{QMA}^+(2) = \mathrm{NEXP}$*?*

Regarding our first motivating question, we show that the nonnegative amplitudes assumption can be almost completely removed by considering unentangled quantum proofs of almost general *real* amplitudes. More precisely, we show that to capture NEXP it suffices to have unentangled proofs of the form $|\psi\rangle = \sqrt{a}|\psi_+\rangle + \sqrt{1-a}|\psi_-\rangle$ where $|\psi_+\rangle$ has nonnegative amplitudes, $|\psi_-\rangle$ only has negative amplitudes and $|a-(1-a)| \geq 1/\mathrm{poly}(n)$ with $a \in [0,1]$. In words, we require the proofs to have slightly more ℓ_2-probability mass ($1/\mathrm{poly}(n)$ extra mass) either on nonnegative or negative amplitudes. We refer to the quantity $|a-(1-a)|$ as the ℓ_2-sign bias of $|\psi\rangle$. We call the associated complexity class almost-$\mathrm{QMA}^{\mathbb{R}}(k)$. Our main complexity result can be stated as follows.

▶ **Theorem 4.** $\mathrm{NEXP} = \text{almost-}\mathrm{QMA}^{\mathbb{R}}(k)$ *with unentangled proofs of* ℓ_2*-sign bias of*[5] $b(n) \geq \mathrm{poly}(1/n)$ *and* $k = \mathrm{poly}(1/b(n))$.

We obtain the above result by investigating the other motivating question: Since the power of $\mathrm{QMA}^+(k)$ ranges from NEXP to $\mathrm{QMA}(k)$ depending on the gap,

Motivating question 2. How much can we amplify the gap of $\mathrm{QMA}^+(k)$*?*

We make significant progress addressing this question. Specifically, we show that a even more relaxed version of $\mathrm{QMA}^+(3)$, featuring a single proof with nonnegative amplitudes and the other two with general amplitudes, equals NEXP, with completeness $1-\exp(-\mathrm{poly}(n))$ and soundness $1/2+1/\mathrm{poly}(n)$. At the first glance, this looks like a "just so so" gap amplification. It is noteworthy that achieving a slightly improved soundness of $1/2-1/\mathrm{poly}(n)$ would imply $\mathrm{QMA}^{\mathbb{R}}(3) = \mathrm{NEXP}$. In particular, if $\mathrm{QMA}^{\mathbb{R}}(3) \neq \mathrm{NEXP}$, then there is a sharp phase transition in the complexity around the gap of a half.

▶ **Theorem 5.** $\mathrm{NEXP} = \mathrm{QMA}^+(3)$ *with completeness* $c = 1-\exp(-\mathrm{poly}(n))$ *and soundness* $s = 1/2+1/\mathrm{poly}(n)$*. Furthermore, we can assume a particular case of* $\mathrm{QMA}^+(3)$ *in which two unentangled proofs have arbitrary amplitudes whereas only one unentangled proof has nonnegative amplitudes.*

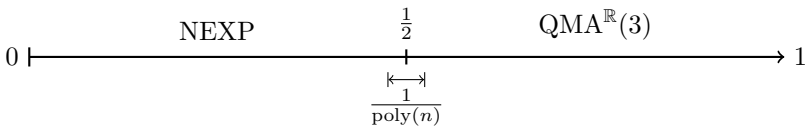

Figure 1 Gap and the complexity regime of the particular version of $\mathrm{QMA}^+(3)$ from Theorem 5. A gap below $1/2-1/\mathrm{poly}(n)$ corresponds to NEXP, whereas a gap above $1/2+1/\mathrm{poly}(n)$ corresponds to $\mathrm{QMA}^{\mathbb{R}}(3)$, illustrating a sharp phase transition.

1.3 Organization

We introduce notations and review basic concepts and facts in Section 2. In Section 3, we present an efficient multipartite disentangler (like) channel from bipartite unentanglement. This construction relies on new de Finetti type properties concerning the interplay between

[5] The letter n represents the input size and $b(n)$ is any polynomial time computable function bounded from below by a polynomial, i.e., by $1/n^c$ for some constant $c > 0$.

entanglement and symmetry which we explore in Section 4. In Section 5, we delve into the utility of our disentangler where we elaborate a generic framework in the context of property testing. As one example, we present a new proof that QMA(2) admits strong gap amplification. The final two sections are devoted to design new tests and derive the main complexity results in this paper. In Section 6, we present the super swap and super product test which leverage unentanglement to achieve much improved soundness than the well-known swap and product tests. Finally, we provide protocols for NEXP in Section 7 leading to the main complexity results of this paper, Theorem 4 and Theorem 5.

2 Preliminaries

General

As usual, $\mathbb{N}, \mathbb{R}, \mathbb{C}$ stand for the natural, real, and complex numbers, respectively. We adopt the Dirac notation for vectors representing quantum states, e.g., $|\psi\rangle, |\phi\rangle$, etc. In this paper, all the vectors of the form $|\psi\rangle$ are unit vectors. Given any pure state $|\psi\rangle$, we adopt the convention that its density operator is denoted by the Greek letter without the "ket", e.g. $\psi = |\psi\rangle\langle\psi|$. The set of density operators in an arbitrary Hilbert space $\mathcal{H}$ is denoted $\mathcal{D}(\mathcal{H})$. A symmetric state $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ is that invariant under any permutation $\pi \in \mathrm{Sym}_n$ where Sym_n is the symmetric group. The action of π on $(\mathbb{C}^d)^{\otimes n}$ is

$$\pi : |\psi_1, \psi_2, \ldots, \psi_n\rangle \mapsto |\psi_{\pi(1)}, \psi_{\pi(2)}, \ldots, \psi_{\pi(n)}\rangle.$$

The *symmetric subspace* is the subspace of $(\mathbb{C}^d)^{\otimes n}$ that is invariant under Sym_n, denoted by $\vee^n(\mathbb{C}^d)$. Given any set $H \subseteq \mathcal{H}$ for some Hilbert space $\mathcal{H}$, $\mathrm{conv}\,(H)$ is the convex hull of H.

One other particularly interesting set of states is the *separable* states. We adopt the following notation for the set of density operators regarding separable states,

$$\mathrm{SEP}(d, r) \coloneqq \mathrm{conv}\left(\psi_1 \otimes \cdots \otimes \psi_r \mid |\psi_1\rangle, \ldots, |\psi_r\rangle \in \mathbb{C}^d\right).$$

A related notion is that of separable measurement, whose formal definition is given below.

▶ **Definition 6** (Separable measurement). *A measurement $M = (M_0, M_1)$ is separable if in the yes case, the corresponding positive semi-definite matrix M_1 can be represented as a conical combination of two operators acting on the first and second parts, i.e., for some distribution μ over the tensor product of positive semi-definite matrices α and β on the corresponding space,*

$$M_1 = \int \alpha \otimes \beta \, \mathrm{d}\mu.$$

We record the following well-known fact. An interested reader is referred to [16] for a formal proof.

▶ **Fact 7** (Folklore). *The swap test is separable.*

Matrix Analysis

Given any matrix $M \in \mathbb{C}^{n\times n}$, $M^\dagger$ is its conjugate transpose. Let $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$ denote its singular values. Then the trace norm $\|\cdot\|_1$, Frobenius norm $\|\cdot\|_F$ are defined as below

$$\|M\|_1 = \sum_i \sigma_i, \qquad \|M\|_F = \sqrt{\sum_i \sigma_i^2}.$$

The Frobenius norm also equals the square root of sum of squared modulus of each entry, i.e., $\|M\|_F = \sqrt{\sum_{i,j} |M(i,j)|^2}$.

For a positive semi-definite (PSD) matrix M, $\|M\|_F = \sqrt{\operatorname{Tr} M^2}$. For two PSD matrices, there is one (of many) analogous matrix Cauchy-Schwarz inequality.

$$\operatorname{Tr}(\sigma\rho) \leq \|\sigma\|_F \cdot \|\rho\|_F. \tag{2.1}$$

We adopt the notation $\succeq$ to denote the partial order that $\sigma \succeq \rho$ if $\sigma - \rho$ is positive semi-definite.

Distances between Quantum States

A standard notion of distance for quantum states is that of the *trace distance*. The trace distance between ψ and ϕ, denoted $\mathrm{D}(\psi, \phi)$, is

$$\frac{1}{2}\|\psi - \phi\|_1 = \frac{1}{2}\operatorname{Tr}\sqrt{(\psi-\phi)^\dagger(\psi-\phi)}. \tag{2.2}$$

We also use the notation $\mathrm{D}(|\psi\rangle, |\phi\rangle)$ if we want to emphasize that ψ and ϕ are pure states. The following fact provides an alternative definition for trace distance between pure states.

▶ **Fact 8.** *The trace distance between $|\phi\rangle$ and $|\psi\rangle$ is given by* $\mathrm{D}(|\phi\rangle, |\psi\rangle) = \sqrt{1 - |\langle\phi \mid \psi\rangle|^2}$.

Two states with small trace distance are indistinguishable to quantum protocols.

▶ **Fact 9.** *If a quantum protocol accepts a state ϕ with probability at most p, then it accepts ψ with probability at most $p + \mathrm{D}(\phi, \psi)$.*

Trace distance enjoys the triangle inequality. For pure states, we can actually strengthen it.

▷ Claim 10. Given unit vectors $|\alpha\rangle, |\phi\rangle, |\beta\rangle \in \mathcal{H}$ for some Hilbert space $\mathcal{H}$. Suppose

$$|\langle\alpha \mid \phi\rangle|^2 = 1 - \varepsilon, \qquad |\langle\beta \mid \phi\rangle|^2 = 1 - \delta.$$

Then for any $\varepsilon + \delta \leq 1$,[6]

$$|\langle\alpha \mid \beta\rangle|^2 \geq (\sqrt{(1-\varepsilon)(1-\delta)} - \sqrt{\varepsilon\delta})^2. \tag{2.3}$$

In general, we always have

$$|\langle\alpha \mid \beta\rangle|^2 \geq 1 - \varepsilon - \delta - 2\sqrt{\varepsilon\delta}. \tag{2.4}$$

Proof. Without loss of generality assume that

$$|\alpha\rangle = \sqrt{1-\varepsilon}|\phi\rangle + \sqrt{\varepsilon}|\mu\rangle,$$
$$|\beta\rangle = \sqrt{1-\delta}|\phi\rangle + \sigma\sqrt{\eta}|\mu\rangle + \sqrt{\delta-\eta}|\rho\rangle,$$

where $|\mu\rangle, |\rho\rangle, |\phi\rangle$ are orthogonal, $0 \leq \eta \leq \delta$ and $\sigma \in \mathbb{C}$ is a relative phase. Using the basis $\{|\phi\rangle, |\mu\rangle, |\rho\rangle\}$, we can write down explicitly the density matrix of α and β:

$$\alpha = \begin{pmatrix} 1-\varepsilon & \sqrt{\varepsilon(1-\varepsilon)} & 0 \\ \sqrt{\varepsilon(1-\varepsilon)} & \varepsilon & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\beta = \begin{pmatrix} 1-\delta & \sigma\sqrt{(1-\delta)\eta} & \sqrt{(1-\delta)(\delta-\eta)} \\ \sigma^*\sqrt{(1-\delta)\eta} & \eta & \sigma\sqrt{\eta(\delta-\eta)} \\ \sqrt{(1-\delta)(\delta-\eta)} & \sigma^*\sqrt{\eta(\delta-\eta)} & \delta-\eta \end{pmatrix}.$$

[6] When $\varepsilon + \delta > 1$, then $|\alpha\rangle$ and $|\beta\rangle$ in general can be orthogonal.

Now by definition,

$$\begin{aligned}
\mathrm{D}(|\alpha\rangle,|\beta\rangle)^2 &= \left(\frac{1}{2}\operatorname{Tr}\sqrt{(\alpha-\beta)^\dagger(\alpha-\beta)}\right)^2 \\
&= \frac{1}{2}\|\alpha-\beta\|_F^2 \\
&= \frac{1}{2}((\varepsilon-\delta)^2+(\varepsilon-\eta)^2+(\delta-\eta)^2)+\eta(\delta-\eta) \\
&\qquad + |\sqrt{\varepsilon(1-\varepsilon)}-\sigma\sqrt{(1-\delta)\eta}|^2+(1-\delta)(\delta-\eta) \\
&\le \frac{1}{2}((\varepsilon-\delta)^2+(\varepsilon-\eta)^2+(\delta-\eta)^2)+\eta(\delta-\eta) \\
&\qquad + (\sqrt{\varepsilon(1-\varepsilon)}+\sqrt{(1-\delta)\eta})^2+(1-\delta)(\delta-\eta), \qquad (2.5)
\end{aligned}$$

where the second step holds because $\alpha-\beta$ is Hermitian with trace 0 and rank 0 or 2. We claim that the RHS of (2.5), denote by f, is non-decreasing for $\eta\in[0,\delta]$. By routine calculation,

$$\begin{aligned}
\frac{df}{d\eta} = -\varepsilon+\sqrt{\frac{\varepsilon}{\eta}(1-\varepsilon)(1-\delta)} \ge 0 \quad &\iff \quad (1-\varepsilon)(1-\delta)\ge\eta\varepsilon \\
&\Longleftarrow \quad (1-\varepsilon)(1-\delta)\ge\delta\varepsilon \quad \iff \quad 1\ge\varepsilon+\delta.
\end{aligned}$$

As we assumed that $1\ge\varepsilon+\delta$, $df/d\eta$ is always non-negative. Since the RHS of (2.5) is non-decreasing for $\eta\in[0,\delta]$, plug $\eta=\delta$ into the RHS of (2.5), we obtain

$$\mathrm{D}(|\alpha\rangle,|\beta\rangle)^2 \le (\varepsilon-\delta)^2+(\sqrt{\varepsilon(1-\varepsilon)}+\sqrt{(1-\delta)\delta})^2,$$

In view of Fact 8, (2.3) is proved. The "in general" part is trivially true when $\varepsilon+\delta>1$ and otherwise follows from (2.3). ◁

Another widely used distance measure between quantum states is that of *fidelity*. For any density operators ρ,σ from the same Hilbert space,

$$F(\rho,\sigma)=\left(\operatorname{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2.$$

For our purposes, we only need the fact that when one of the two density operators corresponds to a pure state, then

$$F(\rho,\sigma)=\operatorname{Tr}(\rho\sigma).$$

The well-known data processing inequality for fidelity states that applying quantum operation never decreases the fidelity.

▶ **Fact 11.** *For any quantum channel (CPTP map) Φ,*

$$F(\Phi(\rho),\Phi(\sigma))\ge F(\rho,\sigma).$$

Schmidt Decomposition and Partial Trace

For $|\psi\rangle$ describing quantum states over two subsystems A,B, e.g., $|\psi\rangle\in\mathbb{C}^m\otimes\mathbb{C}^n$, there are two sets of orthonormal states $\{|\alpha_1\rangle,|\alpha_2\rangle,\ldots,|\alpha_k\rangle\}\subseteq\mathbb{C}^m$, $\{|\beta_1\rangle,|\beta_2\rangle,\ldots,|\beta_k\rangle\}\subseteq\mathbb{C}^n$, and positive numbers $\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_k$ for some $k\le\min\{n,m\}$ such that

$$|\psi\rangle=\sum_{i=1}^k\sqrt{\lambda_i}|\alpha_i\rangle|\beta_i\rangle, \qquad \text{and } \sum_{i=1}^k\lambda_i=1. \qquad (2.6)$$

The formula (2.6) is called the *Schmidt decomposition* of $|\psi\rangle$. The set of $\sqrt{\lambda_i}$ is unique, and is called the *Schmidt coefficient* of $|\psi\rangle$. We call $\sqrt{\lambda_1}$ the *top Schmidt coefficient* and $|\alpha_1\rangle|\beta_1\rangle$ the *top Schmidt component*. Note that the top Schmidt component may not be unique ignoring the global phases, in that case we break tie arbitrarily. Since Schmidt decomposition follows from singular value decomposition, the (top) Schmidt coefficients can also be formulated as some optimization problem.

▷ Claim 12. Given any state $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$. Then

$$\lambda_1 = \max_{|\sigma\rangle\in\mathcal{H}_1, |\rho\rangle\in\mathcal{H}_2} |\langle\psi \mid \sigma, \rho\rangle|^2$$

Often we want to study the density operator of a quantum state $|\psi\rangle$ over the subsystem A, mathematically described by tracing out B, denoted $\mathrm{Tr}_B(\psi)$. We also abbreviate $\psi_A = \mathrm{Tr}_B(\psi)$. Note that fidelity never increases under partial trace due to Fact 11, and similarly, the trace distance never increases under partial trace:

▶ **Fact 13.** *For any quantum states ψ and ϕ over systems A and B,*

$$\mathrm{D}(\psi, \phi) \geq \mathrm{D}(\psi_A, \phi_A).$$

We use subscript to emphasize the systems that an operator is describing, e.g., ψ^{AB} simply means that ψ is a state over systems A and B.

Quantum Merlin-Arthur Systems

We now formally define the class almost-$\mathrm{QMA}^{\mathbb{R}}(k)$, but first we will need the ℓ_2-sign bias definition, which, roughly speaking, quantifies the imbalance in ℓ_2 mass between the positive and negative amplitudes parts of a state.

▶ **Definition 14** (ℓ_2-sign bias). *Given $|\psi\rangle \in \mathbb{R}^n$, we can uniquely write it as $|\psi\rangle = \sqrt{a}|\psi_+\rangle + \sqrt{1-a}|\psi_-\rangle$, where $a \in [0,1]$, $|\psi_+\rangle$ and $|\psi_-\rangle$ are unit vectors with only positive and negative amplitudes, respectively. The ℓ_2-sign bias of $|\psi\rangle$ is defined as $|a - (1-a)|$.*

Note that a non-negative amplitude state has ℓ_2-sign bias of 1 whereas a general state has bias at least 0. Almost-$\mathrm{QMA}^{\mathbb{R}}(k)$ will be defined based on ℓ_2-sign as a natural relaxation of $\mathrm{QMA}^{+}(k)$ towards the general $\mathrm{QMA}(k)$.

▶ **Definition 15** (almost-$\mathrm{QMA}^{\mathbb{R}}(k)$). *Let $k\colon \mathbb{N} \to \mathbb{N}$ be a polynomial time computable function. A promise problem $\mathcal{L}_{\mathrm{yes}}, \mathcal{L}_{\mathrm{no}} \subseteq \{0,1\}^*$ is in* almost-$\mathrm{QMA}^{\mathbb{R}}(k)$ *if there exists a* BQP *verifier V such that for every $n \in \mathbb{N}$ and every $x \in \{0,1\}^n$,*

- **Completeness:** *If $x \in \mathcal{L}_{\mathrm{yes}}$, then there exist unentangled states $|\psi_1\rangle, \ldots, |\psi_{k(n)}\rangle$, each of ℓ_2-sign bias $1/\mathrm{poly}(n)$ and on at most $\mathrm{poly}(n)$ qubits, s.t. $\Pr[V(x, |\psi_1\rangle \otimes \cdots \otimes |\psi_{k(n)}\rangle)$ accepts$] \geq 9/10$.*
- **Soundness:** *If $x \in \mathcal{L}_{\mathrm{no}}$, then for every unentangled states $|\psi_1\rangle, \ldots, |\psi_{k(n)}\rangle$, each of each of ℓ_2-sign bias $1/\mathrm{poly}(n)$ and on at most $\mathrm{poly}(n)$ qubits, we have $\Pr[V(x, |\psi_1\rangle \otimes \cdots \otimes |\psi_{k(n)}\rangle)$ accepts$] \leq 1/10$.*

3 The Disentangler from Unentanglement

In this section, we show how to obtain the dimension independent k-partite disentangler (like) channel from bi-partite unentanglement establishing Theorem 2. We will actually work mainly with a more refined procedure which we call quantum probably approximately product output (PAPO) procedure, from which the claimed disentangler can be easily constructed. We define PAPO as follows.

▶ **Definition 16** (PAPO). *Let $d, \ell, k \in \mathbb{N}$ and $\varepsilon, \delta \in [0,1]$. A $(d, \ell, k, \varepsilon, \delta)$-*PAPO *is a quantum procedure Λ satisfying:*

- **Completeness:** $\forall |\psi\rangle \in \mathbb{C}^d$, $\Lambda(\rho_1 \otimes \rho_2) = |\psi\rangle\langle\psi|^{\otimes k}$ *where* $\rho_1 = \rho_2 = |\psi\rangle\langle\psi|^{\otimes \ell}$,
- **Soundness:** $\forall \rho \in \mathrm{SEP}(d^\ell, 2)$, *with probability at least* $1 - \delta$, $\Lambda(\rho)$ *either rejects or outputs a state ε-close in trace distance to a separable state.*

The main result in this section is an efficient PAPO procedure with parameter ℓ that is independent of the dimension d.

▶ **Theorem 17.** *For every $d, k \in \mathbb{N}$ and $\varepsilon, \delta \in [0,1]$, there is an efficient $(d, \ell, k, \varepsilon, \delta)$-*PAPO *with $\ell = O(k^3 \varepsilon^{-2} \delta^{-2} \log \delta^{-1})$.*

In Algorithm 1, we give a detailed description of our PAPO procedure. The procedure takes input two unentangled states, each over ℓ subsystems. We name the ℓ systems $A_1, A_2, \ldots, A_\ell$ for the first state, and $B_1, B_2, \ldots, B_\ell$ for the second state. The PAPO procedure is very simple, which we consider an advantage for such a fundamental task. It should be compared with the product test [17]: the PAPO procedure further takes advantage of symmetric subspace and that projection onto the symmetric subspace is efficient for quantum algorithms.

Algorithm 1 PAPO.

Input: $\rho^{A_1, A_2, \cdots, A_\ell} \otimes \rho^{B_1, B_2, \cdots, B_\ell} \in \mathrm{SEP}(d^\ell, 2)$.

- Sample $\ell' \in [\ell - k]$ uniformly at random.
- For $i = 1, \ldots, \ell'$:
  1. Project $\rho^{A_i, \cdots, A_\ell}$ onto the symmetric space.
  2. Project $\rho^{B_i, \cdots, B_\ell}$ onto the symmetric space.
  3. If any of the projections fails: *Reject*.
  4. If $i \neq \ell'$, SwapTest(ρ^{A_i}, ρ^{B_i}).
  5. If the SwapTest fails: *Reject*.
- Output $\rho^{A_{\ell'}, \cdots, A_{\ell'+k-1}}$.

3.1 Analysis of PAPO

The efficiency of the protocol is trivial. Indeed projection onto the symmetric subspace can be implemented efficiently, see for example [2], and swap test is a special case of projection onto the symmetric subspace. So in the remainder of the section, we argue that our procedure satisfies the completeness and soundness criterion in Definition 16. We start with the following definition of *termination index*.

▶ **Definition 18** (Termination Index). *We set i^* to be the least element in $[\ell - k]$ such that either $\rho^{A_{i^*}, \ldots, A_\ell}$ or $\rho^{B_{i^*}, \ldots, B_\ell}$ is orthogonal to the symmetric subspace; we set $i^* = \infty$ if no such element exists.*

Here the "termination" means absolute termination (rejection) by projection into the symmetric subspace and has nothing to do with a particular execution of Algorithm 1. Most likely, projecting a general state into the symmetric subspace can success or fail. When a state can be successfully projected into the symmetric subspace with nonzero probability, then PAPO continues to run with nonzero probability. Such case is not counted as absolute termination. [7]

[7] Note that the swap test has no danger of absolute termination since it is always applied to separable states in Algorithm 1 and the swap test has soundness 1/2. Thus in the definition of termination index, we don't worry about the swap test.

▷ Claim 19. The state $\rho^{A_i,\ldots,A_\ell,B_i\ldots,B_\ell}$ at the i^{th} iteration of the for loop in Algorithm 1 is separable across $\rho^{A_i,\ldots,A_\ell}$ and $\rho^{B_i,\ldots,B_\ell}$.

Proof. Because the SwapTest is separable across A and B part given it accepts by Fact 7 and projection into the symmetric subspace for A and B part individually is also separable. Therefore $\rho^{A_i\ldots A_\ell B_i\ldots B_\ell}$ is separable across A and B part. ◁

▶ **Definition 20** (Bad Index). *We say that an index $i \in [\ell]$ is* η-bad
1. *If $i \geq i^*$, (see Definition 18)*
2. *or if* $\text{SwapTest}(\rho^{A_i}, \rho^{B_i})$ *accepts with probability at most* $1-\eta$.

▷ Claim 21. $\text{SwapTest}(\rho,\sigma)$ accepts with probability $\frac{1+\text{Tr}(\rho\sigma)}{2} \leq \frac{3}{4} + \frac{\text{Tr}(\sigma^2)}{4}$.

Proof. Apply (2.1) for the density operators,

$$\text{Tr}(\rho\sigma) \leq \sqrt{\text{Tr}(\sigma^2)\cdot\text{Tr}(\rho^2)} \leq \frac{\text{Tr}\,\sigma^2 + \text{Tr}\,\rho^2}{2},$$

where the second step uses the AM-GM inequality. Note that $\text{Tr}\,\rho^2 \leq 1$, we are done. ◁

One more technical tool that we are going to need is the following, whose proof we defer to the next section.

▶ **Theorem 22.** *Given state $\sigma^{A_1\ldots A_k} \in \text{conv}\left(\vee^k(\mathbb{C}^d)\right)$. Then there is some distribution μ on pure states $|\phi\rangle \in \mathbb{C}^d$, such that*

$$\left\|\sigma - \int \phi^{\otimes k} d\mu\right\|_1 \leq O\left(\sqrt{k^3(1-\text{Tr}\,(\sigma_{A_1})^2)}\right).$$

Proof of Theorem 17.

Completeness: For a desired output of $|\psi\rangle\langle\psi|^{\otimes k}$, we give two unentangled copies of $|\psi\rangle^{\otimes\ell}$ to Λ as input. In this case, Algorithm 1 indeed outputs $|\psi\rangle\langle\psi|^{\otimes k}$ w.p. 1.

Soundness: Let $\rho \in \text{SEP}(d^\ell, 2)$ be the input of Λ. Set

$$\eta = \varepsilon^2/k^3.$$

Due to Claim 19, $\rho^{A_{\ell'}\ldots A_\ell, B_{\ell'}\ldots B_\ell}$ is separable just before the ℓ'^{th} iteration (assuming successfully reaching this iteration). For ℓ' that is not a bad index, after projection onto the symmetric subspace, $\rho^{A_{\ell'}\ldots A_{\ell'+k-1}} \in \text{conv}\left(\vee^k(\mathbb{C}^d)\right)$. It follows from Claim 21 that $\text{Tr}_{A_{\ell'}}(\rho^{A_{\ell'}\ldots A_\ell})^2 \geq 1-4\eta$. Thus we conclude that if ℓ' is not a bad index, then the output (if no rejection) is ε-close in trace distance to a convex combination of product states by Theorem 22 and our choice of parameter η. Therefore to prove the theorem, it suffices to bound the probability that Algorithm 1 outputs (not rejects) when ℓ' is a bad index.

Next we consider two cases. The first case: If the number of the η-bad indices among the first $\ell-k$ subsystems are less than $\delta(\ell-k)$, then with probability at least $1-\delta$, the random index ℓ' is not η-bad. Therefore, Definition 16 is satisfied.

The second case: This fraction is larger than δ. Now conditioning on the event that ℓ' is a bad index, then ℓ' is a uniformly random bad index. Therefore, the chance that the set of indices $\{1,2,\ldots,\ell'\}$ contains less than $\delta/2$ fraction of bad indices is at most $\delta/2$. Thus with probability at least $1-\delta/2$, we have seen at least $\delta/2\cdot\delta(\ell-k)-1$ bad indices in the execution of Algorithm 1 in the first ℓ' iterations. Since for each bad index the probability of not rejecting by the swap test is at most $1-\eta$, the total probability of not rejecting is at most

$$(1-\eta)^{\delta^2(\ell-k)-1} = \exp(-\Omega(\eta\delta^2\ell)) = \exp\left(-\Omega\left(\frac{\ell}{\varepsilon^{-2}\delta^{-2}k^3}\right)\right). \tag{3.1}$$

For $\ell = \Omega\left(k^3\varepsilon^{-2}\delta^{-2}\log\delta^{-1}\right)$, we have $e^{-\eta\delta^2\ell} \leq \delta/2$. In this case, Definition 16 is also satisfied. ◀

3.2 The Disentangler from Unentanglement

We now construct our disentangler using the PAPO procedure, thereby proving Theorem 2 (restated below).

▶ **Theorem 2** (Disentangler from unentanglement). *Let $d, \ell \geq k \in \mathbb{N}^+$. There is an efficient channel $\Lambda\colon (\mathbb{C}^d)^{\otimes\ell} \otimes (\mathbb{C}^d)^{\otimes\ell} \to (\mathbb{C}^d)^{\otimes k}$ such that for any density operators $\rho_1, \rho_2 \in \mathbb{C}^{d\ell}$ there is a distribution μ on pure states $|\psi\rangle \in \mathbb{C}^d$ satisfying*

$$\left\|\Lambda(\rho_1\otimes\rho_2) - \int |\psi\rangle\langle\psi|^{\otimes k} d\mu\right\|_1 \leq \tilde{O}\left(\left(\frac{k^3}{\ell}\right)^{1/4}\right).$$

Furthermore, product states of the form $\rho_1 = \rho_2 = |\psi\rangle\langle\psi|^{\otimes\ell}$ are mapped to $|\psi\rangle\langle\psi|^{\otimes k}$.

Proof. We set $\varepsilon = \delta$, whose exact values will be determined later. Let Λ_0 be the $(d, \ell, k, \varepsilon, \delta)$-PAPO procedure guaranteed by Theorem 17. Suppose that we have an input state $\rho \in \mathrm{SEP}(d^\ell, 2)$. The channel Λ will be defined as follows. Run the PAPO procedure Λ_0 on input ρ, then

1. If $\Lambda_0(\rho)$ succeeds, Λ outputs $\Lambda_0(\rho)$.
2. Otherwise, Λ outputs a fixed product state say $|0\rangle\langle 0|^{\otimes k}$.

If $\rho = \rho_1 \otimes \rho_2$ with $\rho_1 = \rho_2 = |\psi\rangle\langle\psi|^{\otimes\ell}$, then Λ outputs $|\psi\rangle\langle\psi|^{\otimes k}$ as desired. If the Λ_0 rejects, Λ outputs a product state. Therefore by the soundness of Λ_0, firstly, with probability at least $1-\delta$, Λ outputs a state σ which is ε-close to a mixture of product states, i.e., for some distribution μ on $\mathcal{D}(\mathbb{C}^d)$,

$$\left\|\sigma - \int_{|\psi\rangle} |\psi\rangle\langle\psi|^{\otimes k} d\mu\right\|_1 \leq \varepsilon;$$

and secondly, with probability $\leq \delta$, we output a state ρ_{error}. Overall, we have

$$\Lambda(\rho) = (1-\delta')\sigma + \delta'\rho_{\text{error}}.$$

Therefore,

$$\begin{aligned}
&\left\|\Lambda(\rho) - \int |\psi\rangle\langle\psi|^{\otimes k} d\mu\right\|_1 \\
&\qquad = \left\|(1-\delta')\sigma + \delta'\rho_{\text{error}} - \int |\psi\rangle\langle\psi|^{\otimes k} d\mu\right\|_1 \\
&\qquad \leq \left\|\sigma - \int |\psi\rangle\langle\psi|^{\otimes k} d\mu\right\|_1 + \|-\delta'\sigma + \delta'\rho_{\text{error}}\|_1 \\
&\qquad \leq \varepsilon + 2\delta.
\end{aligned}$$

In view of Theorem 17, for $\varepsilon = \delta$,

$$\left\|\Lambda(\rho) - \int |\psi\rangle\langle\psi|^{\otimes k} d\mu\right\|_1 \leq \tilde{O}\left(\left(\frac{k^3}{\ell}\right)^{1/4}\right),$$

concluding the proof. ◀

4 Quantum Slicing de Finetti Theorem

In this section, we prove Theorem 22. In spirit, it is a de Finetti type theorem with the contraint that there is little entanglement across some cut. We refer to such type of theorem as the slicing de Finetti theorem.

4.1 One-versus-Many Slicing de Finetti

To start, we study the following most basic scenario that a given permutation-invariant pure quantum state from $\vee^k(\mathbb{C}^d)$ has a large top Schmidt coefficient over cut between the first and the remaining subsystems. We obtain a dimension independent quantum de Finetti theorem under slicing constraints from first principles.

▶ **Theorem 23** (One-versus-many Slicing de Finetti). *Let $|\sigma\rangle^{A_1\cdots A_k} \in \vee^k(\mathbb{C}^d)$. If the largest Schmidt coefficient across the cut $A_1 : A_2\cdots A_k$ is at least $\sqrt{1-\varepsilon}$, then*

$$\max_{|\phi\rangle\in\mathbb{C}^d} \left|\langle\sigma|^{A_1\cdots A_k}|\phi\rangle^{\otimes k}\right|^2 \geq 1 - 8k^3\cdot\varepsilon.$$

To prove this theorem, we first establish the following duplicate lemma. It says that when a symmetric state $|\sigma\rangle$ is close to some product state $|\phi\rangle|\rho\rangle$, then you can find a new state close to $|\sigma\rangle$ that with two $|\phi\rangle$ and harms the closeness only mildly.

▶ **Lemma 24** (Duplicate Lemma). *Let $|\sigma\rangle \in \vee^k(\mathbb{C}^d)$. Consider some arbitrary decomposition of $\{A_1, A_2, \ldots, A_k\} = \mathcal{A}\cup\mathcal{B}\cup\mathcal{C}$, such that $|\mathcal{A}| = |\mathcal{B}|$. Suppose $|\langle\sigma|^{\mathcal{ABC}}|\phi\rangle^{\mathcal{A}}|\rho\rangle^{\mathcal{BC}}|^2 \geq 1-\varepsilon$. Then, there is a state $|\zeta\rangle^{\mathcal{ABC}}$ such that $|\zeta\rangle = |\phi\rangle^{\mathcal{A}}|\phi\rangle^{\mathcal{B}}\,|\gamma\rangle^{C}$ for some $|\gamma\rangle^{C}$, and*

$$|\langle\sigma \mid \zeta\rangle|^2 \geq 1 - 8\varepsilon.$$

Furthermore, if $\rho_{\mathcal{C}}$ is a pure state, then $\gamma = \rho_{\mathcal{C}}$.

Proof. We assume that $\varepsilon < 1/8$, otherwise the statement is trivially true. Apply Schmidt decomposition to $|\rho\rangle^{\mathcal{BC}}$ for the $\mathcal{B}:\mathcal{C}$ cut,

$$|\rho\rangle^{\mathcal{BC}} = \sum_i \sqrt{\lambda_i}|\beta_i\rangle^{\mathcal{B}}|\gamma_i\rangle^{\mathcal{C}}.$$

Let

$$|\rho'\rangle^{\mathcal{AC}} = \sum_i \sqrt{\lambda_i}|\beta_i\rangle^{\mathcal{A}}|\gamma_i\rangle^{\mathcal{C}}.$$

Since $|\sigma\rangle \in \vee^k(\mathbb{C}^d)$, we have

$$|\langle\sigma|^{\mathcal{ABC}}|\phi\rangle^{\mathcal{A}}|\rho\rangle^{\mathcal{BC}}|^2 = |\langle\sigma|^{\mathcal{ABC}}|\phi\rangle^{\mathcal{B}}|\rho'\rangle^{\mathcal{AC}}|^2 = 1-\varepsilon.$$

By Claim 10,

$$(1-2\varepsilon)^2 \leq |\langle\phi|^{\mathcal{A}}\langle\rho|^{\mathcal{BC}}|\phi\rangle^{\mathcal{B}}|\rho'\rangle^{\mathcal{AC}}|^2 = \left(\sum_i \lambda_i|\langle\phi \mid \beta_i\rangle|^2\right)^2. \tag{4.1}$$

Abbreviate $\eta_i = |\langle\phi \mid \beta_i\rangle|^2$. Note that

$$\sum \eta_i \leq 1, \quad \sum \lambda_i = 1.$$

Therefore, immediately from (4.1),

$$\lambda_1, \max \eta_i \geq 1 - 2\varepsilon, \tag{4.2}$$

which is at least $3/4$ since $\varepsilon < 1/8$. If $\eta_1 \neq \max \eta_i$, then

$$1 - 2\varepsilon \leq \sum_i \lambda_i \eta_i \leq \lambda_1(1 - \max \eta_i) + \max \eta_i \cdot (1 - \lambda_1) \leq 4\varepsilon,$$

which is impossible as $\varepsilon < 1/8$. Therefore, $\eta_1 \geq 1 - 2\varepsilon$.

We push it further,

$$\begin{aligned} 1 - 2\varepsilon \leq \sum_i \lambda_i \eta_i &\leq \lambda_1\eta_1 + (1 - \lambda_1)(1 - \eta_1) = 2\lambda_1\eta_1 - \lambda_1 - \eta_1 + 1 \\ &\leq 2\lambda_1\eta_1 - 2\sqrt{\lambda_1\eta_1} + 1 \\ &= 2\left(\sqrt{\lambda_1\eta_1} - \frac{1}{2}\right)^2 + \frac{1}{2}, \end{aligned}$$

where the second step is due to AM-GM inequality. Since $\lambda_1, \eta_1 > 3/4$, and $\varepsilon < 1/8$,

$$\lambda_1\eta_1 \geq \left(\frac{1}{2} + \sqrt{\frac{1}{4} - \varepsilon}\right)^2 \geq 1 - 3\varepsilon,$$

where the last inequality holds for $\varepsilon \in [0, 1/8]$. Note that

$$|\langle\phi|^{\mathcal{A}}\langle\rho|^{\mathcal{BC}}|\phi\rangle^{\mathcal{A}}|\phi\rangle^{\mathcal{B}}|\gamma_1\rangle^{\mathcal{C}}|^2 \geq \lambda_1\eta_1 \geq 1 - 3\varepsilon.$$

By Claim 10 and that $1 - 3\varepsilon > 1/2$, it can be verified that

$$\begin{aligned} |\langle\sigma|^{\mathcal{ABC}}|\phi\rangle^{\mathcal{A}}|\phi\rangle^{\mathcal{B}}|\gamma_1\rangle^{\mathcal{C}}|^2 &\geq (\sqrt{(1-\varepsilon)(1-3\varepsilon)} - \sqrt{3}\varepsilon)^2 \\ &= 1 - 4\varepsilon + 6\varepsilon^2 - 2\sqrt{3}\varepsilon\sqrt{(1-\varepsilon)(1-3\varepsilon)} \\ &\geq 1 - 8\varepsilon. \end{aligned}$$

◀

Now Theorem 23 is a simple consequence of Lemma 24: Duplicate the the first subsystem taken from the top Schmidt component of $|\sigma\rangle$.

Proof of Theorem 23. Let $|\sigma_0\rangle = |\phi\rangle|\gamma\rangle$ be the top Schmidt component of $|\sigma\rangle$ for the $A_1 : A_2 \ldots A_k$ cut. By assumption of the theorem statement,

$$|\langle\sigma \mid \sigma_0\rangle|^2 \geq 1 - \varepsilon.$$

Let $m = \lfloor \log k \rfloor, m^* = \lceil \log k \rceil$. For $i = 1, 2, \ldots, m$, apply the Duplicate Lemma on $|\sigma_{i-1}\rangle$ with $\mathcal{A} = \{A_1, A_2, \ldots, A_{2^{i-1}}\}, \mathcal{B} = \{A_{2^{i-1}+1}, A_{2^{i-1}+2}, \ldots, A_{2^i}\}$. Let $|\sigma_i\rangle$ be the $|\zeta\rangle$ guaranteed by the Duplicate Lemma.

If $2^m < k$, apply the Duplicate Lemma one more time on $|\sigma_m\rangle$ with $\mathcal{A} = \{A_1, A_2, \ldots, A_{k-2^m}\}, \mathcal{B} = \{A_{2^m+1}, A_{2^m+2}, \ldots, A_k\}$, and let $|\sigma_{m^*}\rangle$ be the state guaranteed by the Duplicate Lemma. Then, a straightforward induction shows

1. $|\langle\sigma \mid \sigma_{m^*}\rangle|^2 \geq 1 - 8^{m^*}\varepsilon \geq 1 - 8k^3\varepsilon$,
2. $|\sigma_{m^*}\rangle = |\phi\rangle^{\otimes k}$.

That finishes the proof. ◀

We make a remark about Theorem 23. Note that some polynomial dependence on k is unavoidable in this analysis for our procedure. Consider the following state:

$$\frac{1}{\sqrt{k+1}}|\vec{0}\rangle + \frac{1}{\sqrt{k+1}}\sum_{i=1}^{k}|\vec{e_i}\rangle.$$

To obtain a tight version of the above theorem with linear dependency on k is an interesting problem.

4.2 Many-versus-Many Slicing de Finetti

In Theorem 23, we considered top Schmidt coefficient being large on a 1 vs $k-1$ cut for pure state. By looking at the example we mentioned in the end of the previous subsection, it is natural to think that if the top Schmidt coefficient is large among a balanced cut, then we can obtain better trace distance. That is indeed the case. In fact, that top Schmidt coefficient is large for a balanced cut always implies the top Schmidt coefficient is large for a less balanced cut for a symmetric state. In this subsection, our goal is to formalize this intuition.

▶ **Theorem 25** (Many-versus-many Slicing de Finetti). *Let $|\sigma\rangle^{A_1\dots A_k} \in \vee^k(\mathbb{C}^d)$. Suppose for some $1 \le \ell \le k/2$, the top Schmidt coefficient of $|\sigma\rangle$ over the $A_1\dots A_\ell : A_{\ell+1}\dots A_k$ cut is $\sqrt{1-\varepsilon}$. Then there is $|\phi\rangle \in \mathbb{C}_d$, such that*

$$|\langle\sigma, \phi^{\otimes k}\rangle|^2 \ge 1 - O((k/\ell)^3\varepsilon).$$

We start by collecting a couple of useful facts. The first one says that if a symmetric state from $(\mathbb{C}^d)^{\otimes k}$ is close to a product state, then it is also close to a symmetric product state, i.e., $|\phi^{\otimes k}\rangle$ for some $|\phi\rangle \in \mathbb{C}^d$.

▶ **Lemma 26.** *Given a symmetric state $|\sigma\rangle \in \vee^k(\mathbb{C}^d)$ and a k-partite product state $|\psi\rangle \in (\mathbb{C}^d)^{\otimes k}$. Suppose $|\langle\psi \mid \sigma\rangle|^2 \ge 1-\varepsilon$. Then there is $|\phi\rangle \in \mathbb{C}^d$ that satisfies*

$$|\langle\sigma \mid \phi^{\otimes k}\rangle|^2 \ge 1 - 9\varepsilon.$$

Proof. We take advantage of $|\sigma\rangle$ being symmetric in a way similar to that of Lemma 24. As $|\sigma\rangle \in \vee^k(\mathbb{C}^d)$, we have for any permutation $\pi \in \mathrm{Sym}_k$, $|\langle\sigma \mid \pi\psi\rangle|^2 \ge 1-\varepsilon$. By Claim 10,

$$|\langle\psi \mid \pi\psi\rangle|^2 \ge 1 - 4\varepsilon.$$

Say $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_k\rangle$, then,

$$\begin{aligned}(1-4\varepsilon)^{k!} \le \prod_{\pi\in\mathrm{Sym}_k} |\langle\psi \mid \pi\psi\rangle|^2 &= \left(\prod_{i\in[k]}\prod_{j\in[k]} |\langle\psi_i \mid \psi_j\rangle|^2\right)^{(k-1)!} \\ &\le \left(\mathop{\mathbb{E}}_{i\in[k]}\prod_{j\in[k]} |\langle\psi_i \mid \psi_j\rangle|^2\right)^{k!}, \end{aligned} \tag{4.3}$$

where the last step uses the AM-GM inequality. It follows from (4.3), there must exist $i \in [k]$ such that

$$1-4\varepsilon \le \prod_{j\in[k]} |\langle\psi_i \mid \psi_j\rangle|^2 \quad\Longleftrightarrow\quad 1-4\varepsilon \le |\langle\psi_i^{\otimes k} \mid \psi\rangle|^2.$$

Apply Claim 10 one more time, we obtain our lemma. ◀

The second fact due to Harrow and Montanaro [17, Appendix B Lemma 2] and Soleimanifar and Wright [31] establishes some criteria when a pure state is close to a product state.

▶ **Lemma 27.** *Given any quantum state $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_k$ for some arbitrary Hilbert space $\mathcal{H}_1, \ldots, \mathcal{H}_k$. Suppose*

$$\mathop{\mathbb{E}}_{\mathcal{S}\subseteq[k]}[\operatorname{Tr} \psi_{\mathcal{S}}^2] \geq 1 - \varepsilon.$$

Then for some product state $|\phi\rangle = |\phi_1\rangle \otimes |\phi_2\rangle \otimes \cdots \otimes |\phi_k\rangle$,

$$|\langle \psi \mid \phi \rangle|^2 \geq 1 - 3\varepsilon.$$

Combining the above two lemmas, we obtain

▶ **Corollary 28.** *Given any state $|\sigma\rangle \in \vee^k(\mathbb{C}^d)$. Suppose*

$$\mathop{\mathbb{E}}_{\mathcal{S}\subseteq[k]}[\operatorname{Tr} \sigma_{\mathcal{S}}^2] \geq 1 - \varepsilon.$$

Then for some state $|\phi\rangle \in \mathbb{C}^d$,

$$|\langle \sigma \mid \phi^{\otimes k} \rangle|^2 \geq 1 - 27\varepsilon.$$

From the above discussion, to prove Theorem 25, it suffices to bound $\operatorname{Tr} \psi_{\mathcal{S}}^2$ for any subset $\mathcal{S}$. The following "cut lemma" establishes such bounds.

▶ **Lemma 29** (Cut Lemma)**.** *Let $|\sigma\rangle \in \vee^k(\mathbb{C}^d)$. Suppose for some $1 \leq \ell \leq k/2$, the top Schmidt coefficient of $|\sigma\rangle$ over the $A_1 \ldots A_\ell : A_{\ell+1} \ldots A_k$ cut is $\sqrt{1-\varepsilon}$. Let $\mathcal{S} \subseteq [k]$ be some arbitrary subset. Then,*

$$\operatorname{Tr} \sigma_{\mathcal{S}}^2 \geq \begin{cases} 1, & |\mathcal{S}| = 0; \\ 1 - 6\varepsilon, & \min\{|\mathcal{S}|, k - |\mathcal{S}|\} \in \{1, 2, \ldots, \ell - 1\}; \\ 1 - O((|\mathcal{S}|/\ell)^3 \varepsilon), & \min\{|\mathcal{S}|, k - |\mathcal{S}|\} \in \{\ell, \ldots, k/2\}. \end{cases}$$

Proof. For $\mathcal{S} = \varnothing$, the statement is trivial as σ is pure. Since $|\sigma\rangle \in \vee^k(\mathbb{C}^d)$, without loss of generality, assume that $\mathcal{S} = \{1, 2, \ldots, m\}$ for some $1 \leq m \leq k/2$. This is because $\operatorname{Tr} \sigma_{\mathcal{S}}^2 = \operatorname{Tr} \sigma_{\overline{\mathcal{S}}}^2$ when σ is a pure state. Let $|\phi\rangle^{A_1 \ldots A_\ell} |\zeta\rangle^{A_{\ell+1} \ldots A_k}$ be the top Schmidt component associated with the coefficient $\sqrt{1-\varepsilon}$.

Case 1. $m < \ell$. Let $\mathcal{A} = \{1, 2, \ldots, m\}, \mathcal{B} = \{m+1, \ldots, \ell\}, \mathcal{C} = \{\ell+1, \ldots, k-\ell+m\}, \mathcal{D} = \{k-\ell+m+1, \ldots, k\}$. Write down the Schmidt decomposition of $|\phi\rangle$ over the $\mathcal{A}$ and $\mathcal{B}$ cut, $|\zeta\rangle$ over the $\mathcal{C}$ and $\mathcal{D}$ cut,

$$|\phi\rangle = \sum_i \sqrt{\lambda_i} |\alpha_i\rangle |\beta_i\rangle, \qquad |\zeta\rangle = \sum_i \sqrt{\eta_i} |\gamma_i\rangle |\delta_i\rangle.$$

Since $\mathcal{B}$ and $\mathcal{D}$ has the same size, and that $|\sigma\rangle \in \vee^k(\mathbb{C}^d)$, we have for the state $|\phi\rangle|\zeta\rangle$, if we switch the subsystem of $\mathcal{B}$ and $\mathcal{D}$, then the overlap with $|\sigma\rangle$ is still $1 - \varepsilon$. Therefore, by Claim 10, we have

$$
\begin{aligned}
(1-2\varepsilon)^2 &\le \left| \left\langle \sum_{i,j} \sqrt{\lambda_i \eta_j} \langle \alpha_i|^{\mathcal{A}} \langle \beta_i|^{\mathcal{B}} \langle \gamma_j|^{\mathcal{C}} \langle \delta_j|^{\mathcal{D}}, \sum_{i,j} \sqrt{\lambda_i \eta_j} |\alpha_i\rangle^{\mathcal{A}} |\beta_i\rangle^{\mathcal{D}} |\gamma_j\rangle^{\mathcal{C}} |\delta_j\rangle^{\mathcal{B}} \right\rangle \right|^2 \\
&= \left(\sum_{i,j} \lambda_i \eta_j |\langle \beta_i \mid \delta_j \rangle|^2 \right)^2 \le \lambda_1^2 \left(\sum_{i,j} \eta_j |\langle \beta_i \mid \delta_j \rangle|^2 \right)^2 \\
&\le \lambda_1^2 \left(\sum_j \eta_j \sum_i |\langle \beta_i \mid \delta_j \rangle|^2 \right)^2 \le \lambda_1^2.
\end{aligned}
$$

Immediately,

$$
\begin{aligned}
\operatorname{Tr} \sigma_{\mathcal{A}}^2 &\ge (1-\varepsilon)^2 \operatorname{Tr}[(\operatorname{Tr}_{\mathcal{BCD}}(\phi \otimes \zeta))^2] \\
&= (1-\varepsilon)^2 \operatorname{Tr} \left[\left(\operatorname{Tr}_{\mathcal{BCD}} \left(\sum_i \lambda_i \alpha_i \otimes \beta_i \otimes \zeta \right) \right)^2 \right] \\
&\ge (1-\varepsilon)^2 \lambda_1^2 \ge (1-\varepsilon)^2 (1-2\varepsilon)^2 \\
&\ge 1 - 6\varepsilon. \qquad (4.4)
\end{aligned}
$$

The first step is true because $\sigma \succeq (1-\varepsilon)\phi \otimes \zeta$, therefore $\operatorname{Tr}_{\mathcal{BCD}} \sigma \succeq (1-\varepsilon) \operatorname{Tr}_{\mathcal{BCD}}(\phi \otimes \zeta)$ as partial trace is completely positive. It then follows that $\operatorname{Tr} \sigma_{\mathcal{A}}^2 \succeq (1-\varepsilon)^2 \operatorname{Tr}(\operatorname{Tr}_{\mathcal{BCD}}(\phi \otimes \zeta))^2$.

Case 2. $\ell < m \le k/2$. We are much like the situation of Theorem 23. Let $t = \lceil \log(m/\ell) \rceil$. For $i = 1$ to t, we apply the Duplicate Lemma and obtain a state $|\sigma_i\rangle$, such that for $i = 1, 2, \ldots, t$

$$
\begin{aligned}
&\operatorname{Tr}_{\{\ell \cdot 2^i + 1, \ldots, k\}} \sigma_i = \phi^{\otimes 2^i}, \\
&\operatorname{Tr}_{\{\ell \cdot 2^i + 1, \ldots, k\}} \sigma_i^2 = 1, \qquad (4.5) \\
&|\langle \sigma \mid \sigma_i \rangle|^2 \ge 1 - 8^i \varepsilon. \qquad (4.6)
\end{aligned}
$$

By our choice of parameter, $2^{t-1}\ell < m \le 2^t \ell$. If $m = 2^t \ell$, then $(\sigma_t)_{\mathcal{S}} = \operatorname{Tr}_{\ell \cdot 2^t + 1, \ldots, k} \sigma_t$ is pure by (4.5). Then

$$
\begin{aligned}
\sqrt{\operatorname{Tr} \sigma_{\mathcal{S}}^2} = \sqrt{\operatorname{Tr} \sigma_{\mathcal{S}}^2 \cdot \operatorname{Tr}(\sigma_t)_{\mathcal{S}}^2} &\ge \operatorname{Tr}(\sigma_{\mathcal{S}} \cdot (\sigma_t)_{\mathcal{S}}) = F(\sigma_{\mathcal{S}}, (\sigma_t)_{\mathcal{S}}) \\
&\ge F(\sigma, \sigma_t) = |\langle \sigma \mid \sigma_t \rangle|^2 \ge (1 - 8^{\log(m/\ell)})\varepsilon,
\end{aligned}
$$

where the first step and third step are true because $(\sigma_t)_{\mathcal{S}}$ is pure; the second step uses (2.1); the fourth step is by Fact 11, the data processing inequality for fidelity; then fifth step is again by purity of the states; and the final step uses (4.6). It follows that

$$\operatorname{Tr} \sigma_{\mathcal{S}}^2 \ge 1 - O((m/\ell)^3 \varepsilon).$$

If $m < 2^t \ell$, then we can apply Case 1. Let $\mathcal{A} = \{1, 2, \ldots, 2^t \ell\}, \mathcal{B} = \{2^t \ell + 1, \ldots, k\}$. Then in view of (4.6), the top Schmidt coefficient of $|\sigma\rangle$ among the $\mathcal{A} : \mathcal{B}$ cut is at least $\sqrt{1 - 8^t \varepsilon}$ by Claim 12. Thus by (4.4),

$$\operatorname{Tr} \sigma_{\mathcal{S}}^2 \ge 1 - 6 \cdot 8^t \varepsilon \ge 1 - O((m/\ell)^3 \varepsilon).$$ ◀

Now Theorem 25 follows from Corollary 28 and Lemma 29.

4.3 Proof of Theorem 22

Now we record a version of the slicing de Finetti theorem for the mixture of symmetric states. A natural generalization of the top Schmidt coefficient among some $A : B$ cut for a state σ being large is that $\operatorname{Tr} \sigma_A^2$ being large. In particular,

▶ **Lemma 30.** *Let $\sigma \in \mathbb{C}^n \otimes \mathbb{C}^m$ be some density operator, and A, B are the systems with respect to the space $\mathbb{C}^n$ and $\mathbb{C}^m$, respectively. Suppose*

$$\operatorname{Tr} \sigma_A^2 \geq 1 - \varepsilon.$$

Let μ be some distribution on pure states induced by σ, then

$$\mathop{\mathbb{E}}_{\rho \sim \mu} \lambda_1(\rho) \geq 1 - \varepsilon.$$

Proof. Let $m = |\operatorname{supp} \mu|$ be a finite number, this is without loss of generality. Let $\rho_1, \rho_2, \ldots, \rho_m$ be the pure states in $\operatorname{supp} \mu$. Further, write the Schmidt decomposition for each ρ_i

$$|\rho_i\rangle = \sum_j \sqrt{\lambda_{ij}} |\phi_{ij}\rangle^A |\sigma_{ij}\rangle^B, \qquad \lambda_{i1} \geq \lambda_{i2} \geq \cdots .$$

Then

$$\sigma_A = \sum_i \mu(\rho_i) \sum_j \lambda_{ij} |\phi_{ij}\rangle\langle\phi_{ij}|.$$

Thus,

$$\begin{aligned}
\operatorname{Tr} \sigma_A^2 &= \sum_i \mu(\rho_i)^2 \sum_j \lambda_{ij}^2 + \sum_{i \neq i'} \mu(\rho_i)\mu(\rho_{i'}) \sum_{j,j'} \lambda_{ij}\lambda_{i'j'} |\langle \phi_{ij} \mid \phi_{i'j'} \rangle|^2 \\
&\leq \sum_i \mu(\rho_i)^2 \sum_j \lambda_{ij}^2 + \sum_{i \neq i'} \mu(\rho_i)\mu(\rho_{i'})\lambda_{i1} \sum_{j,j'} \lambda_{i'j'} |\langle \phi_{ij} \mid \phi_{i'j'} \rangle|^2 \\
&\leq \sum_i \mu(\rho_i)^2 \sum_j \lambda_{ij}^2 + \sum_{i \neq i'} \mu(\rho_i)\mu(\rho_{i'})\lambda_{i1} \sum_{j'} \lambda_{i'j'} \\
&\leq \sum_i \mu(\rho_i)^2 \sum_j \lambda_{ij}^2 + \sum_{i \neq i'} \mu(\rho_i)\mu(\rho_{i'})\lambda_{i1} \\
&= \sum_i \mu(\rho_i)^2 \sum_j \lambda_{ij}^2 + \sum_i \mu(\rho_i)(1 - \mu(\rho_i))\lambda_{i1} \\
&\leq \sum_i \mu(\rho_i)^2 \lambda_{i1} + \sum_i \mu(\rho_i)(1 - \mu(\rho_i))\lambda_{i1} \\
&= \sum_i \mu(\rho_i)\lambda_{i1},
\end{aligned}$$

where the third step holds because for fixed i, i', j', $\sum_j |\langle \phi_{ij} \mid \phi_{i'j'} \rangle|^2 \leq 1$. ◀

▶ **Theorem 31.** *Given density operator $\sigma^{A_1 \cdots A_k}$ that describes states from $\operatorname{conv}\left(\vee^k(\mathbb{C}^d)\right)$. For any $1 \leq \ell \leq k/2$ and $\mathcal{A} = \{A_1, A_2, \cdots, A_\ell\}$, there is some distribution μ on $|\phi\rangle \in \mathbb{C}^d$,*

$$\left\| \sigma - \int |\phi\rangle\langle\phi|^{\otimes k} d\mu \right\|_1 \leq O\left(\sqrt{(k/\ell)^3 (1 - \operatorname{Tr} \sigma_{\mathcal{A}}^2)} \right). \tag{4.7}$$

Proof. Let μ be the distribution on pure symmetric states induced by σ. Let $\operatorname{Tr} \sigma_{\mathcal{A}}^2 = 1 - \varepsilon$. The theorem follows immediately by combining Fact 8, Lemma 30, Theorem 25, and triangle inequality. ◀

5 A Framework: Multiplexing Unentangled States for Property Testing

In this section, we present a general template illustrating the utility of our disentangler Theorem 2. We will then use this template multiple of times. Initially, we provide two examples as warm-ups for what is to come. Subsequently, in later sections, we apply this template with carefully designed testers to obtain new complexity results.

Our disentangler leverages a bipartite unentanglement assumption between two states of the form $\rho_1 \otimes \rho_2$ into an (approximate) multipartite unentanglement assumption of the form $\int |\psi\rangle\langle\psi|^{\otimes k} d\mu$. Having sufficiently many unentangled copies of a state ψ is particularly important in the context of quantum property testing as some properties require this assumption for testability. Indeed, many of other information processing tasks like quantum state tomography often assumes the input is of this form $|\psi\rangle\langle\psi|^{\otimes k}$. Moreover, multiple copies allow the tester to be executed multiple times amplifying its probability of distinguishing the closeness to the desired property. Finally, a property tester may end up destroying the copies $\psi^{\otimes k}$ when it measures this state, so it is desirable to have additional copies that can be used in further information processing tasks once the closeness to the desired property is certified. In Figure 2, we provide an illustration of a property tester being used in conjunction with our disentangler in order to obtain the aforementioned benefits.

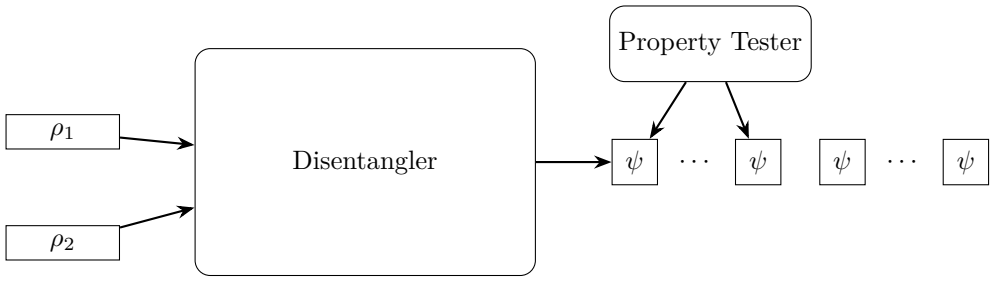

Figure 2 Schematic picture of our disentangler being used to (approximately) ensure multiple unentangled copies of a state as output. Part of these copies are used to test a given desired property. If the test passes, the remaining "certified" copies can be used in further information processing tasks.

Product Tester and Preparing Multipartite Separable States

To make this illustration more concrete, first we consider a scenario where the tester is the product test [17]. More precisely, the product test requires two unentangled copies of $|\psi\rangle \in \mathbb{C}^d$ and checks whether $|\psi\rangle$ is close to a product state of the form $|\phi_1\rangle \otimes \cdots \otimes |\phi_s\rangle \in \mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_s}$, where $d = d_1 \cdots d_s$. For context, recall that (an abridged version of) their main result provides the following guarantees for this tester.

▶ **Theorem 32** (Product Test [17]). *Given* $|\psi\rangle \in \mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_s}$, *let*

$$1 - \varepsilon = \max \left\{ |\langle \psi \,|\, \phi_1, \ldots, \phi_s \rangle|^2 : |\phi_i\rangle \in \mathbb{C}^{d_i}, 1 \leq i \leq s \right\}.$$

Let $P_{test}(|\psi\rangle\langle\psi|)$ *be the probability that the product test passes when applied to* $|\psi\rangle$. *Then, we have* $P_{test}(|\psi\rangle\langle\psi|) = 1 - \Theta(\varepsilon)$.

Combining our disentangler from Theorem 2 and the product test from Theorem 32, we obtain the following corollary giving all the desired qualities alluded above in a more quantitative way.

▶ **Corollary 33.** *Let $\mathcal{H} = \mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_s}$. For every $k, k', \ell \in \mathbb{N}$ such that $\ell \geq k + 2k'$, there is a channel $\Gamma\colon \mathcal{D}(\mathcal{H}^{\otimes \ell} \otimes \mathcal{H}^{\otimes \ell}) \to \mathcal{D}(\mathcal{H}^{\otimes k} \otimes \mathbb{C}^2)$ such that for every $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H}^{\otimes \ell})$, there exists $\sigma \in \mathcal{D}(\mathcal{H}^{\otimes k} \otimes \mathbb{C}^2)$ defined as*

$$\sigma = \int |\psi\rangle\langle\psi|^{\otimes k} \otimes \left(P_{test}(|\psi\rangle\langle\psi|)^{k'} |1\rangle\langle 1| + (1 - P_{test}(|\psi\rangle\langle\psi|)^{k'}) |0\rangle\langle 0| \right) d\mu \,,$$

such that

$$\|\Gamma(\rho_1 \otimes \rho_2) - \sigma\|_1 \leq \tilde{O}\left(\left(\frac{(k+2k')^3}{\ell} \right)^{1/4} \right).$$

Furthermore, $\Gamma(\rho_1 \otimes \rho_2) = (|\psi\rangle\langle\psi|)^{\otimes k} \otimes |1\rangle\langle 1|$ provided $\rho_1 = \rho_2 = (|\psi\rangle\langle\psi|)^{\otimes \ell}$, where $|\psi\rangle = |\phi_1\rangle \otimes \ldots \otimes |\phi_s\rangle$ for some $|\phi_i\rangle \in \mathbb{C}^{d_i}$ for $1 \leq i \leq s$.

Proof. Define another channel $\Gamma'\colon \mathcal{D}(\mathcal{H}^{\otimes (k+2k')}) \to \mathcal{D}(\mathcal{H}^{\otimes k} \otimes \mathbb{C}^2)$ that takes as input the output of the disentangler Λ which is comprised of $k + 2k'$ registers of the space $\mathcal{H}$. We define the channel Γ' to act as identity on the first k registers. On the last $2k'$ registers it performs the product test on each pair of registers, outputting a single qubit $|1\rangle\langle 1|$ if all tests pass, otherwise outputting $|0\rangle\langle 0|$. Next we show $\Gamma = \Gamma' \circ \Lambda$, the composed channel, satisfies the statement.

Given general input $\rho_1 \otimes \rho_2$, by the guarantee of our disentangler, $\Lambda(\rho_1 \otimes \rho_2)$ satisfies

$$\left\| \Lambda(\rho_1 \otimes \rho_2) - \int |\psi\rangle\langle\psi|^{\otimes k+2k'} d\mu \right\|_1 \leq \tilde{O}\left(\left(\frac{(k+2k')^3}{\ell} \right)^{1/4} \right).$$

Note that Γ' applied to $\int |\psi\rangle\langle\psi|^{\otimes k+2k'} d\mu$ results in

$$\int |\psi\rangle\langle\psi|^{\otimes k} \otimes \left(P_{\text{test}}(|\psi\rangle\langle\psi|)^{k'} |1\rangle\langle 1| + (1 - P_{\text{test}}(|\psi\rangle\langle\psi|)^{k'}) |0\rangle\langle 0| \right) d\mu \,. \tag{5.1}$$

Thus, the composed channel output $\Gamma(\rho_1 \otimes \rho_2))$ is $\tilde{O}(((k+2k')^3/\ell)^{1/4})$ close, in trace distance, to the state of (5.1).

The furthermore part is straightforward. Suppose that $|\psi\rangle = |\phi_1\rangle \otimes \ldots \otimes |\phi_s\rangle$, where $|\phi_i\rangle \in \mathbb{C}^{d_i}$ for $1 \leq i \leq s$, and $\rho_1 = \rho_2 = (|\psi\rangle\langle\psi|)^{\otimes \ell}$. In this case, $\Lambda(\rho_1 \otimes \rho_2) = (|\psi\rangle\langle\psi|)^{\otimes k+2k'}$ and $\Gamma'(\Lambda(\rho_1 \otimes \rho_2)) = (|\psi\rangle\langle\psi|)^{\otimes k} \otimes |1\rangle\langle 1|$ since $|\psi\rangle$ is a product state and product test accepts with probability 1. ◀

QMA(2) Tester – Gap Amplification for QMA(2)

The gap amplification of QMA(2) was first proved in the seminar work of Harrow and Montanaro [17]. Using our template, we provide a conceptually more straightforward proof: Take the old QMA(2) protocol as the property tester in Figure 2.

▶ **Theorem 34.** *Given a language $\mathcal{L} = (\mathcal{L}_{\text{yes}}, \mathcal{L}_{\text{no}})$. Suppose that $\mathcal{L} \in \text{QMA}(2)$ with completeness c and soundness s, where $c - s > 1/\text{poly}(n)$. Then, $\mathcal{L} \in \text{QMA}(2)$ with completeness $c' = 1 - \exp(-\text{poly}(n))$ and soundness $s' = 1/\text{poly}(n)$.*

Proof. Let $\mathcal{P}$ be the protocol for $\mathcal{L}$ with the promised completeness c and soundness s. Therefore, for any fixed input x there is a measurement M acting on a space $\mathcal{H}^{\otimes 2}$ where $\mathcal{H} = \mathbb{C}^d$, such that,

$$\begin{aligned} &\exists \sigma \otimes \rho \in \mathcal{D}(\mathcal{H}^{\otimes 2}),\ \text{Tr}(M(\sigma \otimes \rho)) \geq c, && \text{if } x \in \mathcal{L}_{\text{yes}} \\ &\forall \sigma \otimes \rho \in \mathcal{D}(\mathcal{H}^{\otimes 2}),\ \text{Tr}(M(\sigma \otimes \rho)) \leq s, && \text{if } x \in \mathcal{L}_{\text{no}}. \end{aligned}$$

In the new protocol, choose $k = \text{poly}(n)/(c-s)^2$ and $\ell = \text{poly}(k)$ for some large enough polynomial. We ask for two proofs $|\rho_1\rangle, |\rho_2\rangle \in \mathcal{D}(\mathcal{H}'^{\otimes \ell})$, where $\mathcal{H}' = \mathbb{C}^2 \otimes \mathcal{H}$. In words, $\mathcal{H}'$ is $\mathcal{H}$ with one extra qubit. Apply the disentangler Λ from Theorem 2 on $\rho_1 \otimes \rho_2$, obtaining a separable state $\phi = \int d\mu |\psi\rangle\langle\psi|^{\otimes k}$, such that

$$\left\| \Lambda(\rho_1 \otimes \rho_2) - \int d\mu |\psi\rangle\langle\psi|^{\otimes k} \right\|_1 = \frac{1}{\text{poly}(n)}. \tag{5.2}$$

Consider the new measurement $M' = |01\rangle\langle 01| \otimes M$. We apply $M'^{\otimes (k/2)}$ to $\Lambda(\rho_1 \otimes \rho_2)$. Accept if more than $(c+s)/2$ fraction of the applications of M' accepts; reject otherwise. Next, we calculate the completeness and soundness of the new protocol.

Completeness. Suppose that $x \in \mathcal{L}_{\text{yes}}$, then the faithful prover will provide

$$|\rho_1\rangle = |\rho_2\rangle = \left(\frac{|0,\sigma\rangle + |1,\rho\rangle}{\sqrt{2}} \right)^{\otimes \ell}, \text{ and } \Lambda(\rho_1 \otimes \rho_2) = \left(\frac{|0,\sigma\rangle + |1,\rho\rangle}{\sqrt{2}} \right)^{\otimes k}.$$

Calculating the probability that M' accepts $(|0,\sigma\rangle + |1,\rho\rangle)^{\otimes 2}/2$,

$$\text{Tr}\left(M' \left(\frac{|0,\sigma\rangle + |1,\rho\rangle}{\sqrt{2}} \right)^{\otimes 2} \right) = \frac{1}{4} \text{Tr}(M(\sigma \otimes \rho)) \geq c/4.$$

By Chernoff bound, with probability at least $1 - \exp(-\Omega((c-s)^2 k)) = 1 - \exp(-\text{poly}(n))$, the new protocol accepts.

Soundness. Suppose that $x \in \mathcal{L}_{\text{no}}$. Calculating the probability that M' accepts $(\alpha|0,\sigma\rangle + \beta|1,\rho\rangle)^{\otimes 2}$ for arbitrary $\alpha, \beta \in \mathbb{C}$ and arbitrary $\sigma, \rho \in \mathcal{H}$ such that $|\alpha|^2 + |\beta|^2 = 1$,

$$\text{Tr}(M'(\alpha|0,\sigma\rangle + \beta|1,\rho\rangle)^{\otimes 2} = |\alpha\beta|^2 \text{Tr}(M(\sigma \otimes \rho)) \leq s/4.$$

Therefore the probability to accept ϕ, an arbitrary convex combination of $|\psi\rangle^{\otimes k}$ is at most $\exp(-\Omega((c-s)^2 k))$ by Chernoff bound. Finally, by (5.2), the probability of accepting $\Lambda(\rho_1 \otimes \rho_2)$ is at most $1/\text{poly}(n)$. ◀

6 The Super Swap and Super Product Tests

In this section, we take another look at the product test as well as the swap test, considering one of the strongest possible generalization of the two.

We start with the more elementary swap test, which is a widely used to test if two quantum states, say $|\psi\rangle$ and $|\phi\rangle$, are equal. One fundamental limitation of the swap test is that it always accepts with probability at least $1/2$ even if the states are orthogonal. More precisely, its acceptance probability is $(1 + |\langle\psi \,|\, \phi\rangle|^2)/2$. Ideally, it would be much more useful to have a test with acceptance probability of $|\langle\psi \,|\, \phi\rangle|^2$, which is impossible with only one copy for each state. In the presence of many unentangled copies of $|\phi\rangle$ but just a single copy of $|\psi\rangle$, we show that it is possible to approach this goal with an arbitrarily small error overcoming the inherent limitation of the swap test. Therefore, we call this test the *super swap* test and we provide a description of it in Algorithm 2. In particular, this super swap test can be useful when it is difficult to produce a state $|\psi\rangle$, but much easier to produce copies of $|\phi\rangle$ and we want the tester's acceptance probability to more accurately capture how close $|\psi\rangle$ is to $|\phi\rangle$. In Section 7, the special state $|\psi\rangle$ will be a nonnegative amplitudes state which has a greater cost in the context of complexity protocols there, whereas $|\phi\rangle$ will have general amplitudes being a cheaper resource in that context.

The acceptance probability of the super swap test is established next.

Algorithm 2 SuperSwap($|\psi\rangle, |\phi\rangle^{\otimes \ell}$).

Input: $|\psi\rangle, |\phi\rangle^{\otimes \ell}$.
1. Project $|\psi\rangle|\phi\rangle^{\otimes \ell}$ onto the symmetric space $\vee^{\ell+1}(\mathbb{C}^d)$.
2. If the projection succeeds *accept*; else *reject*.

▶ **Lemma 35.** *The super swap test accepts with probability*

$$\frac{\ell \cdot |\langle \psi \,|\, \phi \rangle|^2}{\ell + 1} + \frac{1}{\ell + 1}.$$

Proof. Let $\Pi = (1/(\ell+1)!) \sum_{\pi \in \mathrm{Sym}_{\ell+1}} \pi$ be the projector onto $\vee^{\ell+1}(\mathbb{C}^d)$. Indeed, we have

$$\langle \psi | \langle \phi |^{\otimes \ell} \Pi |\psi\rangle |\phi\rangle^{\otimes \ell} = \frac{1}{\ell+1} \langle \psi \,|\, \psi \rangle \langle \phi \,|\, \phi \rangle^{\ell} + \frac{\ell}{\ell+1} |\langle \psi \,|\, \phi \rangle|^2 \langle \phi \,|\, \phi \rangle^{\ell-2} ,$$

concluding the proof. ◀

At first glance, it may seem inconvenient to assume multiple (ℓ-many) unentangled copies of $|\phi\rangle$. However, due to our disentangler channel, we can enforce a distribution over product states $|\phi\rangle^{\otimes \ell}$ by assuming only bipartite unentanglement.

Next we turn to the product test which checks whether a state is close to a k-partite product state [17]. It has a similar drawback to the usual swap test, namely, it always accepts with probability at least $1/2$ even if the state $|\psi\rangle$ is very far from product. As before, we will arbitrarily improve the soundness of the product test by having multiple unentangled copies. We call this new test the *super product* test and we describe it in Algorithm 3.

Algorithm 3 SuperProduct($|\psi\rangle, (|\phi_1\rangle \dots |\phi_k\rangle)^{\otimes \ell}$).

Input: $|\psi\rangle, (|\phi_1\rangle \dots |\phi_k\rangle)^{\otimes \ell}$
1. Project $|\psi\rangle(|\phi_1\rangle \dots |\phi_k\rangle)^{\otimes \ell}$ onto the symmetric space $\vee^{\ell+1}((\mathbb{C}^d)^{\otimes k})$.
2. If the projection succeeds *accept*; else *reject*.

▶ **Lemma 36.** *The super product test accepts with probability*

$$\frac{\ell}{(\ell+1)} \cdot |\langle \psi \,|\, \phi_1 \rangle \dots |\phi_k\rangle|^2 + \frac{1}{(\ell+1)} .$$

Proof. We view each copy of the state $|\phi_1\rangle \dots |\phi_k\rangle$ as a single state $|\phi\rangle$ and apply the super swap test to $|\psi\rangle$ and $|\phi\rangle^{\otimes \ell}$. The acceptance probability of the super product test now follows from Lemma 35. ◀

Analogously, it may seem inconvenient to assume multiple (ℓ-many) unentangled copies of $|\phi_1\rangle \dots |\phi_k\rangle$. However, that is not an issue by Corollary 33: We can enforce a distribution over product states $(|\phi_1\rangle \dots |\phi_k\rangle)^{\otimes \ell}$ by assuming only 2 unentangled states.

7 Gap Amplification for QMA$^+(k)$ up to Criticality and Almost-QMA(k) = NEXP

In the previous section, we described a very strong version of swap test and product test, noting that our disentangler channel has a good synergy with the new tests to overcome the drawbacks in their original versions. In this section, we put the tools in the context of quantum Merlin-Arthur games with unentangled provers, establishing our main complexity results Theorems 4 and 5.

7.1 Gap Amplification for QMA$^+(k)$ up to Criticality

The gap amplification for QMA$^+(k)$ is much less straightforward than QMA(2). Indeed, a full gap amplification would imply QMA(2) = NEXP. To give our half gap amplification promised in Theorem 5, we start by showing how to simulate a QMA$^+(k)$ protocol $\mathcal{P}$ given the following kinds of proofs:

1. one nonnegative-amplitudes proof $|\psi\rangle$;
2. abundant equal copies of an arbitrary proofs over reals $|\phi\rangle$.

Note we are relaxing k nonnegative-amplitudes proofs in a QMA$^+(k)$ protocol with only one nonnegative-amplitudes proof and general-amplitudes states. The motivation is, roughly, to remove as many nonnegative-amplitudes proofs in a QMA$^+(k)$ protocol as possible, so we get closer to a general QMA(k) protocol.

We will check whether $|\phi\rangle^{\otimes k}$ is close to $|\psi\rangle$. Either they are close and then we can use the many copies of $|\phi\rangle^{\otimes k}$ to simulate $\mathcal{P}$, or else they are far apart and an application of the super product test can detect this condition. A description of this simulation procedure is given in Algorithm 4, which we denote as the symmetric simulator (since it assumes many equal copies of $|\phi\rangle$).

Algorithm 4 SymSimulator.

Input: QMA$^+(k)$ protocol $\mathcal{P}$, $|\psi\rangle = \sum_i \beta_i |i\rangle : \beta_i \geq 0, |\phi\rangle^{\otimes 2k\ell}$.

- If SuperProduct$(|\psi\rangle, (|\phi\rangle^{\otimes k})^{\otimes \ell})$ fails, then *reject*.
- For $i = 1, \ldots, \ell$
  - Run the QMA$^+(k)$ protocol $\mathcal{P}$ on a new copy of $|\phi\rangle^{\otimes k}$.
  - If protocol rejects, then *reject*.
- *Accept*.

We now analyze the completeness and soundness of this simulation.

▶ **Lemma 37.** *Suppose $\mathcal{P}$ is a* QMA$^+(k)$ *protocol with completeness c and soundness s. Let $p(n)$ be a non-decreasing function such that $p(n) \geq C_0$ for a sufficiently large constant $C_0 > 0$. If $\ell \geq 8p(n)^2 \ln(2)$ and $s \leq 1/8p(n)^2$, then SymSimulator has completeness c^ℓ and soundness at most $1/2 + 1/p(n)$.*

Proof. In the completeness case, we can assume that the proofs $|\phi\rangle$ have nonnegative amplitudes and $|\psi\rangle = |\phi\rangle^{\otimes k}$. Thus, SymSimulator accepts with probability at least c^ℓ.

Now, suppose that we are in the soundness case. Set $\varepsilon = \left|\left\langle \psi \,\middle|\, \phi^{\otimes k} \right\rangle\right|^2$. By Lemma 3, the super product test accepts with probability

$$\left(\frac{\varepsilon\ell}{\ell+1} + \frac{1}{\ell+1}\right).$$

Since $\ell \geq 2p(n)$, if $\varepsilon < 1/2 + 1/2p(n)$, then the acceptance probability due to the super product test alone is at most $1/2 + 1/p(n)$ and we are done. Therefore, from now on, we assume that $\varepsilon \geq 1/2 + 1/2p(n)$.

Suppose $|\phi\rangle = \sum_i \alpha_i |i\rangle$, and let $|\phi_+\rangle = \sum_i |\alpha_i| \, |i\rangle$. Thus, $|\phi_+\rangle$ is a valid nonnegative-amplitudes state. Since $|\psi\rangle$ has nonnegative amplitudes by assumption, we should have

$$|\langle \psi \mid \phi_+^{\otimes k} \rangle|^2 \geq |\langle \psi \mid \phi^{\otimes k} \rangle|^2 = \varepsilon. \tag{7.1}$$

This is because the latter inner product incurs some cancellations due to negative values, which are avoided in the former inner product. (7.1) together with Claim 10 implies that

$$|\langle \phi^{\otimes k}, \phi_+^{\otimes k} \rangle|^2 \geq 2\varepsilon - 1.$$

Since we are assuming $\varepsilon > 1/2$, the trace distance between $|\phi\rangle^{\otimes k}$ and $|\phi_+\rangle^{\otimes k}$ can be bounded as below

$$D(\phi^{\otimes k}, \phi_+^{\otimes k}) \leq 2\sqrt{\varepsilon(1-\varepsilon)} \tag{7.2}$$

Note that $\mathcal{P}$ accepts $|\phi_+^{\otimes k}\rangle$ with probability at most s by the soundness of $\mathcal{P}$. Therefore, each execution of the protocol $\mathcal{P}$ on $|\phi\rangle^{\otimes k}$ accepts with probability, by Fact 9, at most

$$\min\{1, 2\sqrt{\varepsilon(1-\varepsilon)} + s\}.$$

The overall soundness of SymSimulator becomes

$$\left(\varepsilon \frac{\ell}{\ell+1} + \frac{1}{\ell+1}\right) \left(\min\{1, 2\sqrt{\varepsilon(1-\varepsilon)} + s\}\right)^{\ell}.$$

Now take $\varepsilon \geq 1/2 + 1/2p(n)$, and compute, we have

$$2\sqrt{\varepsilon(1-\varepsilon)} \leq 2\sqrt{\frac{1}{4} - \frac{1}{4p(n)^2}} \leq 1 - \frac{1}{2p(n)^2} + O\left(\frac{1}{p(n)^4}\right) \leq 1 - \frac{1}{4p(n)^2},$$

where the last inequality relies on $p(n) \geq C_0$ for a large enough constant $C_0 > 0$. Using that $s \leq 1/8p(n)^2$ and $\ell \geq 8p(n)^2 \ln(2)$, the final acceptance probability is

$$\left(2\sqrt{\varepsilon(1-\varepsilon)} + s\right)^{\ell} \leq \left(1 - \frac{1}{8p(n)^2}\right)^{\ell} \leq \frac{1}{2},$$

concluding the proof. ◀

To remove 2 the symmetric assumption of having multiple identical copies of $|\phi\rangle$ in SymSimulator, we use the PAPO channel Λ and the PAPO channel takes just two unentangled proofs $|\phi'\rangle$ and $|\phi''\rangle$ (of arbitrary amplitudes) as its input. In other words, we now simulate a $\mathrm{QMA}^+(k)$ protocol $\mathcal{P}$ with:

(i) one nonnegative-amplitudes proof $|\psi\rangle$;
(ii') two general states $|\phi'\rangle, |\phi''\rangle$.

A formal description of the new simulation is given in Algorithm 5.

The analysis of Algorithm 5 is similar to that of Lemma 37. Therefore, instead of presenting an analysis of Algorithm 5 in isolation, we now apply this simulation for a $\mathrm{QMA}^+(k)$ protocol $\mathcal{P}$ that solves a NEXP-complete problem. In particular, we will need the following characterization of $\mathrm{QMA}^+(2)$ from [20] as shown in the following theorem.

Algorithm 5 Simulator.

Input: QMA$^+(k)$ protocol $\mathcal{P}$, $|\psi\rangle = \sum_i \beta_i |i\rangle : \beta_i \geq 0, |\phi'\rangle, |\phi''\rangle$
- Let ρ be the output of our disentangler $\Lambda(\phi' \otimes \phi'')$ (i.e. Theorem 2).
- If SymSimulator$(\mathcal{P}, |\psi\rangle, \rho)$ accepts, then *accept*; else *reject*.

▶ **Theorem 38** ([20]). QMA$^+(2)$ = NEXP.

Algorithm 5 gives rise to a protocol for NEXP that improves the above theorem in two aspects. First, the new protocol uses three unentangled proofs among which only one is required to have nonnegative amplitudes. Second, the completeness and soundness gap of this protocol is about 1/2. This seemingly mediocre gap is in fact a critical point, which we discuss in the next section.

▶ **Theorem 5.** NEXP = QMA$^+(3)$ *with completeness* $c = 1 - \exp(-\text{poly}(n))$ *and soundness* $s = 1/2 + 1/\text{poly}(n)$. *Furthermore, we can assume a particular case of* QMA$^+(3)$ *in which two unentangled proofs have arbitrary amplitudes whereas only one unentangled proof has nonnegative amplitudes.*

Proof. From Theorem 38, we apply the standard gap amplification by asking for more unentangled proofs to obtain a QMA$^+(k)$ protocol $\mathcal{P}$ with completeness $c = 1-\exp(-\text{poly}(n))$ and soundness $s = \exp(-\text{poly}(n))$, where $k = \text{poly}(n)$. Simulate $\mathcal{P}$ using Algorithm 5. By Theorem 2, $\rho = \Lambda(\phi' \otimes \phi'')$ is $1/\text{poly}(n)$-close to a convex combination of product states $\int |\phi\rangle\langle\phi|^{\otimes 2k\ell} d\mu$ with $\ell = \text{poly}(n)$. Invoking the symmetric simulator, by Lemma 37, the completeness becomes $c^\ell \geq 1 - \exp(-\text{poly}(n))$ and the soundness $1/2 + 1/\text{poly}(n)$ for a suitable choice of polynomial $\ell = \text{poly}(n)$. ◀

7.2 Almost-QMA$^{\mathbb{R}}(k)$ = NEXP

Next, we show how to go from the nonnegative amplitudes assumptions to almost general amplitudes. Recall that the ℓ_2-sign bias of a state $|\psi\rangle = \sqrt{a}|\psi_+\rangle + \sqrt{1-a}|\psi_-\rangle$, where $|\psi_+\rangle$ and $|\psi_-\rangle$ are the normalized nonnegative and negative amplitudes parts of $|\psi\rangle$, is defined as $|a - (1-a)|$ (see Definition 14).

▶ **Theorem 4.** NEXP = almost-QMA$^{\mathbb{R}}(k)$ *with unentangled proofs of* ℓ_2*-sign bias of*[8] $b(n) \geq \text{poly}(1/n)$ *and* $k = \text{poly}(1/b(n))$.

Proof. We start with the QMA$^+(3)$ protocol from Theorem 5 with two general proofs $|\phi'\rangle, |\phi''\rangle$ and only one nonnegative proof $|\psi\rangle$. Let M be the verifier measurement. In the completeness case, we can assume that $|\psi\rangle$ has nonnegative amplitudes so we proceed to analyze the soundness case.

In the almost-QMA$^{\mathbb{R}}(3)$ protocol, $|\psi\rangle$ will no-longer be assumed to have nonnegative amplitudes. Instead, we write $|\psi\rangle = \sqrt{a}|\psi_+\rangle + \sqrt{1-a}|\psi_-\rangle$, where $|\psi_+\rangle$ and $|\psi_-\rangle$ are its nonnegative- and negative-amplitudes normalized states. Without loss of generality, suppose that $a \geq 1/2$. Furthermore, under the ℓ_2-sign bias assumption, we may assume that

$$a \geq 1/2 + \sqrt{100/p(n)}. \tag{7.3}$$

[8] The letter n represents the input size and $b(n)$ is any polynomial time computable function bounded from below by a polynomial, i.e., by $1/n^c$ for some constant $c > 0$.

Let $|\phi'\rangle$ and $|\phi''\rangle$ be some quantum states (ignoring the ℓ_2-bias requirement) as to be used in the simulation Algorithm 5. The combined proofs of the almost-QMA$^{\mathbb{R}}$(3) protocol can be expressed as $|\xi\rangle = \sqrt{a}|\xi_0\rangle + \sqrt{1-a}|\xi_1\rangle$, where $|\xi_0\rangle = |\phi'\rangle \otimes |\phi''\rangle \otimes |\psi_+\rangle$ and $|\xi_1\rangle = |\phi'\rangle \otimes |\phi''\rangle \otimes |\psi_-\rangle$. Denote s the soundness of QMA$^+$(3) protocol from Theorem 5. Then we can assume

$$s \leq 1/2 + 6/p(n). \tag{7.4}$$

Calculating the accepting probability of M on ξ,

$$\begin{aligned}\langle\xi|M|\xi\rangle &= a\langle\xi_0|M|\xi_0\rangle + (1-a)\langle\xi_1|M|\xi_1\rangle \\ &\qquad + \sqrt{a(1-a)}\langle\xi_0|M|\xi_1\rangle + \sqrt{a(1-a)}\langle\xi_1|M|\xi_0\rangle \\ &\leq s + \sqrt{a(1-a)}\left(\langle\xi_0|M|\xi_0\rangle + \langle\xi_1|M|\xi_1\rangle\right) \\ &\leq (1 + 2\sqrt{a(1-a)})s. \end{aligned} \tag{7.5}$$

where the first inequality follows from M being PSD, i.e., since $(\langle\xi_0| - \langle\xi_1|)M(|\xi_0\rangle - |\xi_1\rangle) \geq 0$ implies $\langle\xi_0|M|\xi_0\rangle + \langle\xi_1|M|\xi_1\rangle \geq \langle\xi_0|M|\xi_1\rangle + \langle\xi_1|M|\xi_0\rangle$. By (7.3) and (7.4), we have

$$(1 + 2\sqrt{a(1-a)})s \leq \left(2 - \frac{8}{p(n)}\right)s \leq \left(2 - \frac{8}{p(n)}\right)\left(\frac{1}{2} + \frac{1}{p(n)}\right) \leq 1 - \frac{2}{p(n)}.$$

Note that by a suitable choice of polynomial $p(n)$ and the initial completeness $c = 1 - \exp(-\text{poly}(n))$ of the QMA$^+$(3) protocol of Theorem 5, we obtain a gap of $\Omega(1/p(n))$. To conclude the proof, we apply standard gap amplification using $k = \text{poly}(p(n))$ proofs in almost-QMA$^{\mathbb{R}}(k)$. ◀

We emphasize an important observation following from the above analysis: The "half" gap amplification in Theorem 5 is almost optimal. A larger gap in Theorem 5 by an additive term $1/\text{poly}(n)$ (e.g., if the soundness was at most $1/2 - 1/\text{poly}(n)$) would allow us to completely discard the ℓ_2-sign bias assumption in Theorem 4, showing NEXP = QMA$^{\mathbb{R}}(k)$. This can be easily seen in (7.5), when $s < 1/2 - 1/\text{poly}(n)$, the RHS will be at most $1 - 1/\text{poly}(n)$. It means that $s = 1/2 \pm 1/\text{poly}(n)$ in Theorem 5 is a critical point. In the case that QMA$(k)^{\mathbb{R}} \neq$ NEXP, there is a sharp phase transition.

References

1 Scott Aaronson, Salman Beigi, Andrew Drucker, Bill Fefferman, and Peter Shor. The power of unentanglement. In *Proceedings of the 23rd IEEE Conference on Computational Complexity (CCC)*, pages 223–236, 2008. doi:10.1109/CCC.2008.5.

2 Adriano Barenco, André Berthiaume, David Deutsch, Artur Ekert, Richard Jozsa, and Chiara Macchiavello. Stabilization of quantum computations by symmetrization. *SIAM Journal on Computing*, 26(5), 1997.

3 Roozbeh Bassirian, Bill Fefferman, and Kunal Marwaha. Quantum Merlin-Arthur and Proofs Without Relative Phase. In *Proceedings of the 15th Innovations in Theoretical Computer Science Conference (ITCS)*, volume 287, pages 9:1–9:19, 2024. doi:10.4230/LIPIcs.ITCS.2024.9.

4 Salman Beigi. NP vs QMAlog(2). *Quantum Info. Comput.*, 2010. doi:10.5555/2011438.2011448.

5 J. S. Bell. On the einstein podolsky rosen paradox. *Physics Physique Fizika*, 1, November 1964. doi:10.1103/PhysicsPhysiqueFizika.1.195.

6 Hugue Blier and Alain Tapp. All languages in NP have very short quantum proofs. In *2009 Third International Conference on Quantum, Nano and Micro Technologies*, pages 34–37, 2009. doi:10.1109/icqnm.2009.21.

7 Fernando G. S. L. Brandão, Matthias Christandl, and Jon Yard. Faithful squashed entanglement. *Communications in Mathematical Physics*, 2011. doi:10.1007/s00220-011-1302-1.

8 Fernando G. S. L. Brandao and Aram W. Harrow. Estimating operator norms using covering nets, 2015. arXiv:1509.05065.

9 Fernando G.S.L. Brandão and Aram W. Harrow. Quantum de finetti theorems under local measurements with applications. In *Proceedings of the 45th ACM Symposium on Theory of Computing (STOC)*, 2013. doi:10.1145/2488608.2488718.

10 Matthias Christandl, Robert König, Graeme Mitchison, and Renato Renner. One-and-a-half quantum de finetti theorems. *Communications in mathematical physics*, 273(2):473–498, 2007.

11 John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23, October 1969. doi:10.1103/physrevlett.24.549.

12 Andrew C. Doherty, Pablo A. Parrilo, and Federico M. Spedalieri. Complete family of separability criteria. *Physical Review A*, 69, 2004. doi:10.1103/physreva.69.022308.

13 A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47, May 1935. doi:10.1007/978-3-322-91080-6_6.

14 François Le Gall, Shota Nakagawa, and Harumichi Nishimura. On QMA protocols with two short quantum proofs. *Quantum Info. Comput.*, 2012. doi:10.26421/qic12.7-8-4.

15 Jeongwan Haah, Aram W. Harrow, Zhengfeng Ji, Xiaodi Wu, and Nengkun Yu. Sample-optimal tomography of quantum states. *IEEE Transactions on Information Theory*, 2017.

16 Aram W Harrow. The church of the symmetric subspace. *arXiv preprint*, 2013. arXiv:1308.6595.

17 Aram W. Harrow and Ashley Montanaro. Testing product states, quantum merlin-arthur games and tensor optimization. *J. ACM*, 60(1), February 2013. doi:10.1145/2432622.2432625.

18 Aram W. Harrow, Anand Natarajan, and Xiaodi Wu. An improved semidefinite programming hierarchy for testing entanglement. *Communications in Mathematical Physics*, 2017. doi:10.1007/s00220-017-2859-0.

19 Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, June 2009. doi:10.1103/RevModPhys.81.865.

20 Fernando Granha Jeronimo and Pei Wu. The Power of Unentangled Quantum Proofs with Non-negative Amplitudes. In *Proceedings of the 55th ACM Symposium on Theory of Computing (STOC)*, 2023.

21 Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. MIP*=RE, 2020. doi:10.1145/3485628.

22 Masaru Kada, Harumichi Nishimura, and Tomoyuki Yamakami. The efficiency of quantum identity testing of multiple states. *Journal of Physics A: Mathematical and Theoretical*, 41(39):395309, September 2008. doi:10.1088/1751-8113/41/39/395309.

23 Hirotada Kobayashi, Keiji Matsumoto, and Tomoyuki Yamakami. Quantum merlin-arthur proof systems: Are multiple merlins more helpful to arthur? In *Algorithms and Computation*, 2003. doi:10.1007/978-3-540-24587-2_21.

24 Robert König and Renato Renner. A de Finetti representation for finite symmetric quantum states. *Journal of Mathematical Physics*, 46(12), 2005.

25 Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010. doi:10.5555/1972505.

26 Ryan O'Donnell and John Wright. Efficient quantum tomography. In *Proceedings of the 48th ACM Symposium on Theory of Computing (STOC)*, 2016.

27 Attila Pereszlényi. Multi-prover quantum merlin-arthur proof systems with small gap, 2012. arXiv:1205.2761.

28 Renato Renner. Security of quantum key distribution. *International Journal of Quantum Information*, 2008.

29 Adrian She and Henry Yuen. Unitary property testing lower bounds by polynomials. In *Proceedings of the 14th Innovations in Theoretical Computer Science Conference (ITCS)*, 2023.

30 Yaoyun Shi and Xiaodi Wu. Epsilon-net method for optimizations over separable states. In *Proceedings of the 39th International Colloquium on Automata, Languages and Programming (ICALP)*, 2012. doi:10.1016/j.tcs.2015.03.031.

31 Mehdi Soleimanifar and John Wright. Testing matrix product states. In *Proceedings of the 33rd ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1679–1701, 2022. doi:10.1137/1.9781611977073.68.

32 John Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018. doi:10.1017/9781316848142.